Double ionization and excitation ionization in the Compton scattering of high-energy photons for metastable states of heliumlike ions

A. I. Mikhaĭlov*) and I. A. Mikhaĭlov

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

(Submitted 3 March 1999)
Zh. Éksp. Teor. Fiz. **116**, 1889–1902 (December 1999)

Double ionization and excitation ionization in Compton scattering for heliumlike ions in metastable states are investigated. The electron energy distribution for double ionization and the total cross sections for both processes are calculated. The calculations are carried out in the zeroth order of perturbation theory with respect to electron-electron interaction, using Coulomb wave functions as the first approximation. The resulting equations are valid only in the high-energy nonrelativistic range. It is assumed that $Z \gg 1$, but $\alpha Z \ll 1$ (Z is the charge of the nucleus, and α is the fine-structure constant). © *1999 American Institute of Physics.* [S1063-7761(99)00112-2]

1. INTRODUCTION

The interaction of electrons with photons is described by a one-particle operator. Consequently, in the absorption or scattering of a single photon the ejection of two electrons from an atom (double ionization) or the ejection of a single electron with transition of the second electron to a higher state (excitation ionization) are possible only through the agency of electron-electron interaction. These processes are currently the subject of extensive theoretical and experimental studies aimed at acquiring information about electron correlations in atoms and ions.[1] Researchers are devoting considerable attention to the helium atom and the helium isoelectronic sequence as the simplest many-electron systems. Their application as targets provides a relatively simple means of testing various models of the role of electron-electron interaction. In lieu of experimental data, the comparison of calculations utilizing perturbation theory with calculations based on diverse variational functions affords the possibility of choosing correlated wave functions that best describe the behavior of the system in the investigated process.

A great many papers address the ground state of helium. Processes in metastable helium are important as well, because they can be used to gain information about correlations between shells. The first such calculations using multiparameter variational wave functions were carried out for double ionization associated with the absorption of a photon having an energy ω of the order of several keV (Refs. 2 and 3). However, if $\omega > \eta$ (η is the average momentum of a $1s$ electron, $\eta \simeq 7$ keV[1)] for He) the dominant process of electron-photon interaction is not photoabsorption, but the scattering of a photon by an electron (Compton scattering).[4] Double ionization in Compton scattering (the double Compton effect) in the helium ground $1\,^1S$ state has been investigated in several papers.[5–8] In the present article we investigate double ionization and excitation ionization in the scattering of high-energy photons in the $2\,^1S$ and $2\,^3S$ states of heliumlike ions in the nonrelativistic energy range. The calculations are based on perturbation theory with respect to electron-electron interaction, Coulomb wave functions serving as the zeroth approximation. We assume that $Z \gg 1$, but $\alpha Z \ll 1$ (Z is the charge of the nucleus, and α is the fine-structure constant). The equations derived below can also be used to calculate the double ionization of stable multiply charged ions containing more than two electrons.

2. DERIVATION OF A GENERAL EQUATION FOR THE AMPLITUDE OF THE PROCESS

We consider the asymptotic nonrelativistic range of photon energies $\eta \ll \omega \ll m$ (m is the electron mass).

The scattering of a photon by an electron is graphically represented by three Feynman diagrams (Fig. 1). The graphs in Figs. 1a and 1b are called pole diagrams and describe a two-photon transition through intermediate states. In momentum representation, which we shall use from now on, a vertex with an incoming photon line corresponds to the operator γ_1:

$$\langle \mathbf{f}' | \gamma_1 | \mathbf{f} \rangle = \sqrt{\frac{2\pi\alpha}{\omega_1}} \frac{\mathbf{e}_1}{m} \left(\mathbf{f} + \frac{i}{2} [\boldsymbol{\sigma} \cdot \mathbf{k}_1] \right) \langle \mathbf{f}' | \mathbf{f} + \mathbf{k}_1 \rangle,$$
$$\langle \mathbf{f}' | \mathbf{f} \rangle = (2\pi)^3 \delta(\mathbf{f}' - \mathbf{f}), \tag{1}$$

and a vertex with an outgoing photon line corresponds to the operator γ_2^*:

$$\langle \mathbf{f}' | \gamma_2^* | \mathbf{f} \rangle = \sqrt{\frac{2\pi\alpha}{\omega_2}} \frac{\mathbf{e}_2^*}{m} \left(\mathbf{f} - \frac{i}{2} [\boldsymbol{\sigma} \cdot \mathbf{k}_2] \right) \langle \mathbf{f}' | \mathbf{f} - \mathbf{k}_2 \rangle, \tag{2}$$

where $\mathbf{e}_1$ and $\mathbf{k}_1$ ($\mathbf{e}_2$ and $\mathbf{k}_2$) are the polarization vector and momentum of the absorbed (emitted) photon, respectively. A vertex with two photon lines [Fig. 1c] corresponds to the operator Q:

$$\langle \mathbf{f}' | Q | \mathbf{f} \rangle = \frac{2\pi\alpha}{\sqrt{\omega_2 \omega_1}} \frac{\mathbf{e}_2^* \cdot \mathbf{e}_1}{m} \langle \mathbf{f}' | \mathbf{f} + \mathbf{k} \rangle, \quad \mathbf{k} = \mathbf{k}_1 - \mathbf{k}_2. \tag{3}$$

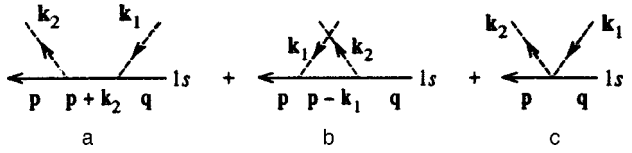

FIG. 1. Feynman diagrams for the scattering of a photon by an electron. Photons are represented by dashed lines, electrons are represented by solid lines, and $\mathbf{q}=\mathbf{p}+\mathbf{k}_2-\mathbf{k}_1$.

For the scattering of a photon by a bound electron the solid lines in Fig. 1 correspond to Coulomb functions.

We now estimate the graphs of Fig. 1 in the investigated energy range. We know[9] that in scattering by a free electron the energy of the incident photon (ω_1) is distributed as follows between the scattered photon (ω_2) and the electron (E_p):

$$\omega_2 \simeq \omega_1 - \frac{\omega_1^2}{m}(1-t), \quad E_p = \omega_1 - \omega_2 \simeq \frac{\omega_1^2}{m}(1-t), \tag{4}$$

where $t=\cos\theta$, and θ is the photon scattering angle.

In scattering by a bound electron the main contribution to the cross section of the process is from the kinematic domain where the momentum q transferred to the nucleus is small ($q\sim\eta$). As a result, the energies of the scattered photon and the ejected electron are also described by Eqs. (4). Accordingly, the momentum of the electron is $p\sim\omega_1$, for $\omega_1 \gg \eta$ the wave function of the ionized electron can be interpreted as a plane wave, and the Coulomb Green's function can be replaced by the free Green's function, because the intermediate electron has high energy E' and momentum p':

$$E' = \omega_1 + E_{1s} \approx \omega_1 \gg I, \quad p' \gg \eta, \quad I = \eta^2/2m, \quad \eta = m\alpha Z.$$

Omitting spin terms in the operators γ_i, we obtain the following estimate for the pole diagrams:

$$F_a \simeq \langle \mathbf{p}|\gamma_2^* G \gamma_1|1s\rangle \sim \frac{\alpha}{\omega_1}\frac{\mathbf{e}_2^*\mathbf{p}}{m}\frac{1}{E'-E_p}\frac{\mathbf{e}_1\mathbf{p}}{m}\langle \mathbf{q}|1s\rangle$$

$$\sim \frac{\alpha}{m^2}\langle \mathbf{q}|1s\rangle, \tag{5}$$

$$\mathbf{q}=\mathbf{p}+\mathbf{k}_2-\mathbf{k}_1,$$

$$F_b \simeq \langle \mathbf{p}|\gamma_1 G \gamma_2^*|1s\rangle \sim \frac{\alpha}{m^2}\langle \mathbf{q}|1s\rangle. \tag{6}$$

The spin terms of the operators γ_i provide the same contribution, since $p\sim\omega_2\sim\omega_1$. The contact diagram in Fig. 1c, calculated under the same conditions, is of the order of magnitude

$$F_c \simeq \langle \mathbf{p}|Q|1s\rangle \sim \frac{\alpha}{m\omega_1}\langle \mathbf{q}|1s\rangle. \tag{7}$$

Comparing Eqs. (5)–(7),

$$\frac{F_a}{F_c} \sim \frac{F_b}{F_c} \sim \frac{\omega_1}{m}, \tag{8}$$

we see that in the energy range $I \ll \omega_1 \ll m$ the pole diagrams are small and can be discarded. The entire discussion thus far is also valid for the double Compton effect (double ionization in photon scattering).[8]

Inasmuch as the operator Q is a one-electron operator, the amplitude of the two-electron transition with allowance for the symmetry properties of the wave functions can be written in the form

$$M^{++} = \langle \Psi_f(1,2)|Q(1)+Q(2)|\Psi_i(1,2)\rangle$$
$$= 2\langle \Psi_f(1,2)|Q(1)|\Psi_i(1,2)\rangle. \tag{9}$$

Here the numbers in parentheses are abbreviated symbols for the variables of integration associated with electrons 1 and 2, and Ψ_i (Ψ_f) is the two-electron wave function of the initial (final) state.

In the nonrelativistic approximation the wave functions can be represented by the product of a spatial function and a spin function. The operator Q does not contain spin matrices, so that the spin part of the wave function in Eq. (9) is invariant and can be dropped from further consideration. Since the total wave function must be antisymmetric with respect to permutation of electrons, the character of the symmetry of the spatial function, being governed by the total electron spin, is left unchanged.

We seek the wave functions $\Psi_{i,f}$ in the lowest perturbation order with respect to electron-electron interaction, using the Coulomb functions $\Psi_{i,f}^{(0)}$ as the zeroth approximation:

$$\Psi_i = \Psi_i^{(0)} + G_i^{(2)}(1-P_i)V_{12}\Psi_i^{(0)},$$

$$\Psi_f = \Psi_f^{(0)} + G_f^{(2)}V_{12}\Psi_f^{(0)}. \tag{10}$$

Here V_{12} is the electron-electron interaction operator [in coordinate representation $V_{12}=\alpha/|\mathbf{r}_1-\mathbf{r}_2|$, $P_i=|\Psi_i^{(0)}\rangle\langle\Psi_i^{(0)}|$ is the projection operator, $G_{i,f}^{(2)}$ is the Green's function for two noninteracting electrons situated in the field of the nucleus Z:

$$G_{i,f}^{(2)} = (E_{i,f}^{(0)} - H_1 - H_2 + i0)^{-1},$$
$$E_i^{(0)} = E_{1s} + E_{2s}, \quad E_f^{(0)} = E_{p_1} + E_{p_2}, \tag{11}$$

and H_1 and H_2 are the one-particle Coulomb Hamiltonians acting on the one-electron functions $\psi(i)$:

$$H_i\psi_{n\swarrow}(i) = E_{n\swarrow}\psi_{n\swarrow}(i), \quad H_i\psi_p(i) = E_p\psi_p(i).$$

Here the subscripts $n\swarrow$ refer to the discrete state, and the subscript p refers to the state of the continuous spectrum. The energy $E_f^{(0)}$ in Eqs. (11) is written for the case in which both electrons are free in the final state. The normalized zeroth-approximation functions have the form

$$\Psi_i^{(0)} = \frac{1}{\sqrt{2}}\{\psi_{1s}(1)\psi_{2s}(2) \pm \psi_{1s}(2)\psi_{2s}(1)\}, \tag{12}$$

$$\Psi_f^{(0)} = \frac{1}{\sqrt{2}}\{\psi_{p_1}(1)\psi_{p_2}(2) \pm \psi_{p_1}(2)\psi_{p_2}(1)\}, \tag{13}$$

where p_1 and p_2 are the ionized electron momenta, and the sign $+(-)$ refers to the singlet (triplet) state. Since the sym-

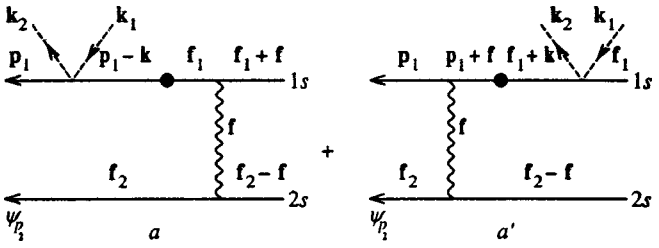

FIG. 2. Feynman amplitudes for double ionization in Compton scattering, taking into account electron-electron interaction in the initial (a) and final (a′) states. A wavy line represents electron-electron interaction (exchange by a Coulomb photon). A line between two vertices with a large dot in the middle corresponds to the Coulomb Green's function. Energy, but not momentum, is conserved on this line.

metry of the system is fixed in the investigated process, the same sign—either plus or minus—is used in Eqs. (12) and (13).

Substituting Eqs. (10)–(13) into (9), we obtain

$$M^{++}=\{\langle\psi_{p_1}\psi_{p_2}|QG_c(E_i)V_{12}+V_{12}G_c(E_f)Q|\psi_{1s}\psi_{2s}\rangle \pm(1s\leftrightarrow 2s)\}\pm\{p_1\leftrightarrow p_2\}. \tag{14}$$

Here $G_c(E)=(E-H_1)^{-1}$ is the one-particle Coulomb Green's function with energy E,

$$E_i=E_i^{(0)}-E_{p_2}=-\frac{5}{4}I-E_{p_2},$$

$$E_f=E_f^{(0)}-E_{2s}=\omega-I,\quad \omega=\omega_1-\omega_2. \tag{15}$$

Equation (15) takes into account the law of conservation of energy:

$$E_{p_1}+E_{p_2}=\omega+E_{1s}+E_{2s}=\omega-\frac{5}{4}I. \tag{16}$$

The terms occurring in the amplitude expression (14) and containing the Green's function $G_c(E_i)$ take into account electron-electron interaction in the initial state, whereas terms containing $G_c(E_f)$ take into account the same interaction in the final state.

The first two terms in Eq. (14) are graphically represented in Fig. 2. The momenta are distributed on the lines for the case in which ψ_{p_1} can be replaced by a plane wave.

It is a well-known fact[10,11] that the energy of the absorbed photon in double photoionization (double photoeffect) is distributed extremely nonuniformly between electrons: one of the electrons carries away almost all the energy, while the other is a low-energy (slow) electron. In the double Compton effect the energy of the incident photon is shared by three particles (two electrons and a photon), so that both electrons in the final state can be slow. As in photoionization, however, the most probable situation is when one electron is slow, and the second acquires large energy in accordance with the laws of kinematics (4) for Compton scattering by a free electron.[8] Consequently, the main contribution to the cross section is from two extreme ranges of the electron energy spectrum: $E_{p_1}\gg E_{p_2}$ for the graphs in Figs. 2a and 2a′ and $E_{p_2}\gg E_{p_1}$ for the same graphs with the momenta p_1 and p_2 transposed. In determining the total cross section it is necessary to take into account the identity of the photoelectrons and to introduce a factor of 1/2. However, it is sufficient consider only one extreme range, say $E_{p_1}\gg E_{p_2}$, and not to divide the cross section by 2.

Let us compare the amplitude contributions from the graphs in Figs. 2a and 2a′. Inasmuch as these graphs differ only in their upper lines, we estimate the matrix elements involved in the vertex lines. The estimates are given for the case in which $p_1\gg\eta$ and, hence, $|\psi_{p_1}\rangle\approx|\mathbf{p}_1\rangle$. Omitting the common factor for both graphs from the operator Q, we obtain

$$F_a=\int\langle\mathbf{p}_1-\mathbf{k}|G_c(E_i)|\mathbf{f}_1\rangle\langle\mathbf{f}_1+\mathbf{f}|1s\rangle\frac{d\mathbf{f}_1}{(2\pi)^3} \sim\langle\mathbf{q}|G_c(E_i)|1s\rangle=\frac{\langle\mathbf{q}|1s\rangle}{E_i-E_{1s}}\sim\frac{\langle\mathbf{q}|1s\rangle}{\eta\alpha Z}, \tag{17}$$

$$F_{a'}=\int\langle\mathbf{p}_1+\mathbf{f}|\frac{1}{E_f-H_0}|\mathbf{f}_1+\mathbf{k}\rangle\langle\mathbf{f}_1|1s\rangle\frac{d\mathbf{f}_1}{(2\pi)^3} \sim\frac{\langle\mathbf{q}|1s\rangle}{E_f-E_{a'}}, \tag{18}$$

$$E_{a'}=\frac{(\mathbf{p}_1+\mathbf{f})^2}{2m},\quad H_0|\mathbf{f}\rangle=\frac{\mathbf{f}^2}{2m}|\mathbf{f}\rangle.$$

In estimating the diagrams of Figs. 2a and 2a′, we have assumed that $f\sim\eta$ (the integrals over f are saturated for $f\sim\eta$). In Fig. 2(a′) the momentum of the intermediate electron is $p_{a'}\simeq p_1\gg\eta$, and the Coulomb Green's function in Eq. (18) is therefore replaced by the free Green's function. Using Eqs. (15) and (4) with E_{p_1} in place of E_p, we obtain

$$E_f-E_{a'}\simeq E_{p_1}-\left(E_{p_1}+\frac{1}{m}\mathbf{p}_1\mathbf{f}\right)\approx-\frac{1}{m}\mathbf{p}_1\mathbf{f}\sim\alpha Z\omega_1, \tag{19}$$

because $E_f\approx E_{p_1}\approx(\omega_1^2/m)(1-t)$ and $p_1\sim\omega_1$. Substituting Eq. (19) into (18), we have

$$F_{a'}\sim\frac{\langle\mathbf{q}|1s\rangle}{\omega_1\alpha Z},\quad \frac{F_{a'}}{F_a}\sim\frac{\eta}{\omega_1}. \tag{20}$$

If the one-electron final states are transposed ($p_1\leftrightarrow p_2$), the energy E_f (15) does not change, but the energy E_i becomes large: $|E_i|\sim E_{p_1}\sim\omega\gg I$. As a result, graphs with the lines ψ_{p_1} and ψ_{p_2} transposed have the same order of smallness as the graph in Fig. 2a′. All these graphs can be discarded in the investigated photon energy range ($\eta\ll\omega_1\ll m$). The amplitude expression (14) then retains only two terms:

$$M^{++}=\langle\psi_{p_1}\psi_{p_2}|QG_c(E_i)V_{12}|\psi_{1s}\psi_{2s}\rangle \pm(1s\leftrightarrow 2s)=M_a\pm M_b, \tag{21}$$

which correspond to the two Feynman diagrams in Fig. 3.

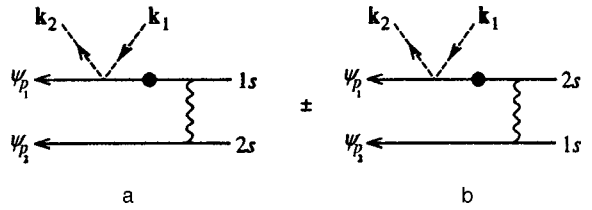

FIG. 3. Feynman diagrams showing the main contribution to the double Compton effect in the $2\,^1S$ and $2\,^3S$ states of heliumlike ions.

3. CALCULATION OF THE AMPLITUDE OF THE DOUBLE COMPTON EFFECT

Here we calculate the amplitude M^{++}. In momentum representation the electron-electron interaction operator V_{12} corresponds to the photon propagator $D(\mathbf{f})$, which is equal to

$$D(\mathbf{f})=\frac{4\pi\alpha}{f^2}. \tag{22}$$

In the amplitude expression we remove the factor $4\pi\alpha$ from the photon propagator (22), and the factor $2\pi\alpha(\omega_1\omega_2)^{-1/2}\mathbf{e}_2^*\mathbf{e}_1/m$ from the operator Q (3), inserting them into the expression for the cross section. The two extreme ranges provide identical contributions to the cross section; hence, we need consider only one of them, say the one in which $p_1\sim\omega_1\gg\eta$ and $p_2\ll p_1$. Using a plane wave as the fast-electron wave function ψ_{p_1}, we obtain

$$M_a=\int\frac{d\mathbf{f}}{(2\pi)^3}\frac{1}{f^2}F_1(\mathbf{p}_1,\mathbf{f})F_2(\mathbf{p}_2,\mathbf{f}), \tag{23}$$

$$F_1(\mathbf{p}_1,\mathbf{f})=\int\frac{d\mathbf{f}_1}{(2\pi)^3}\langle\mathbf{q}|G_c|\mathbf{f}_1\rangle\langle\mathbf{f}_1+\mathbf{f}|1s\rangle=N_1\left(-\frac{\partial}{\partial\eta_1}\right)\times\langle\mathbf{q}|G_cV_{i\eta_1}|-\mathbf{f}\rangle, \tag{24}$$

$$\mathbf{q}=\mathbf{p}_1+\mathbf{k}_2-\mathbf{k}_1,\quad \eta_1=\eta=m\alpha Z,\quad N_1^2=\eta^3/\pi,$$

$$F_2(\mathbf{p}_2,\mathbf{f})=\int\frac{d\mathbf{f}_2}{(2\pi)^3}\langle\psi_{p_2}|\mathbf{f}_2\rangle\langle\mathbf{f}_2-\mathbf{f}|2s\rangle N_2\Gamma_{\eta_2}\left(-\frac{\partial}{\partial\eta_2}\right)\times\langle\psi_{p_2}|V_{i\eta_2}|\mathbf{f}\rangle, \tag{25}$$

$$\Gamma_{\eta_2}=1+\frac{\eta}{2}\frac{\partial}{\partial\eta_2},\quad \eta_2=\frac{\eta}{2},\quad N_2^2=\frac{\eta_2^3}{\pi},$$

$$\langle\mathbf{f}'|V_{i\lambda}|\mathbf{f}\rangle=\frac{4\pi}{(\mathbf{f}'-\mathbf{f})^2+\lambda^2}. \tag{26}$$

The calculations require the wave functions of $1s$ and $2s$ states in momentum representation, which we borrow from Ref. 12. The energy E of an electron in the intermediate state (energy of the Green's function G_c) is identical for the diagrams in Figs. 3a and 3b and coincide with the energy E_i (15):

$$E=-\left(\frac{5}{4}+\varepsilon_2\right)I,\quad \varepsilon_2=\frac{E_{p_2}}{I}=\frac{p_2^2}{\eta^2}. \tag{27}$$

Substituting Eqs. (24) and (25) into (23), we write M_a in the form

$$M_a=\langle\psi_a|G_c|\mathbf{q}\rangle, \tag{28}$$

$$\langle\psi_a|=N_1N_2\Gamma_{\eta_2}\frac{\partial^2}{\partial\eta_2\partial\eta_1}\times\int\frac{d\mathbf{f}}{(2\pi)^3}\frac{1}{f^2}\langle\psi_{p_2}|V_{i\eta_2}|\mathbf{f}\rangle\langle-\mathbf{f}|V_{i\eta_1}. \tag{29}$$

An analogous calculation of the diagram in Fig. 3b gives

$$M_b=\langle\psi_b|G_c|\mathbf{q}\rangle, \tag{30}$$

$$\langle\psi_b|=N_1N_2\Gamma_{\eta_2}\frac{\partial^2}{\partial\eta_2\partial\eta_1}\int\frac{d\mathbf{f}}{(2\pi)^3}\frac{1}{f^2}\times\langle\psi_{p_2}|V_{i\eta_1}|\mathbf{f}\rangle\langle-\mathbf{f}|V_{i\eta_2}. \tag{31}$$

Combining Eqs. (28) and (30), we find the total amplitude M^{++} (21):

$$M^{++}=\langle\psi|G_c|\mathbf{q}\rangle,\quad \psi=\psi_a\pm\psi_b. \tag{32}$$

The sign $+(-)$ is chosen for the singlet (triplet) state.

4. CROSS SECTION OF THE DOUBLE COMPTON EFFECT

The cross section of the double Compton effect, averaged and summed over the photon polarizations, is equal to

$$d\sigma_C^{++}=\frac{(4\pi\alpha)^4}{8m^2}(1+t^2)|M^{++}|^2d\Gamma,\quad t=\cos\theta, \tag{33}$$

$$d\Gamma=\frac{d\mathbf{p}_1d\mathbf{p}_2d\mathbf{k}_2}{\omega_1\omega_2(2\pi)^8}\delta(E_{p_1}+E_{p_2}+\omega_2-\omega_1-E_{1s}-E_{2s}). \tag{34}$$

If G_c is replaced by the free Green's function in Eq. (32), we can readily show that $|M^{++}|^2\sim(q^2+\eta^2)^{-4}$ and, hence, the cross section is governed by the range of low transferred momenta $q\lesssim\eta$. Replacing $d\mathbf{p}_1$ by $d\mathbf{q}$ and integrating over ω_2, we obtain an equation for the phase volume $d\Gamma$ in the nonrelativistic domain:[8]

$$d\Gamma=\frac{d\mathbf{q}d\mathbf{p}_2}{(2\pi)^8}d\Omega_{k_2}. \tag{35}$$

Extending the integral over $\mathbf{q}$ to all space (this operation is permissible, because the contribution of large momenta q to the integral is small), we find

$$S(p_2)=\int\frac{d\mathbf{q}}{(2\pi)^3}|M^{++}|^2=\langle\psi|G_c^2|\psi\rangle=\sum_N\frac{|\langle\psi|N\rangle|^2}{(E-E_N)^2}. \tag{36}$$

The summation in Eq. (36) is carried out over all one-electron Coulomb states of the discrete and continuous spectra, the number N replaces the set of quantum numbers, and the energy E is determined in Eq. (27).

Integrating over the photon emission angle, we obtain

$$d\sigma_C^{++}=\frac{\alpha^2}{\pi}\sigma_C^+S(p_2)d\mathbf{p}_2, \tag{37}$$

$$\sigma_C^+ = 2\sigma_T, \quad \sigma_T = \frac{8}{3}\pi r_e^2, \quad r_e = \frac{\alpha}{m}. \tag{38}$$

Here σ_T is the Thomson cross section of photon scattering by a free electron, σ_C^+ is the single ionization cross section for the scattering of a photon by a two-electron atom (ion).

The exact calculation of $S(p_2)$ poses a difficult problem. It can be simplified considerably on the basis of two essential considerations. First, as will be shown below, the differential cross section decreases rapidly as the slow-electron energy ε_2 increases. The total cross section of the process is determined by the range of low energies $\varepsilon_2 < 1$, and for its determination it is sufficient to calculate $S(p_2)$ in this range. Second, the contribution of the lowest state ($N=1$) in Eq. (36) for small ε_2 is substantially greater than the contribution from all other states. Within acceptable error limits Eq. (36) can be restricted to the first term of the sum, i.e., we can set

$$S(p_2) \simeq \frac{|\langle \psi_a \pm \psi_b | 1s\rangle|^2}{(E - E_{1s})^2}. \tag{39}$$

Using the expression

$$|1s\rangle = N_1 \left(-\frac{\partial}{\partial \eta} \right) V_{i\eta} |0\rangle \tag{40}$$

for the vector state $|1s\rangle$ and invoking the readily verifiable identity

$$-\frac{\partial}{\partial \eta} V_{i\eta} V_{i\lambda} = -\frac{\partial}{\partial \eta} V_{i\lambda} V_{i\eta} = V_{i(\lambda+\eta)}, \tag{41}$$

we obtain

$$\langle \psi_a | 1s\rangle = N_1^2 N_2 \Gamma_{\eta_2} \left(-\frac{\partial}{\partial \nu} \right) \frac{1}{\nu^2} \langle \psi_{p_2} | V_{i\eta_2} - V_{i(\eta_2+\nu)} | 0\rangle, \tag{42}$$

$$\langle \psi_b | 1s\rangle = N_1^2 N_2 \Gamma_{\eta_2} \left(-\frac{\partial}{\partial \eta_2} \right) \frac{1}{(\eta_2+\eta)^2} \langle \psi_{p_2} | V_{i\eta_1} - V_{i(\eta_2+\nu)} | 0\rangle, \tag{43}$$

where $\nu = 2\eta$, and $|0\rangle$ is the zero-momentum plane wave state. The matrix element in Eqs. (42) and (43) with a Coulomb wave function of the continuous spectrum has been calculated previously:[13,14]

$$\langle \psi_{p_2} | V_{i\lambda} | 0\rangle = 4\pi N_{p_2} \frac{(\lambda + ip_2)^{i\xi_2 - 1}}{(\lambda - ip_2)^{i\xi_2 + 1}},$$

$$N_{p_2} \approx \sqrt{2\pi\xi_2}, \quad \xi_2 = \frac{\eta}{p_2}. \tag{44}$$

Substituting Eq. (44) into (42) and (43) and performing the differentiation operations, we arrive at the expressions

$$\langle \psi_a | 1s\rangle = \frac{N_{p_2} N_2}{\eta^2} \left\{ \varphi\left(\frac{1}{2},1\right) - \varphi(a,1) + \frac{1}{2}\varphi\left(\frac{1}{2},2\right) - \frac{5}{2}\varphi(a,2) + 12\varphi(a,3) \right\}, \tag{45}$$

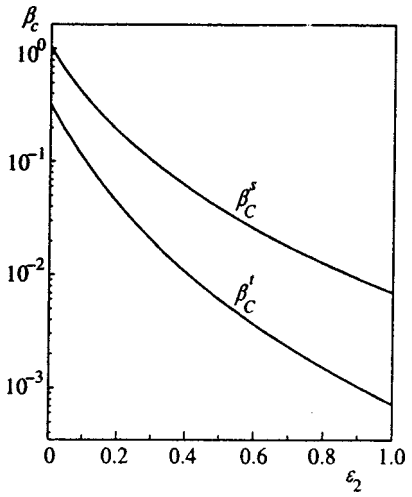

FIG. 4. Energy distributions of slow electrons ($\beta_C = Z^2 d\sigma_C^{++}/d\varepsilon_2 \sigma_C^+$) in the double Compton effect in the $2\,^1S$ (β_C^s) and $2\,^3S$ (β_C^t) states of heliumlike ions.

$$\langle \psi_b | 1s\rangle = \frac{N_{p_2} N_2}{\eta^2} \frac{64}{3} \varphi(a,3), \quad a = \frac{5}{2}, \tag{46}$$

$$\varphi(q,k) = (q^2 + \varepsilon_2)^{-k} \exp\left(-\frac{2}{\sqrt{\varepsilon_2}} \arctan\frac{\sqrt{\varepsilon_2}}{q} \right). \tag{47}$$

Inasmuch as the matrix elements (45) and (46) and, hence, $S(p_2)$ do not depend on the angles, slow electrons are emitted isotropically (the angular distribution of fast electrons is analogous to the angular distribution of electrons in the ordinary Compton effect). Replacing $d\Omega_{p_2}$ by 4π and transforming from $d\mathbf{p}_2$ to $d\varepsilon_2$, we obtain

$$N_{p_2}^2 d\mathbf{p}_2 = (2\pi)^2 \eta^3 d\varepsilon_2. \tag{48}$$

Substituting Eqs. (45) and (46) into (39), substituting Eq. (39) into (37), and taking Eq. (48) into account, we find the energy distribution of slow electrons in the ionization of singlet (s) and triplet (t) states of heliumlike ions:

$$\frac{d\sigma_C^{++}/d\varepsilon_2}{\sigma_C^+} = \frac{\beta_C^{s,t}(\varepsilon_2)}{Z^2}, \quad \beta_C^{s,t}(\varepsilon_2) = 2\left(\frac{\phi^{s,t}(\varepsilon_2)}{1/4 + \varepsilon_2} \right)^2, \tag{49}$$

$$\phi^{s,t}(\varepsilon_2) = \varphi\left(\frac{1}{2},1\right) - \varphi(a,1) + \frac{1}{2}\varphi\left(\frac{1}{2},2\right) - \frac{5}{2}\varphi(a,2) + A^{s,t}\varphi(a,3), \tag{50}$$

$$A^s = \frac{100}{3}, \quad A^t = -\frac{28}{3}. \tag{51}$$

The functions $\beta_C^{s,t}(\varepsilon_2)$ are plotted in Fig. 4.

As ε_2 increases from 0 to 1, the values of β_C decrease more than a hundredfold. In the singlet state both electrons can reside at a single point in space, but this situation is impossible in the triplet state. On the average, the distances between electrons in the singlet state are shorter, and the interaction between them is stronger, than in the triplet state. Consequently, the values of β_C^s are higher than those of β_C^t

for all ε_2. As ε_2 increases, β_C^t decreases much more rapidly than β_C^s, because the role of short distances increases as the transferred energy (ε_2) increases.

The values of β_C for $\varepsilon_2=0$ will be required below. Their exact values are

$$\beta_C^s(0)=1.0424, \quad \beta_C^t(0)=0.3327. \tag{52}$$

Evaluating the integral of (49) over ε_2, we obtain the ratio of the total cross sections of double and single ionization in Compton scattering:

$$R_C=\frac{\sigma_C^{++}}{\sigma_C^{+}}=\frac{B_C^{s,t}}{Z^2}, \quad B_C^s=0.132, \quad B_C^t=0.033. \tag{53}$$

For scattering in the ground state of a two-electron atom we have $R_C=0.048/Z^2$ (Refs. 8 and 15), which is equal to 1.2% for $Z=2$. The recent experimental value of the given quantity for a helium atom, measured at $\omega=97.8$ keV, is equal to $(0.98\pm0.09)\%$ (Ref. 16).

5. EXCITATION IONIZATION IN COMPTON SCATTERING

Like double ionization, excitation ionization into states with $n>2$ is a two-electron process and in the case of Compton scattering is determined entirely by electron-electron interaction. The amplitude $M^{+}*$ of this process can be represented by Feynman diagrams (Fig. 3), except that now the slow-electron wave function ψ_{p_2} must be replaced by the bound-state wave function $\psi_{n\not m}$. Making use of the relation[14]

$$\langle \psi_{n\not m}|V_{i\lambda}|0\rangle = 4\pi N_n \frac{(\lambda-\eta_n)^{n-1}}{(\lambda+\eta_n)^{n+1}}\delta_{\not 0}\delta_{m0},$$

$$N_n^2=\frac{\eta_n^3}{\pi}, \quad \eta_n=\frac{\eta}{n}, \tag{54}$$

and comparing Eq. (54) with (44), we infer that the amplitude $M^{+}*$ is obtained from the amplitude M^{++} by the substitutions $N_{p_2}\to N_n$ and $p_2\to i\eta_n$. The equation for the ionization excitation cross section then follows from Eq. (37) if the factor $d\mathbf{p}_2/(2\pi)^3$ is discarded, and $S(p_2)$ is replaced by

$$S'(n)=\frac{N_n^2}{N_{p_2}^2}S(p_2)\bigg|_{p_2\to i\eta_n}. \tag{55}$$

This cross section has the form

$$\frac{\sigma_C^{+}*(n)}{\sigma_C^{+}}=8\pi^2\alpha^2 S'(n)=\frac{Q_C^{s,t}(n)}{Z^2}, \tag{56}$$

$$Q_C^{s,t}(n)=\frac{2}{n^3}\beta_C^{s,t}(\varepsilon_2\to -n^{-2})=\frac{4}{n^3}\left(\frac{\phi_n^{s,t}}{2^{-2}-n^{-2}}\right)^2. \tag{57}$$

The function $\phi_n^{s,t}$ is given by Eq. (50), in which it is now required to make the substitution

$$\varphi(q,k)\to\varphi_n(q,k)=\frac{(q-n^{-1})^{n-k}}{(q+n^{-1})^{n+k}}. \tag{58}$$

The values of $Q_C^{s,t}$ for various n are given in Table I. It is evident from Eq. (54) that only s states are excited under the stated assumptions. The probability of excitation of states with a different angular momentum is substantially lower. It can be determined by including higher-order terms of the expansion in Eq. (36).

TABLE I. Values of $Q_C^{s,t}$ for various n.

n	$Q_C^s(n)$	$Q_C^t(n)$	n	$Q_C^s(n)$	$Q_C^t(n)$
3	0.3953	0.1602	8	4.889(−3)	1.608(−3)
4	0.0731	0.0265	9	3.301(−3)	1.079(−3)
5	2.728(−2)*	0.943(−2)	10	2.341(−3)	0.762(−3)
6	1.347(−2)	0.454(−2)	20	2.681(−4)	0.860(−4)
7	7.736(−3)	2.569(−3)	100	2.087(−6)	0.666(−6)

*$A(-n)\equiv A\cdot 10^{-n}$.

It follows from Eq. (57) that the following approximate equality holds for large n:

$$Q_C^{s,t}(n\gg 1)\simeq\frac{2}{n^3}\beta_C^{s,t}(0). \tag{59}$$

Substituting the values (52) into this equation, for $n=10$ and $n=100$ we obtain

$$Q_C^s(10)\simeq 2.085\cdot 10^{-3}, \quad Q_C^t(10)\simeq 0.665\cdot 10^{-3},$$

$$Q_C^s(100)\simeq 2.085\cdot 10^{-6}, \quad Q_C^t(100)\simeq 0.665\cdot 10^{-6}.$$

Comparing these numbers with tabulated values, we note that the discrepancies are approximately 10% for $n=10$, whereas far better agreement is obtained for $n=100$.

The authors are indebted to M. Ya. Amus'ya for suggesting the topic of the paper.

APPENDIX

Here we obtain numerical estimates of the discarded terms in Eq. (39). The exact value of $S(p_2)$ (36) differs from the approximate value (39) by the amount

$$\Delta S=\sum_{N=2}^{\infty}\frac{\langle\psi|N\rangle\langle N|\psi\rangle}{(E-E_N)^2}\approx\frac{1}{E^2}(\langle\psi|\psi\rangle-\langle\psi|1s\rangle^2). \tag{A1}$$

Since $E=-I(5/4+\varepsilon_2)$ and $E_N=-I/N^2$ (or $E_N=E_p>0$), when $(E-E_N)^2$ is replaced by E^2, the contribution of the discrete spectrum to ΔS decreases slightly, and the contribution of the continuous spectrum increases. As a result, the approximate estimate (A1) is a little too high; in its derivation we have invoked the completeness condition

$$\sum_{N=1}^{\infty}|N\rangle\langle N|=1.$$

Taking Eq. (32) into account, we have

$$\langle\psi|\psi\rangle=\langle\psi_a|\psi_a\rangle+\langle\psi_b|\psi_b\rangle\pm 2\langle\psi_a|\psi_b\rangle. \tag{A2}$$

Making use of Eq. (29), we obtain

$$\langle\psi_a|\psi_a\rangle = N^2 DD' \int \frac{d\mathbf{f}d\mathbf{f}'}{(2\pi)^6}\frac{1}{f^2} \times\langle\psi_{p_2}|V_{i\eta_2}|\mathbf{f}\rangle\frac{8\pi\nu}{[(\mathbf{f}-\mathbf{f}')^2+\nu^2]^2} \times\langle\mathbf{f}'|V_{i\eta_2'}|\psi_{p_2}\rangle\frac{1}{f'^2},$$

$$N=N_1N_2, \quad D=\Gamma_{\eta_2}\frac{\partial}{\partial\eta_2}, \quad D'=\Gamma_{\eta_2'}\frac{\partial}{\partial\eta_2'}, \quad \nu=2\eta. \tag{A3}$$

The calculation of (A3) is greatly facilitated by making the substitution $[(\mathbf{f}-\mathbf{f}')^2+\nu^2]^2\to(f^2+\nu^2)(f'^2+\nu^2)$. As a result, the integrals are uncoupled, and the ensuing calculations become elementary:

$$\langle\psi_a|\psi_a\rangle \approx \frac{N^2\nu}{2\pi}\left(D\int\frac{d\mathbf{f}}{(2\pi)^3}\langle\psi_{p_2}|V_{i\eta_2}|\mathbf{f}\rangle\frac{1}{f^2}\langle\mathbf{f}|V_{i\nu}|0\rangle\right)^2$$
$$= \frac{N^2}{2\pi\nu^3}(\Gamma_{\eta_2}\langle\psi_{p_2}|V_{i\eta_2}-V_{i(\eta_2+\nu)}|0\rangle)^2$$
$$= \frac{1}{4p_2}\left[\varphi\left(\frac{1}{2},1\right)-\varphi\left(\frac{5}{2},1\right)+\frac{1}{2}\varphi\left(\frac{1}{2},2\right)+\frac{3}{2}\varphi\left(\frac{5}{2},2\right)\right]^2. \tag{A4}$$

An analysis has shown (see also Ref. 8) that the indicated substitution yields an insignificant overestimation of the matrix element (A3).

The other matrix elements on the right hand side of Eq. (A2) are calculated analogously. The final expressions for them are as follows:

$$\langle\psi_b|\psi_b\rangle = \frac{2}{p_2}\{F^2-F[\varphi(2,2)+6\varphi(2,3)]+2\varphi^2(2,2)\},$$
$$F=\varphi(1,1)-\varphi(2,1), \tag{A5}$$
$$\langle\psi_a|\psi_b\rangle = \frac{8}{9p_2}\left\{\left[\varphi(1,1)-\varphi\left(\frac{5}{2},1\right)\right][\varphi(2,2)-2\varphi(2,3)]+\varphi\left(\frac{5}{2},2\right)\left[\varphi\left(\frac{1}{2},1\right)-\varphi(2,1)+\frac{1}{2}\varphi\left(\frac{1}{2},2\right)+\varphi(2,2)\right]\right\}. \tag{A6}$$

We denote by S_1 the approximate value of (39) for $S(p_2)$. Using Eqs. (A4)–(A6) and (45)–(47), we obtain

$$\left.\frac{\Delta S}{S_1}\right|_{\varepsilon_2=0}<2\%, \quad \left.\frac{\Delta S}{S_1}\right|_{\varepsilon_2=1/4}<8\%, \quad \left.\frac{\Delta S}{S_1}\right|_{\varepsilon_2=3/4}<25\%. \tag{A7}$$

The relative contribution of the terms discarded in Eq. (39) increases as ε_2 increases, but the absolute values of S_1 and ΔS decrease rapidly as ε_2 increases. For example, for $\varepsilon_2=1/4$ the quantities S_1 and ΔS are an order of magnitude smaller, and for $\varepsilon_2=3/4$ they are two orders of magnitude smaller than the corresponding quantities for $\varepsilon_2=0$. Consequently, the replacement of the infinite sum (36) by its single term (39) introduces an error of 10% or less in the total cross section (53).

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

1)The system of units $\hbar=c=1$ is used in this paper.

[1] J. H. McGuire, N. Berrah, R. J. Bartlett *et al.*, J. Phys. B **28**, 913 (1995).
[2] Z. Teng and R. Shakeshaft, Phys. Rev. A **49**, 3597 (1994).
[3] R. C. Forrey, H. R. Sadeghpour, J. D. Baker *et al.*, Phys. Rev. A **51**, 2112 (1995).
[4] J. R. Samson, C. H. Greene, and R. J. Bartlett, Phys. Rev. Lett. **71**, 201 (1993).
[5] L. R. Andersson and J. Burgdörfer, Phys. Rev. Lett. **71**, 50 (1993).
[6] K. Hino, P. M. Bergstrom, and J. H. Macek, Phys. Rev. Lett. **72**, 1620 (1994).
[7] T. Surić, K. Pisk, B. A. Logan, and R. H. Pratt, Phys. Rev. Lett. **73**, 790 (1994).
[8] M. Ya. Amusia and A. I. Mikhailov, Phys. Rev. Lett. **199**, 209 (1995); J. Phys. B **28**, 1723 (1995).
[9] A. I. Akhiezer and V. B. Berestetskii, *Quantum Electrodynamics* (Wiley, New York, 1965).
[10] T. A. Carlson, Phys. Rev. **156**, 142 (1967).
[11] M. Ya. Amusia, E. G. Drukarev, V. G. Gorshkov, and M. P. Kazachkov, J. Phys. B **8**, 1248 (1975).
[12] M. Ya. Amus'ya, A. I. Mikhaĭlov, and I. A. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **113**, 539 (1998) [JETP **86**, 299 (1998)].
[13] A. Nordsieck, Phys. Rev. **93**, 785 (1954).
[14] M. Ya. Amus'ya and A. I. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **111**, 862 (1997) [JETP **84**, 474 (1997)].
[15] T. Surić, K. Pisk, and R. H. Pratt, Phys. Lett. A **211**, 289 (1996).
[16] L. Spielberger, H. Bräuning, A. Muthig *et al.*, Phys. Rev. A **59**, 371 (1999).

Translated by James S. Wood

Radiative transition mechanisms in metal clusters

B. M. Smirnov*)

Institute of High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia

H. Weidele

Laboratorium voor Vaste-Stoffysica en Magnetisme, B-3001, Leuven, Belgium

(Submitted 30 March 1999)
Zh. Éksp. Teor. Fiz. **116**, 1903–1911 (December 1999)

Experimental data on the absorption of radiation by cold lithium, potassium, and silver clusters and on the emission of radiation by hot niobium and tungsten clusters are analyzed within the scope of two interpretive schemes of radiative transitions in clusters. The first scheme comprises plasmon model of light absorption by valence electrons of metal clusters. The second scheme treats radiative transitions in metal clusters as transitions of valence electrons interacting with surrounding electrons and atomic cores. The experimental data exhibit better agreement with the second interpretation. © *1999 American Institute of Physics.* [S1063-7761(99)00212-7]

1. INTRODUCTION

The investigation of radiative transitions involving metal clusters can lend insight into the character of the interaction of electromagnetic waves with atomic systems. In this papers we analyze the mechanism of radiative processes involving clusters on the basis of experimental data. The latter have been obtained, for the most part, by two experimental methods. The first method is based on the concept of photoinduced dissociation,[1] whereby the absorption of a photon causes the cluster to decay and, hence, change its mass. The absorption cross section of a cluster ion is deduced from an analysis of the mass spectrum of its absorption products as a function of the intensity of the incident laser radiation. The absorption cross sections for a number of metal clusters in the visible and ultraviolet regions of the spectrum have been measured on the basis of this method.[2–7]

The second method is based on measurement of the spectral power of radiation from hot clusters.[8,9] The measured wavelength distribution of the radiation power for clusters of certain sizes can be used determine the radiative temperature of the clusters during their relaxation after excitation by laser radiation or as a result of oxidation. Moreover, various spectroscopic measurements provide additional information about individual aspects of radiative processes involving clusters,[10–13] and the details of these processes can be learned from theoretical studies.[14–20] The objective of the present study is to summarize the data on radiative processes involving clusters and to choose a simple concept to describe such processes.

2. TWO INTERPRETATIONS OF ABSORPTION BY METAL CLUSTERS

Regarding a metal cluster as a system of bound atoms, we portray a radiative transition in this system as the result of transition of valence electrons. Choosing different bases for the interaction of valence electrons, we can propose two schemes characterizing the process of absorption by a cluster. In the first scheme interaction between valence electrons is more significant than their interaction with atomic cores, so that radiative transitions are an aftermath of interaction of the radiation field with collective degrees of freedom of the system, which in the given situation are plasmon oscillations of the electronic subsystem. This fact imparts a bell-shaped profile to the absorption spectrum.

In the second case, when the main part of the interaction takes place between valence electrons and their atomic cores, radiative transitions are similar to those occurring in solitary atoms, so that the absorption spectrum consists of individual lines broadened by interaction involving neighboring atoms. The number of broadened lines or bands in the absorption spectrum of metal clusters decreases as the temperature increases (see, e.g., the experiment of Haberland *et al.*[12] for sodium clusters), and the explanation of this behavior in the given scheme entails an attendant increase in the number of new cluster atomic configurations, which create the absorption spectrum. Our goal here is to decide which is the more appropriate interpretation for radiative transitions in a cluster from the standpoint of experimental data.

We note that the absorption spectrum of sodium clusters includes several broadened resonances,[6] lending preference to the second interpretation of interaction of the radiation field with clusters in application to sodium clusters and silver Ag_{11} and Ag_{15} clusters.[2] For lithium, potassium, and silver Ag_9 and Ag_{21} clusters the absorption spectrum is bell-shaped, consistent with either interpretation. Consequently, further analysis of these cases will be needed before we can choose the most suitable mechanism of interaction between a radiation field and clusters in the sense of achieving the best fit with measurement data. Such an analysis follows.

3. CLUSTER ABSORPTION CROSS SECTION

Initially we treat absorption by a large cluster as by a macroscopic particle, using the liquid-drop model for the

TABLE I. Parameters of the absorption cross section for metal clusters.

Cluster	$\hbar\omega_0$, eV	Γ, eV	$\sigma_{\max}/n$, Å^2	$\omega_0\sqrt{3}/\omega_p$	ξ	f
Li^+_{139}	2.92	0.90	4.5	0.64	2.8	0.58
Li^+_{270}	3.06	1.15	4.4	0.68	3.2	0.73
Li^+_{440}	3.17	1.32	6.4	0.70	4.9	1.20
Li^+_{820}	3.21	1.10	5.4	0.71	3.3	0.85
Li^+_{1500}	3.25	1.15	5.5	0.72	3.5	0.91
Li_n (average)	3.1±0.1	1.10±0.12	5.2±0.8	0.69±0.03	3.5±0.8	0.85±0.23
K^+_9	1.93	0.22	2.9	0.79	2.9	0.91
K^+_{21}	1.98	0.16	4.2	0.81	2.9	0.96
K^+_{500}	2.03	0.28	3.5	0.84	4.0	1.40
K^+_{900}	2.05	0.40	2.8	0.84	4.5	1.59
K_n (average)	2.00±0.05	0.26±0.10	3.4±0.6	0.82±0.02	3.6±0.8	1.2±0.3
Ag^+_9	4.02	0.62	1.0	0.82	2.6	0.87
Ag^+_{21}	3.82	0.56	0.9	0.78	2.1	0.64
Ag_n (average)	3.9±0.1	0.59±0.03	0.9±0.1	0.80±0.02	2.4±0.2	0.75±0.16

cluster; in this model the cluster is a spherical particle of radius r, which is small in comparison with the radiation wavelength λ:

$$\lambda \gg r. \tag{1}$$

In this model the density ρ of the given particle is assumed to be the same as in a condensed macroscopic system, so that the number of atoms in the cluster n is related to its radius r by the equation

$$n=\left(\frac{r}{r_{\rm WS}}\right)^3, \quad r_{\rm WS}=\left(\frac{3m}{4\pi\rho}\right)^{1/3}. \tag{2}$$

Here $r_{\rm WS}$ is the Wigner–Seitz radius, m is the mass of a single atom, and ρ is the density of the macroscopic system; this relation shows that the investigated cluster can be extracted from the condensed macroscopic system.

The absorption cross section for the spherical macroscopic particle is expressed in terms of the dielectric permittivity of the particle material $\varepsilon(\omega)=\varepsilon'(\omega)+i\varepsilon''(\omega)$ by the equation[21]

$$\sigma_{\rm abs}(\omega)=\frac{12\pi\omega}{c}\,\frac{\varepsilon''}{(\varepsilon'+2)^2+\varepsilon''^2}r^3=\frac{12\pi\omega}{c}r^3 g(\omega). \tag{3}$$

It follows from this equation that the absorption cross section $\sigma_{\rm abs}$ is of the order of $(r/\lambda)r^2$, i.e., is small in comparison with the transverse cross section of the particle πr^2. Moreover, this quantity is proportional to the cluster volume r^3 or to the number of atoms in the cluster n.

For metal particles interacting with a radiation field we use the Drude–Sommerfeld theory,[22,23] which postulates that metal electrons are similar to a gas of free electrons, so that the permittivity of this electron gas is equal to

$$\varepsilon(\omega)=1-\omega_p^2/\omega^2. \tag{4}$$

Here $\omega_p=(4\pi N_e e^2/m_e)^{1/2}$ is the frequency of plasma—or Langmuir—waves; accordingly, N_e is the density of electrons, e is the electron charge, and m_e is the electron mass. The damping of plasma waves is characterized by the imaginary part ε'' of the permittivity. The condition $\varepsilon'' \ll 1$ in conjunction with Eq. (4) transforms Eq. (3) as follows in the vicinity of resonance:

$$\sigma_{\rm abs}(\omega)=2\pi\frac{\hbar\omega^2}{c^2}r^3\frac{\Gamma}{\hbar^2(\omega-\omega_0)^2+\Gamma^2}$$

$$=\sigma_{\max}\frac{\Gamma^2}{\hbar^2(\omega-\omega_0)^2+\Gamma^2}, \tag{5}$$

where $\omega_0=\omega_p/\sqrt{3}$ is the resonance frequency, $\Gamma=\hbar\omega_0\varepsilon''/6$ is the width of the resonance peak, and $\sigma_{\max}$ is the maximum absorption cross section,

$$\sigma_{\max}=2\pi\frac{\hbar\omega_0^2 r^3}{\Gamma c}. \tag{6}$$

From Eq. (5) we deduce the integral equation

$$\int \sigma_{\rm abs}(\omega)d\omega=\frac{\pi}{2}\,\frac{\sigma_{\max}\Gamma}{\hbar}=\pi^2\frac{\omega_0^2 r^3}{c}=\frac{\pi^2\omega_0^2 r_{\rm WS}^3}{c}n, \tag{7}$$

where n is the number of cluster valence electrons, and the width of the resonance peak is assumed to be relatively small.

Although the given model originally applies to a macroscopic particle and is crude in application to a cluster, it takes into account the character of the interaction of a metal cluster with an electromagnetic wave through valence electrons. The absorption spectrum of a metal cluster is therefore localized in the visible region of the spectrum or close to it, and the far infrared region of the spectrum is nonexistent. In reality, the absorption spectrum of metal clusters can have a more complex form than that described by Eq. (5), and it can contain several resonances. Table I shows the parameters of the absorption cross section of metal clusters for cases in which the spectrum consists of a single resonance. The absorption cross sections have been measured for silver,[2] potassium,[3,4] and lithium.[5] We note that in several cases the absorption spectrum as a function of the photon energy is not a bell-shaped curve. In particular, this is true of certain Ag_n (Ref. 2) and Na_n (Ref. 6) clusters, whose absorption spectra have a more complex form than predicted by plasmon theory.

It follows from Eqs. (3) and (4) that the resonance frequency for plasmon-ty8pe interaction of an electromagnetic wave with a metal cluster is equal to

$$\omega_0=\frac{\omega_p}{\sqrt{3}}=\frac{e}{m_e^{1/2}r_{\mathrm{WS}}^{3/2}}, \tag{8}$$

because the density of valence electrons is $N_e=3/(4\pi r_{\mathrm{WS}}^3)$. It follows from this equation that $\hbar\omega_0=4.5\,\mathrm{eV}$ for large lithium clusters, $\hbar\omega_0=2.4\,\mathrm{eV}$ for large potassium clusters, and $\hbar\omega_0=4.9\,\mathrm{eV}$ for large silver clusters. The Wigner-Seitz radii are 1.65 Å, 2.65 Å, and 1.66 Å for liquid clusters of lithium, potassium, and silver, respectively, at the melting point. Also shown in Table I is the ratio of the observed resonance frequencies ω_0 for cluster absorption to the values obtained from Eq. (8). Although these ratios differ from unity, the large-cluster resonance frequency can be corrected by introducing an effective electron mass m_{eff}, which differs from the free-electron mass m_e. Then, according to the data in Table I, for the average effective mass we have $m_{\mathrm{eff}}=(0.57\pm0.10)m_e$. Consequently, even though the positions of the resonances in the cluster absorption cross section differ from those determined on the basis of Eq. (8), the difference can be eliminated by introducing an effective mass for electrons in the clusters.

The plasmon absorption mechanism yields Eq. (6) for the maximum absorption cross section. We can test the validity of this relation for the measured parameters of the absorption cross section. We introduce the parameter

$$\xi=\frac{1}{2\pi}\frac{\sigma_{\max}\Gamma c}{\hbar\omega_0^2r^3}=\frac{1}{2\pi}\frac{\sigma_{\max}}{n}\frac{\Gamma c}{\hbar\omega_0^2r_{\mathrm{WS}}^3}, \tag{9}$$

which is equal to unity if Eq. (6) is valid. The values of this parameter for metal clusters with a bell-shaped absorption spectrum are given in Table I. It is evident that the parameter ξ deviates farther from unity than can be attributed to the error of measurement of the parameters. This discrepancy implies that the underlying assumptions are false. They are based on the notion of a macroscopic absorption process in Eq. (3) and on the notion of plasmon-type interaction of the electromagnetic wave with cluster valence electrons, which leads to Eq. (4). Hence we infer that the interpretation of electromagnetic wave absorption by a metal cluster as the result of interaction between the wave and a macroscopic plasmon is unsuitable, even in cases of a proper cluster absorption spectrum.

The foregoing analysis implies a partial analogy between the system of valence electrons in a metal cluster and free electrons in a plasma. This analogy suggests that electrons are responsible for the interaction of the given atomic system with an electromagnetic wave. The nature of the radiation from metal clusters as systems of bound atoms with interacting valence electrons can be depicted on the basis of the second scheme of interaction between the radiation field and electrons. We consider clusters whose atoms have a resonance excited state, which is linked to the ground state of the atom by dipole radiative transition. For the cases in question the lowest resonance state of an atom is characterized by the greatest oscillator strength for transition from the ground state of the atom. Constructing a cluster of n such atoms and fixing the positions of the nuclei, we obtain a cluster absorption spectrum consisting of n spectral lines. They are obtained from the resonance atomic spectral line, which is split by interaction in the system. Owing to the vibrational motion of the nuclei in a solid cluster, these lines broaden and partially overlap. As a result, the absorption spectrum of the clusters consists of one or several broad resonances. This form of the absorption spectrum follows from calculations of the cluster absorption cross section.[14–20] From these considerations we infer that the cluster absorption cross section is proportional to the number of atoms in the cluster. The same result is consistent with the macroscopic cluster model (3).

The radiative transitions in metal clusters for the given model are similar to radiative transitions involving resonance-excited atoms. This similarity also follows from an analysis of the emission spectrum of mercury clusters as the transition is made from a single atom to a macroscopic system[10,11] by increasing the number of atoms in the system. The indicated nature of the cluster absorption spectrum corresponds to the temperature dependence of the absorption spectrum of sodium Na_{11}^- clusters.[12]

We now consider the sum rule for a metal cluster, making use of the fact that for fixed nuclei the cluster absorption spectrum in the investigated spectral region consists of individual lines comparable in number with the number of atoms in the cluster. In the single-atom limit this spectrum is transformed into one or more resonance lines of the given atom. We introduce an effective oscillator strength f associated with a single valence electron, so that the sum of the oscillator strengths of the given spectrum comprising individual lines is equal to nf, where n is the number of valence electrons in the cluster. As a result of motion of the nuclei, the absorption spectrum acquires the form of several broadened resonances, but the sum rule remains the same. Next we consider the case of a bell-shaped absorption spectrum of the kind encountered for Ag, Li, and K clusters, for which data are given in Table I.

We invoke the general equation for the absorption spectrum of an atomic system[24]

$$\sigma_{\mathrm{abs}}(0\to k)=\frac{\pi^2c^2}{\omega^2}\frac{a_\omega}{\tau_{0k}}\frac{g_k}{g_0}=\frac{2\pi^2e^2}{m_ec}f_{0k}g_ka_\omega. \tag{10}$$

Here m_e is the electron mass, ω is the frequency of the investigated electron transition between states 0 (the lowest state) and k (the upper state), g_0 and g_k are the statistical weights of the transition states, τ_{0k} is the radiative lifetime in relation to this transition, a_ω is the frequency distribution function of emitted photons, $\int a_\omega d\omega=1$, and f_{0k} is the oscillator strength for the given transition; the sum rule for the oscillator strengths of dipole radiative transitions for valence electrons in the investigated spectral region, including resonance transitions, has the form

$$\sum_k f_{0k}=nf.$$

For definiteness we consider clusters consisting of atoms with a single valence electron, as is the case for Ag, Li, and

K clusters (see Table I). Assuming that the investigated spectral region contains two dipole resonance transitions of electrons, integrating over frequencies in the vicinity of the cluster resonance, and summing over all resonance transitions, we obtain an integral equation for the cluster absorption cross section:

$$\int \sigma_{\mathrm{abs}}(\omega)d\omega = \frac{2\pi^2 e^2}{m_e c} nf. \tag{11}$$

If the absorption cross section of the clusters is bell-shaped, as in the cases of Ag, Li, and K clusters (see Table I), the integral equation (7) is valid. We then obtain the following equation from Eqs. (7) and (11):

$$f = \frac{1}{4\pi} \frac{\sigma_{\max}\Gamma m_e c}{e^2 n\hbar}. \tag{12}$$

Table I gives the oscillator strengths f for metal clusters at a single valence electron when the absorption cross section is approximated by a bell-shaped curve. The scatter of values for each element is clearly dictated by the error limits of the cited data. On the average, the values of f for each element correspond to the oscillator strengths for the lowest resonance $^2S_{1/2} \rightarrow {}^2P_{1/2}$, $^2P_{3/2}$ transitions of the corresponding atom. These oscillator strengths are equal to[25] 0.74 for the lithium atom, 1.05 for the potassium atom, and 0.77 for the silver atom. The agreement between the cluster and single-atom oscillator strengths within the error limits of the cluster oscillator strengths confirms the validity of the interpretation that treats the cluster absorption spectrum as the result of transformation of the atomic spectral lines under the influence of their interaction. Radiative transitions in clusters can therefore be regarded as radiative transitions of single valence electrons participating in interaction in the cluster, where these transitions are broadened by the motion of the nuclei.

4. RADIATION FROM HOT CLUSTERS

The width of the spectral absorption band of a cluster is governed by the scattering of single electrons in the field of the atomic cores for both mechanisms of interaction between clusters and the radiation field or is attributable to different configurations of the nuclei in the second interpretation. Additional information about this interaction can be obtained from the temperature dependence of the absorption spectrum and the absorption cross section. In particular, the variation of the absorption spectrum for Na_{11}^- clusters[12] as their temperature varies from low to room temperature can be attributed to new configurations of the cluster nuclei at elevated temperatures. In this paper we use data on the absorption of cold lithium, potassium, and silver clusters at close to room temperature (Table I) and data on the absorption cross sections of hot niobium and tungsten clusters,[26] which follow from an experimental study of the evolution of the spectrum of these clusters[8,9] at high temperatures.

In the experiments[8,9] the emission spectra of Nb, Hf, and W clusters were measured after the cluster beams had been irradiated by a laser pulse. The resultant signal was obtained by summing many pulses, which limited the accuracy of the measurements. These measurements were used to determine the emission spectrum at different times after irradiation, and it was approximated by the emission spectrum of a perfectly black particle having a definite temperature. An analysis[26] shows that the cooling of the irradiated cluster is determined by its radiation, so that the rate of change of the radiation intensity can be used to find the absorption cross section for the hot cluster.

TABLE II. Radiation parameters for the lowest atomic states.

Atom	Transition electron shell	Energy range, eV	Transition time, ns
Li	$2p \rightarrow 2s$	1.85	27
K	$4p \rightarrow 4s$	1.61	27
Ag	$5p \rightarrow 5s$	3.6–3.8	7–8
Nb	$4d^3 5s 5p \rightarrow 4d^4 5s$	2.5–2.6	100–1000
Nb	$4d^4 5p \rightarrow 4d^4 5s$	3.1–3.3	8
W	$5d^4 6s 6p \rightarrow 5d^4 6s^2$	2.7–3.8	60–800

The assumption that the absorption cross section is independent of the wavelength yields an absorption cross section per atom of $(5.2 \pm 0.8) \cdot 10^{-18}\,\mathrm{cm}^2$ for tungsten clusters if they emit in the temperature interval $T = 3170 - 3550\,\mathrm{K}$, corresponding to wavelengths $\lambda_{\max} = 0.68 - 0.76\,\mu\mathrm{m}$ for the maximum spectral power of the radiation. In the case of niobium clusters the absorption cross section per atom is equal to $(5.9 \pm 1.0) \cdot 10^{-18}\,\mathrm{cm}^2$ if the clusters emit at temperatures $T = 3200 - 3600\,\mathrm{K}$, corresponding to wavelengths $\lambda_{\max} = 0.67 - 0.75\,\mu\mathrm{m}$ for the maximum spectral power of the radiation. We note the following relation between the radiative temperature T and the radiation wavelength $\lambda_{\max}$ at the maximum spectral power of the radiation: $\lambda_{\max} T = 0.24\,\mathrm{cm} \cdot \mathrm{K}$. It refers to the dependence of the absorption cross section on the radiation frequency $\sigma(\omega) \propto \omega$, which corresponds to $g(\omega) = \mathrm{const}$ in Eq. (3) for the absorption cross section of a small macroscopic particle.

A significant discrepancy is evident in the specific absorption cross sections of cold and hot metal clusters. We analyze this discrepancy on the basis of the second mechanism of interaction between the electronic subsystem and the radiation field, so that the main radiation parameters of the cluster are atomic, which are transformed by interaction in the cluster. Table II shows the parameters of radiative transitions involving the lowest excited states of the atoms comprised in the investigated clusters. If we assume that radiative transition in a niobium cluster begins with the atomic $5p \rightarrow 5s$ transition, we can discern an analogy between niobium and silver clusters. The maximum spectral powers of niobium cluster radiation, which correspond to the investigated cluster temperatures and have been used to determine the specific absorption cross sections of these clusters, refer to photon energies in the interval $\hbar\omega = 1.4 - 1.5\,\mathrm{eV}$. Using Eq. (5) for the absorption cross section of a cluster with the parameters of silver clusters in Table I, we obtain the specific absorption cross section in the investigated range of photon energies $\sigma/n = (5.1 \pm 1.7) \cdot 10^{-18}\,\mathrm{cm}^2$, corresponding to the above-indicated specific absorption cross section of a niobium cluster. Consequently, the small specific absorption

cross section of niobium clusters can be attributed to the low typical photon energies associated with the tail of the absorption spectrum.

The small specific absorption cross section for tungsten clusters can also be identified with low oscillator strengths for transitions to lower excited states, owing to the long radiative lifetimes of these states (Table II). We note that the high temperature associated with the above cluster absorption cross sections correspond to increased widths of the absorption spectrum, because at high temperatures the nuclei in the cluster acquire contributions that are inadmissible at low temperatures. The broadening of the spectrum can be expected to be of the order of the nuclear thermal energy, and since this amount is small in comparison with the low-temperature width of the spectrum, the broadening effect does not fundamentally change the width of the cluster absorption spectrum at high temperatures. It follows from the foregoing analysis, therefore, that the small specific absorption cross sections for niobium and tungsten clusters, determined from the radiation of hot clusters, can be attributed to the location of the emitted photons in the tail of the cluster emission spectrum and to the weaker interaction of these clusters with the radiation field in comparison with alkali and silver clusters.

5. CONCLUSION

Our analysis of the measured parameters associated with the absorption and emission of radiation by metal clusters and the analysis of other papers show that the model of plasma electrons for the electronic subsystem can lead to conflict with experimental data. In contrast, the metal-cluster spectrum model based on an atomic spectrum containing lines broadened by the interaction of valence electrons with surrounding atomic particles is consistent with various experimental data. It must be emphasized that the noted discrepancy between the two models applies to the mathematical description of the process of light absorption by metal clusters and not to the nature of this process, which is the same for both absorption mechanisms. In fact, light is absorbed by valence electrons in both cases, and the parameters of the absorption spectrum and cross section are dictated by the character of the interaction of these electrons with their surrounding electrons and atomic cores. However, in the plasmon absorption model the width of the absorption spectrum and the maximum absorption cross section are bound by a definite relation, but the experimental data do not corroborate this relation. On the other hand, in the case of weak interaction between valence electrons of the cluster atoms a simple relation exists between the cluster oscillator strength integrated over the spectrum and the oscillator strength in the atom. Even though the interaction between valence electrons in the cluster is not weak, this relation holds for metal clusters for which the absorption cross sections have been measured. It is useful to take these considerations into account in analyzing the radiative parameters of systems containing clusters.

This work has received partial support from the Russian Fund for Fundamental Research (RFFI Grant 99-02-16094).

*)E-mail: smirnov@orc.ru

[1] M. L. Alexander, M. A. Johnson, N. E. Levinger, and W. C. Lindinger, Phys. Rev. Lett. **57**, 976 (1986).
[2] J. Tiggesbumker, L. Keller, H. O. Lutz, and K. H. Meiwes-Broer, Chem. Phys. Lett. **190**, 42 (1992).
[3] C. Bréchignac, P. Cahuzac, F. Carlier, and J. Leygnier, Chem. Phys. Lett. **164**, 433 (1989).
[4] C. Bréchignac, P. Cahuzac, N. Kebaili, J. Leygnier, and A. Sarfati, Phys. Rev. Lett. **68**, 3916 (1992).
[5] C. Bréchignac, P. Cahuzac, J. Leygnier, and A. Sarfati, Phys. Rev. Lett. **70**, 2036 (1993).
[6] C. Bréchignac, P. Cahuzac, F. Carlier, M. de Frutos, and J. Leygnier, Chem. Phys. Lett. **164**, 433 (1989).
[7] H. Fallgren and T. P. Martin, Chem. Phys. Lett. **168**, 223 (1990).
[8] U. Frenzel, U. Kalmbach, D. Kreisle, and E. Recknagel, Surf. Rev. Lett. **3**, 505 (1996).
[9] U. Frenzel, U. Hammer, H. Westje, and D. Kreisle, Z. Phys. D: At., Mol. Clusters **40**, 108 (1997).
[10] H. Haberland, B. von Issendorff, J. Yufeng, and T. Kolar, Phys. Rev. Lett. **69**, 3212 (1992).
[11] C. Ellert, M. Schmidt, C. Schmitt, T. Reiners, and H. Haberland, Phys. Rev. Lett. **75**, 1731 (1995).
[12] H. Haberland, B. von Issendorff, J. Yufeng, T. Kolar, and G. Thanner, Z. Phys. D: At., Mol. Clusters **26**, 8 (1993).
[13] H. Haberland and B. von Issendorff, Phys. Rev. Lett. **76**, 1445 (1996).
[14] W. Ekardt, Phys. Rev. B **31**, 6360 (1985).
[15] V. Kresin, Phys. Rev. B **39**, 3042 (1989); **40**, 12508 (1989); **42**, 3247 (1990).
[16] C. Yannouleas, R. A. Broglia, M. Brack, and P. F. Bortignon, Phys. Rev. Lett. **63**, 255 (1989).
[17] L. Serra, F. Garcia, M. Barranco, J. Navarro, C. Balbas, and A. Mananes, Phys. Rev. B **39**, 8247 (1989).
[18] V. Bonačić-Koutecký, P. Fantucci, and J. Koutecký, Chem. Phys. Lett. **93**, 3802 (1990).
[19] W. Ekardt and Z. Penzar, Phys. Rev. B **43**, 1322 (1991).
[20] K. Selby, V. Kresin, J. Masui, M. Vollmer, A. Scheidemann, and W. D. Knight, Z. Phys. D: At., Mol. Clusters **19**, 41 (1991).
[21] 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).
[22] C. Kittel, *Introduction to Solid State Physics*, 6th ed. (Wiley, New York, 1986).
[23] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
[24] V. P. Krainov, H. R. Reiss, and B. M. Smirnov, *Radiative Transitions in Atomic Physics* (Wiley, New York, 1987).
[25] A. A. Radzig and B. M. Smirnov, *Reference Data on Atoms, Molecules and Ions* (Springer-Verlag, Berlin-New York, 1985).
[26] B. M. Smirnov and H. Weidele, JETP Lett. **69**, 490 (1999).

Translated by James S. Wood

Influence of the boundary of a medium on the coherent backscattering of light

V. L. Kuz'min

St. Petersburg Institute of Trade and Economics, 194021 St. Petersburg, Russia

V. P. Romanov*)

St. Petersburg State University, 198904 Staryi Petergof, St. Petersburg, Russia

(Submitted 5 May 1999)
Zh. Éksp. Teor. Fiz. **116**, 1912–1928 (December 1999)

The multiple scattering of light from an inhomogeneous medium occupying a half-space is investigated on the basis of the Bethe–Salpeter equation. The latter is integrated over the spatial variables to obtain an identity having the significance of the energy balance of the incident and scattered radiations. This relation is then used to derive a length parameter that plays the role of the Milne interpolation length. The use of this parameter in the method of mirror images for describing the shape of the coherent backscattering peak in isotropic single scattering yields results in almost perfect agreement with the predictions of the Milne theory. The application of the given approach for an anisotropic single-scattering diagram yields quantitative agreement of the theory with experiments on the angular dependence of coherent backscattering. The new approach is generalized to an electromagnetic (vector) field, and backscattering polarization effects are investigated. © *1999 American Institute of Physics.* [S1063-7761(99)00312-1]

1. INTRODUCTION

The study of multiple scattering of light in strongly inhomogeneous media has led to the discovery of a whole series of phenomena caused by coherence and correlation effects (see Refs. 1–5). Work continues today on coherent backscattering,[6–9] angular and frequency correlations between reflected and transmitted light,[10,11] and memory and universality effects in the behavior of the time correlation function.[12] Wave propagation in strongly inhomogeneous media especially began to elicit major interest once it had been established that techniques of correlation spectroscopy and, specifically, techniques of the spectroscopy of diffusing photon-density waves[1,2,13] could be used to solve the problem of imaging ordered structures and macroscopic inhomogeneities in opaque media.[14–17] In particular, this realization launched a new trend in medical technology: diagnostics utilizing visible and infrared radiation.[18]

To adequately reproduce the structure of opaque media and achieve a more accurate description of coherence and correlation effects, a concerted effort is underway at the present time to develop a theory of multiple scattering. A major problem in this area is the systematic treatment of boundaries and interfaces.[19–23] The method of mirror images fits quite naturally into the diffusion approximation. However, because the diffusion approximation is valid far from any boundary, a certain leeway is encountered in the choice of an effective boundary. Alternative choices of this kind are discussed in Ref. 19. The position of the effective boundary is usually determined by means of the Milne interpolation length, which is obtained from the exact solution for the scalar field in the model of point scatterers. This choice was first adopted in Ref. 23. However, the method of mirror images itself does not require that the effective boundary coincide with the Milne interpolation length.

In this paper we propose a self-consistent method for choosing the effective boundary in the description of multiple scattering from a half-space. The solution is constructed for scalar and vector fields. For the scalar field in isotropic single scattering we show that in describing the coherent backscattering peak the result obtained by the approach developed here essentially coincides with the result obtained from the exact Milne solution. The application of the given approach to systems with an anisotropic angular scattering function yields good agreement of the theory with experimental coherent backscattering data.

The paper is structured as follows. In Sec. 2 general expressions are given for the intensity of multiply scattered radiation due to the contributions of ladder and cyclic diagrams. The scalar field is discussed in Sec. 3, including a general analysis of the system of equations with zeroth- and first-degree terms included in the expansion in Legendre polynomials. In Sec. 4 an identity having the significance of an energy balance equation for the incident and scattered light is derived for the scalar field by integrating the Bethe–Salpeter equation. A length parameter analogous to the Milne interpolation length is obtained self-consistently from this equation. It is used to calculate the angular dependence of coherent backscattering, and the results of the calculations are compared with experiment. In Sec. 5 the new approach is extended to the case of an electromagnetic field, and the polarized and depolarized components of coherent backscattering are calculated.

2. MULTIPLE LIGHT SCATTERING IN A STRONGLY INHOMOGENEOUS MEDIUM

We consider the propagation of light in a random dielectric medium. Assuming that variations of the inhomogeneities of the medium are negligible during wave propagation,[24] we can write the wave equation for the spectral field component $\mathbf{E}(\mathbf{r},\omega)$

$$\operatorname{curl}\operatorname{curl}\mathbf{E}(\mathbf{r},\omega)-\left(\frac{\omega}{c}\right)^2\varepsilon(\omega)\mathbf{E}(\mathbf{r},\omega)=\left(\frac{\omega}{c}\right)^2\delta\varepsilon(\mathbf{r},\omega)\mathbf{E}(\mathbf{r},\omega), \tag{2.1}$$

where $\delta\varepsilon(\mathbf{r},\omega)=\varepsilon(\mathbf{r},\omega)-\varepsilon(\omega)$ is the fluctuation of the dielectric permittivity, $\varepsilon(\omega)=\langle\varepsilon(\mathbf{r},\omega)\rangle$ is the average permittivity at the frequency ω, c is the light velocity in empty space. From now on we shall drop the argument ω for brevity. The intensity of multiple scattering of radiation at a point $\mathbf{r}$ can be written in the form

$$\langle|\delta E_s(\mathbf{r})|^2\rangle=\int d\mathbf{r}_1 d\mathbf{r}_1' d\mathbf{r}_2 d\mathbf{r}_2' A_{sj}\times(\mathbf{r}-\mathbf{r}_2)A_{sl}^*(\mathbf{r}-\mathbf{r}_2')\Gamma_{jl,ii}(\mathbf{r}_2,\mathbf{r}_2',\mathbf{r}_1,\mathbf{r}_1')\times\langle E_i(\mathbf{r}_1)\rangle\langle E_i^*(\mathbf{r}_1')\rangle. \tag{2.2}$$

Here $\delta\mathbf{E}_s(\mathbf{r})=\mathbf{E}(\mathbf{r})-\langle\mathbf{E}(\mathbf{r})\rangle$ is the fluctuation of the field in the medium, $\langle\mathbf{E}(\mathbf{r})\rangle$ is the average incident field, which we assumed to be a monochromatic plane wave, $\langle\mathbf{E}(\mathbf{r})\rangle=\mathbf{E}\exp(i\mathbf{k}_i\cdot\mathbf{r})$, $\mathbf{k}_i$ is the incident wave vector, and $\hat{A}(\mathbf{r}-\mathbf{r}_2)$ is the Green's function of the wave equation. In the far field at large distances $\mathbf{r}$ from the scattering volume this function has the form

$$\hat{A}(\mathbf{r}-\mathbf{r}_2)=\frac{1}{r}\left(\hat{I}-\frac{\mathbf{k}_s\cdot\mathbf{k}_s}{k^2}\right)e^{-i\mathbf{k}_s\cdot\mathbf{r}_2}, \tag{2.3}$$

where $\hat{I}$ is the unit matrix, $\mathbf{k}_s=k\mathbf{r}/r$ is the unit vector of the scattered wave, $k=\varepsilon^{1/2}k_0$, $k_0=\omega/c=2\pi/\lambda$, and λ is the light wavelength. Summation over repeated subscripts is tacitly understood from now on, except for the subscripts i and s, which designate the polarizations of the incident and scattered light, respectively.

The function $\hat{\Gamma}(\mathbf{r}_2,\mathbf{r}_2',\mathbf{r}_1,\mathbf{r}_1')$, which characterizes the scattered field, is represented in general by a power series in the permittivity fluctuations $\delta\varepsilon$. This series is summed and leads to the Bethe–Salpeter equation. In the weak scattering approximation $\lambda\ll l$ (l is the extinction length) this equation assumes the form

$$\Gamma_{lj,mn}(\mathbf{r}_2,\mathbf{r}_2',\mathbf{r}_1,\mathbf{r}_1')=k_0^4G(\mathbf{r}_2-\mathbf{r}_2')\Big[\delta(\mathbf{r}_2-\mathbf{r}_1)\delta(\mathbf{r}_2'-\mathbf{r}_1')\delta_{lm}\delta_{jn}+\int d\mathbf{r}_3 d\mathbf{r}_3' A_{la}(\mathbf{r}_2-\mathbf{r}_3)A_{jb}^*(\mathbf{r}_2'-\mathbf{r}_3')\times\Gamma_{ab,mn}(\mathbf{r}_3,\mathbf{r}_3',\mathbf{r}_1,\mathbf{r}_1')\Big], \tag{2.4}$$

where

$$G(\mathbf{r}_2-\mathbf{r}_2')=\frac{1}{(4\pi)^2}\langle\delta\varepsilon(\mathbf{r}_2)\delta\varepsilon(\mathbf{r}_2')\rangle \tag{2.5}$$

is the correlation function of the permittivity fluctuations. Equation (2.4) is written in the ladder approximation on the assumption that the fluctuations $\delta\varepsilon$ are essentially Gaussian. This equation is derived on the assumption of smallness of the parameter $\delta n r_c^3\rho$, where δn is the inhomogeneity of the refractive index of the medium, r_c is the correlation length of the inhomogeneities, or radius of the scatterers, as defined by the function (2.5), and ρ is the density of inhomogeneities, or scatterers, in unit volume, $\rho\sim a^{-3}$ (a is the average distance between scatterers).

In the function $\hat{\Gamma}(\mathbf{r}_2,\mathbf{r}_2',\mathbf{r}_1,\mathbf{r}_1')$ it is useful to transform to ''center of gravity'' coordinates $\mathbf{R}_j=(\mathbf{r}_j+\mathbf{r}_j')/2$ and relative coordinates $\mathbf{r}_j''=\mathbf{r}_j-\mathbf{r}_j'$. For the function $\tilde{\Gamma}_{lj,mn}(\mathbf{R}_2,\mathbf{R}_2';\mathbf{k}_s,\mathbf{k}_i)$, which represents the Fourier transform in the relative coordinates,

$$\tilde{\Gamma}_{lj,mn}(\mathbf{R}_2,\mathbf{R}_1;\mathbf{k}_s,\mathbf{k}_i)=\int d\mathbf{r}_1''\mathbf{r}_2''\Gamma_{lj,mn}\left(\mathbf{R}_2+\frac{\mathbf{r}_2''}{2},\mathbf{R}_2-\frac{\mathbf{r}_2''}{2},\mathbf{R}_1+\frac{\mathbf{r}_1''}{2},\mathbf{R}_1-\frac{\mathbf{r}_1''}{2}\right)\times\exp(i\mathbf{k}_i\mathbf{r}_1''-i\mathbf{k}_s\dot{\mathbf{r}}_2''), \tag{2.6}$$

Eq. (2.4) can be written in the form

$$\tilde{\Gamma}_{lj,mn}(\mathbf{R}_2,\mathbf{R}_1;\mathbf{k}_s,\mathbf{k}_i)=k_0^4\tilde{G}(\mathbf{k}_i-\mathbf{k}_s)\delta(\mathbf{R}_{21})\delta_{lm}\delta_{jn}+k_0^4\int d\mathbf{R}_3\tilde{G}(\mathbf{k}_{23}-\mathbf{k}_s)\times\Lambda_{lj,ab}(\mathbf{R}_{23})\tilde{\Gamma}_{ab,mn}\times(\mathbf{R}_3,\mathbf{R}_1;\mathbf{k}_{23},\mathbf{k}_i). \tag{2.7}$$

Here

$$\tilde{G}(\mathbf{q})=\int d\mathbf{r}\,G(\mathbf{r})e^{-i\mathbf{q}\cdot\mathbf{r}} \tag{2.8}$$

is the Fourier transform of the correlation function,

$$\Lambda_{lj;ab}(\mathbf{R})=\frac{e^{-R/l}}{R^2}\left(\delta_{la}-\frac{R_lR_a}{R^2}\right)\left(\delta_{jb}-\frac{R_jR_b}{R^2}\right), \tag{2.9}$$

and $\mathbf{k}_{ij}=k\mathbf{R}_{ij}/R_{ij}$ is the wave vector of the wave propagating between points $\mathbf{R}_i$ and $\mathbf{R}_j$.

We now consider scattering from the half-space $z>0$, where z is the Cartesian coordinate directed along the inward normal to the boundary $z=0$ for angles of incidence close to 180° and for normal incidence. In the far field, taking Eq. (2.3) into account, we can write the mean-square field in the form

$$\langle|\delta\mathbf{E}_s(\mathbf{r})|^2\rangle=\frac{1}{r^2}I_{si}(\mathbf{k}_s,\mathbf{k}_i), \tag{2.10}$$

where

$$I_{si}(\mathbf{k}_s,\mathbf{k}_i)=\int d\mathbf{R}_1 d\mathbf{R}_2 \exp\left(-\frac{Z_1+Z_2}{l}\right)\times\left(\hat{I}-\frac{\mathbf{k}_s\cdot\mathbf{k}_s}{k^2}\right)_{sj}\left(\hat{I}-\frac{\mathbf{k}_s\cdot\mathbf{k}_s}{k^2}\right)_{sl}\times\hat{\Gamma}_{jl,ii}(\mathbf{R}_2,\mathbf{R}_1;\mathbf{k}_s,\mathbf{k}_i)|\mathbf{E}_i|^2. \quad (2.11)$$

The axial symmetry of the problem reduces the dependence on $\mathbf{R}_2$ and $\mathbf{R}_1$ to a dependence on the relative transverse variable $\boldsymbol{\rho}_{21}=[(X_2-X_1),(Y_2-Y_1)]$ and the coordinates Z_2 and Z_1.

It is evident from Eq. (2.11) that for scattering problems at normal incidence it is sufficient to consider the function

$$\hat{\Gamma}(\boldsymbol{\rho}_{21},Z_2;\mathbf{k}_s,\mathbf{k}_i)=\int_0^\infty \tilde{\hat{\Gamma}}(\mathbf{R}_2,\mathbf{R}_1;\mathbf{k}_s,\mathbf{k}_i)\times\exp(-Z_1/l)dZ_1. \quad (2.12)$$

According to Eq. (2.7), this function satisfies the equation

$$\Gamma_{lj,mn}(\boldsymbol{\rho}_{21},Z_2;\mathbf{k}_s,\mathbf{k}_i)=k_0^4\tilde{G}(\mathbf{k}_i-\mathbf{k}_s)\,\theta(Z_2)\,\delta(\boldsymbol{\rho}_{21})\times\exp(-Z_2/l)\,\delta_{lm}\delta_{jn}+k_0^4\times\int\tilde{G}(\mathbf{k}_{23}-\mathbf{k}_s)\Lambda_{lj,ab}(\mathbf{R}_{23})\times\Gamma_{ab,mn}(\boldsymbol{\rho}_{31},Z_3;\mathbf{k}_{23},\mathbf{k}_i)d\mathbf{R}_3. \quad (2.13)$$

In the given geometry we have $\mathbf{k}_i=(0,0,k)$ and $\mathbf{k}_s=(k_0\theta_s,0,-k)$, where θ_s is the scattering angle measured relative to the true backward direction. Equation (2.11) now assumes the form

$$I_{si}(\mathbf{k}_s,\mathbf{k}_i)=|\mathbf{E}_i|^2 S\int d\boldsymbol{\rho}_{21}\int_0^\infty dZ_2\Gamma_{ss,ii}(\boldsymbol{\rho}_{21},Z_2;\mathbf{k}_s,\mathbf{k}_i)\times\exp(-Z_2/l), \quad (2.14)$$

where S denotes the illuminated region.

For small angles θ_s, together with the contribution of ladder diagrams, it is also necessary to include an interference term of the form[19]

$$I_{si}^{(c)}(\mathbf{k}_s,\mathbf{k}_i)=|\mathbf{E}_i|^2 S\int d\boldsymbol{\rho}_{21}\int_0^\infty dZ_2[\Gamma_{si,is}(\boldsymbol{\rho}_{21},Z_2;\mathbf{k}_s,\mathbf{k}_i)-\delta(\boldsymbol{\rho}_{21})\,\delta_{si}k_0^4\tilde{G}(\mathbf{k}_s-\mathbf{k}_i)\exp(-Z_2/l)]\times\exp(-Z_2 l^{-1}+ik_0\theta_s(X_2-X_1)). \quad (2.15)$$

Equation (2.15) represents the contribution of cyclic diagrams. In its derivation from the contribution of ladder diagrams it is required, in particular, to permute subscripts and subtract the single-scattering contribution (see, e.g., Ref. 4).

The anisotropy of single scattering makes the function $\hat{\Gamma}(\boldsymbol{\rho},Z_2;\mathbf{k}_s,\mathbf{k}_i)$ dependent on the orientation of the wavevectors. Owing to the structure of Eq. (2.13), we can fix the orientation of the incident wave vector $\mathbf{k}_i$ and consider only the dependence on the orientation of $\mathbf{k}_s$. In the general case this dependence on $\mathbf{k}_s$ can be written as an expansion in spherical harmonics. We confine the discussion to the zeroth- and first-degree Legendre polynomials, which correspond to the diffusion approximation

$$\hat{\Gamma}(\boldsymbol{\rho}_{21},Z_2,\mathbf{k}_s,\mathbf{k}_i)=\frac{1}{4\pi l}[\hat{\gamma}_0(\boldsymbol{\rho}_{21},Z_2)+\hat{\gamma}_n(\boldsymbol{\rho}_{21},Z_2)\times\cos\theta+\hat{\gamma}_t(\boldsymbol{\rho}_{21},Z_2)\cos\theta_t], \quad (2.16)$$

where

$$\cos\theta=(\mathbf{k}_s\cdot\mathbf{e}_3)k^{-1},\quad \cos\theta_t=(\mathbf{k}_s\cdot\boldsymbol{\rho}_{21})k^{-1}\rho_{21}^{-1},$$

and $\mathbf{e}_3$ is the unit vector along the Z axis. Here we have also made use of the fact that $\mathbf{k}_i\|\mathbf{e}_3$.

Invoking the orthogonality property of the Legendre polynomials, we can reduce Eq. (2.13) to a system of integration equations for the functions $\hat{\gamma}_0$, $\hat{\gamma}_n$, and $\hat{\gamma}_t$:

$$\hat{\gamma}_0(\boldsymbol{\rho}_{21},Z_2)=\frac{3}{2}\theta(Z_2)\exp\left(-\frac{Z_2}{l}\right)\delta(\boldsymbol{\rho}_{21})\hat{I}+\frac{3}{8\pi l}\int d\mathbf{R}_3\hat{\Lambda}(R_{23})\times\left[\hat{\gamma}_0(\boldsymbol{\rho}_{31},Z_3)+\left(\frac{\mathbf{k}_{23}\cdot\boldsymbol{\rho}_{31}}{k\rho_{31}}\right)\hat{\gamma}_t(\boldsymbol{\rho}_{31},Z_3)+\left(\frac{\mathbf{k}_{23}\cdot\mathbf{e}_3}{k}\right)\hat{\gamma}_n(\boldsymbol{\rho}_{31},Z_3)\right],$$

$$\hat{\gamma}_t(\boldsymbol{\rho}_{21},Z_2)=\frac{9\mu}{8\pi l}\int d\mathbf{R}_3\frac{\boldsymbol{\rho}_{21}\cdot\mathbf{R}_{23}}{\rho_{21}R_{23}}\hat{\Lambda}(R_{23})\times\left[\hat{\gamma}_0(\boldsymbol{\rho}_{31},Z_3)+\left(\frac{\mathbf{k}_{23}\cdot\boldsymbol{\rho}_{31}}{k\rho_{31}}\right)\hat{\gamma}_t(\boldsymbol{\rho}_{31},Z_3)+\left(\frac{\mathbf{k}_{23}\cdot\mathbf{e}_3}{k}\right)\hat{\gamma}_n(\boldsymbol{\rho}_{31},Z_3)\right],$$

$$\hat{\gamma}_n(\boldsymbol{\rho}_{21},Z_2,t)=\frac{9}{2}\mu\,\theta(Z_2)\exp(-Z_2/l)\,\delta(\boldsymbol{\rho}_{21})\hat{I}+\frac{9\mu}{8\pi l}\int d\mathbf{R}_3\frac{\mathbf{R}_{23}\cdot\mathbf{e}_3}{R_{23}}\hat{\Lambda}(R_{23})\times\left[\hat{\gamma}_0(\boldsymbol{\rho}_{31},Z_3,t)+\left(\frac{\mathbf{k}_{23}\cdot\boldsymbol{\rho}_{31}}{k\rho_{31}}\right)\hat{\gamma}_t(\boldsymbol{\rho}_{31},Z_3)+\left(\frac{\mathbf{k}_{23}\cdot\mathbf{e}_3}{k}\right)\hat{\gamma}_n(\boldsymbol{\rho}_{31},Z_3)\right], \quad (2.17)$$

where $\theta(Z)$ is the Heaviside theta function, $\mu=\overline{\cos\theta}$ is the cosine of the scattering angle, averaged over the single-scattering angular diagram. In the derivation of Eqs. (2.17) we have relied on the optical theorem

$$l^{-1}=\frac{2}{3}k_0^4\int d\Omega_s\tilde{G}(\mathbf{k}_s-\mathbf{k}_i). \quad (2.18)$$

We now consider a medium in which light attenuation is attributable to scattering rather than to absorption, i.e., the extinction length l is much shorter than the attenuation length.

The system (2.17) is a generalization of the Milne equation to the case of an electromagnetic field. Here the anisotropy of single scattering is taken into account in the P_1-approximation. The solution of this system of equations can be used to find the function $\hat{\Gamma}(\boldsymbol{\rho}_{21},Z_2,\mathbf{k}_s,\mathbf{k}_i)$ and, hence, to determine the scattering intensity, which is given by Eq. (2.11).

The substitution of the solution for an unbounded medium in place of the function $\hat{\Gamma}$ in Eq. (2.11), of course, yields a divergent expression. The standard procedure used to remove the divergence, by analogy with electrostatic problems, is based on the method of mirror images. The solution of the system of equations (2.17) does not result in divergence, because the boundedness of the medium is taken into account explicitly here. However, this system of equations is complex and difficult to solve, even numerically.[20,25] In contrast with an electromagnetic field, for the scalar case an exact solution exists for a system of point scatterers. Inasmuch as the existence of an exact solution can be used to analyze the degree of validity of approximate approaches, we first discuss the case of a scalar field.

3. SCALAR FIELD

In the scalar analog of the system of equations (2.17) we can readily transform to the Fourier spectrum in transverse variables. We have

$$\begin{aligned}
\tilde{\gamma}_0(\mathbf{Q},Z_2)&=\theta(Z_2)\exp\left(-\frac{Z_2}{l}\right)+\frac{1}{4\pi l}\\
&\quad\times\int_0^\infty dZ_3\{\tilde{\Lambda}(\mathbf{Q},Z_{23})\tilde{\gamma}_0(\mathbf{Q},Z_3)+\tilde{\boldsymbol{\Lambda}}_1(\mathbf{Q},Z_{23})\\
&\quad\times[\tilde{\boldsymbol{\gamma}}_t(\mathbf{Q},Z_3)+\mathbf{e}_3\tilde{\gamma}_n(\mathbf{Q},Z_3)]\},\\
\tilde{\boldsymbol{\gamma}}_t(\mathbf{Q},Z_2)&=\frac{3\mu}{4\pi l}\int_0^\infty dZ_3\{\tilde{\boldsymbol{\Lambda}}_1(\mathbf{Q},Z_{23})\tilde{\gamma}_0(\mathbf{Q},Z_3)\\
&\quad+\hat{\tilde{\Lambda}}_2(\mathbf{Q},Z_{23})[\tilde{\boldsymbol{\gamma}}_t(\mathbf{Q},Z_3)+\mathbf{e}_3\tilde{\gamma}_n(\mathbf{Q},Z_3)]\},\\
\tilde{\gamma}_n(\mathbf{Q},Z_2)&=3\mu\left\{\theta(Z_2)\exp\left(-\frac{Z_2}{l}\right)+\frac{1}{4\pi l}\right.\\
&\quad\times\int_0^\infty dZ_3\mathbf{e}_3\{\tilde{\boldsymbol{\Lambda}}_1(\mathbf{Q},Z_{23})\tilde{\gamma}_0(\mathbf{Q},Z_3)\\
&\quad\left.+\hat{\tilde{\Lambda}}_2(\mathbf{Q},Z_{23})[\tilde{\boldsymbol{\gamma}}_t(\mathbf{Q},Z_3)+\mathbf{e}_3\tilde{\gamma}_n(\mathbf{Q},Z_3)]\}\right\},
\end{aligned}\tag{3.1}$$

From now on we denote quantities with a tilde by the corresponding two-dimensional Fourier transforms

$$f(\boldsymbol{\rho},Z)=\int\frac{d^2\mathbf{Q}}{(2\pi)^2}\tilde{f}(\mathbf{Q},Z)\exp(i\mathbf{Q}\cdot\boldsymbol{\rho})\tag{3.2}$$

of the unknown functions $\gamma_0(\boldsymbol{\rho},Z)$, $\boldsymbol{\gamma}_t(\boldsymbol{\rho},Z)=\boldsymbol{\rho}\gamma_t(\boldsymbol{\rho},Z)/\rho$, $\gamma_n(\boldsymbol{\rho},Z)$ and the integral kernels $\Lambda(R)$, $\boldsymbol{\Lambda}_1(R)=\mathbf{R}\Lambda(R)/R$ and $\hat{\Lambda}_2(R)=\mathbf{R}\cdot\mathbf{R}\Lambda(R)/R^2$. The quantity

$$\Lambda(R)=\exp(-R/l)/R^2\tag{3.3}$$

is the scalar analog of the tensor (2.8).

For convenience we transform to the dimensionless variables $z=Z/l$ and $\mathbf{q}=l\mathbf{Q}$. In the system of equations (3.1) we go over to the Laplace transform in the variable z. We determine the Laplace transforms in the form

$$g_m(\mathbf{q},s)=\int_0^\infty dz\,\tilde{\gamma}_m(\mathbf{q},z)e^{-zs},\quad m=0,n,t.\tag{3.4}$$

Here the function $\tilde{\boldsymbol{\gamma}}_t$ has been parametrized in the form $\tilde{\boldsymbol{\gamma}}_t=\mathbf{q}\gamma_t(\mathbf{q},z)$ by virtue of its orthogonality to the z axis. For the Laplace transforms the system (3.1) assumes the form

$$\begin{aligned}
&[1-p_0(w)]g_0(\mathbf{q},s)-q^2p_1(w)g_t(\mathbf{q},s)\\
&\quad+sp_1(w)g_n(\mathbf{q},s)=a_0(\mathbf{q},s),\\
&3\mu p_1(w)g_0(\mathbf{q},s)\\
&\quad+\left\{1-\frac{3\mu}{2}\frac{s^2[p_0(w)-p_1(w)]-2q^2p_1(w)}{w^2}\right\}g_t(\mathbf{q},s)\\
&\quad+\frac{3\mu s}{w^2}[p_0(w)-3p_1(w)]g_n(\mathbf{q},s)=a_t(\mathbf{q},s),\\
&3\mu sp_1(w)g_0(\mathbf{q},s)-\frac{3\mu q^2s[p_0(w)-3p_1(w)]}{w^2}g_t(\mathbf{q},s)\\
&\quad+\left\{1-\frac{3\mu}{2}\frac{2s^2p_1(w)-q^2[p_0(w)-p_1(w)]}{w^2}\right\}g_n(\mathbf{q},s)\\
&\quad=a_n(\mathbf{q},s),
\end{aligned}\tag{3.5}$$

where

$$w^2=s^2-q^2,\quad p_0(w)=\frac{1}{2w}\ln\frac{1+w}{1-w},\quad p_1(w)=\frac{p_0(w)-1}{w^2},$$

$$\begin{aligned}
a_0(\mathbf{q},s)&=\frac{1}{1+s}-\frac{1}{4\pi}\int_0^{2\pi}d\phi\int_1^\infty\frac{dr}{r}\frac{1}{\chi-s}\left[g_0(\mathbf{q},\chi)\right.\\
&\quad\left.-\frac{1}{r}g_n(\mathbf{q},\chi)+\frac{q^2}{r^2}g_t(\mathbf{q},\chi)\right],\\
a_t(\mathbf{q},s)&=\frac{3\mu}{2}\int_1^\infty\frac{dr}{r}\frac{1}{r-s}\left[\frac{1}{r^2}g_0(\mathbf{q},r)\right.\\
&\quad\left.-\frac{1}{2}\left(1-\frac{1}{r^2}\right)g_t(\mathbf{q},r)-\frac{1}{2}\left(\frac{3}{r^3}-\frac{1}{r}\right)g_n(\mathbf{q},r)\right],\\
a_n(\mathbf{q},s)&=\frac{3\mu}{1+s}+\frac{3\mu}{4\pi}\int_0^{2\pi}d\phi\int_1^\infty\frac{dr}{r}\frac{1}{\chi-s}\left[g_0(\mathbf{q},\chi)\right.\\
&\quad\left.+\frac{q^2}{4}\left(\frac{3}{r^2}-1\right)g_t(\mathbf{q},\chi)-\frac{1}{r}g_n(\mathbf{q},\chi)\right],
\end{aligned}\tag{3.6}$$

and $\chi=r+iql\sqrt{r^2-1}\cos\phi$. The system of equations (3.5) is an extension of the Milne equation to anisotropic scattering

($\mu \neq 0$) and includes an interference contribution ($q \neq 0$). We note that the functions $g_n(\mathbf{q},s)$ and $g_t(\mathbf{q},s)$ are first-order in μ.

According to Eqs. (2.13)–(2.15), the scattering intensity is the Laplace transform of (2.13) with respect to z for $s=1$ and the Fourier transform with respect to the transverse variables for $q=lk_0\theta_s$. In the given situation of almost pure backscattering we have $\cos\theta=\cos(\pi-\theta_s)\approx -1$, $\cos\theta_t \approx 0$, and the interference contribution has the form

$$I_{si}^{(c)}(\mathbf{k}_s,\mathbf{k}_i)=|E_i|^2 A\left\{\frac{1}{4\pi}[g_0(q,1)-g_n(q,1)] - \frac{1}{2}k_0^4 l\tilde{G}(2k)\right\}. \tag{3.7}$$

Consequently, to solve the backscattering problem, it is sufficient to find combinations of the components $g_0(\mathbf{q},s) - s^{-1}g_n(\mathbf{q},s)$.

For isotropic scattering ($\mu=0$) the system of equations (3.5) reduces to the single equation

$$[1-p_0(w)]g_0(\mathbf{q},s)=\frac{1}{1+s}-\frac{1}{4\pi} \times \int_0^{2\pi} d\phi \int_1^{\infty}\frac{dr}{r}\frac{g_0(\mathbf{q},\chi)}{\chi-s}. \tag{3.8}$$

In solving Eq. (3.8) (see Ref. 26), it is essential to utilize its following properties. The right-hand side is regular for $\mathrm{Re}\, s<0$, the unknown function $g_0(\mathbf{q},s)$ by definition is regular for $\mathrm{Re}\, s>0$, and the function $1-p_0(w)$ is even in s and analytic in the complex plane of s, with two cuts $(-\infty,-1)$ and $(1,\infty)$. These properties can be used to find a solution of the given equation by a method analogous to the Wiener–Hopf method. The solution has the form[21]

$$g_0(\mathbf{q},s)=\frac{1}{p_1(iq)(1+s)(1+q)(s+q)} \times \exp\left\{\frac{1}{2\pi i}\int_{-i\infty}^{i\infty}\frac{ds'}{s'}\left(\frac{s}{s'-s}+\frac{1}{s'-1}\right)\ln\left[\frac{p_1(w')}{p_1(iq)}\right]\right\}, \tag{3.9}$$

where $w'=\sqrt{s'^2-q^2}$.

Setting $s=1$ in Eq. (3.9) and transforming to the real variable of integration $s'=iq_z$, we can write the angular dependence of the coherent backscattering intensity for $\mu=0$, according to (3.7), in the form

$$I_{\mathrm{Milne}}^{(c)}(q)\sim\frac{1}{2(1+q)^2 p_1(iq)} \times \exp\left\{-\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{dq_z}{q_z^2+1}\ln\left[\frac{p_1(iv)}{p_1(iq)}\right]\right\}-\frac{1}{2}, \tag{3.10}$$

where $v=\sqrt{q_z^2+q^2}$. We now compare the angular dependence of coherent backscattering according to the exact equation (3.10) with the angular dependence obtained by the well-known method of mirror images according to the equation[4,27]

$$I_{\mathrm{mir}}^{(c)}(q)\sim\int_0^{\infty}dq_z\frac{(1/v)\arctan v}{1-(1/v)\arctan v}f(q_z), \tag{3.11}$$

where

$$f(q_z)=\frac{1+q_z^2-(1-q_z^2)\cos(2z^*q_z)+2q_z\sin(2z^*q_z)}{(1+q_z^2)^2}. \tag{3.12}$$

Notice that Eq. (3.12) contains the parameter z^*, which can only be evaluated on the basis of additional considerations. In the general case the position of the mirror image plane is usually chosen as follows in accordance with the Milne solution: $z^*\ z^*=0.7104(1-\mu)^{-1}$ (Ref. 23).

We have calculated $I^{(c)}(q)$ for $\mu=0$ according to the exact equation (3.10) and according to Eqs. (3.11) and (3.12) with $z^*=0.7104$. The results of the calculations are given in Table I (columns 2 and 4). Clearly, the results obtained from these equations differ by approximately one percent, at least up to values of $q\leqslant 1$. If only the diffusion pole is included in the integrand of Eq. (3.11), so that

$$\frac{v^{-1}\arctan v}{1-v^{-1}\arctan v}\to\frac{3}{v^2}, \tag{3.13}$$

the discrepancy with the exact result increases to 10% (see column 5 in Table I).

TABLE I. Coherent backscattering intensity, normalized to the height of the peak, for a scalar field, $\overline{\cos\theta}=0$, at various scattering angles ($q=kl\sin\theta_s$).

q	$I_{\mathrm{Milne}}^{(c)}$	$I_{\mathrm{mir}}^{(c)}$ $z^*=0.74$	$I_{\mathrm{mir}}^{(c)}$ $z^*=0.71$	$I_{\mathrm{dif}}^{(c)}$ $z^*=0.71$
0	1	1	1	1
0.1	0.806	0.806	0.808	0.793
0.2	0.663	0.664	0.667	0.642
0.3	0.556	0.556	0.560	0.527
0.4	0.473	0.473	0.478	0.439
0.5	0.409	0.409	0.413	0.374

4. SELF-CONSISTENT DETERMINATION OF THE INTERPOLATION LENGTH FROM THE ENERGY BALANCE

In general, setting z^* equal to $z^*=0.7104$ in the above-described scheme is not an exclusive choice. For its evaluation we use an exact integral relation deduced from the Bethe–Salpeter equation. We consider Eqs. (3.1) for $q=0$. These equations are also formally valid outside the medium, i.e., at $Z<0$. Integrating the first equation (3.1) over Z_2, i.e., evaluating the integral of $\gamma_0(\boldsymbol{\rho}_{21},Z_2)$ over the entire volume, we obtain

$$\int_{-\infty}^{\infty}\tilde{\gamma}_0(0,Z_2)dZ_2=l+\frac{1}{4\pi l}\int d\mathbf{R}_2\int_0^{\infty}dZ_3\Lambda(\mathbf{R}_{23})\tilde{\gamma}_0(0,Z_3). \tag{4.1}$$

The integral of the vector function $\mathbf{\Lambda}_1$ over the volume vanishes by virtue of symmetry considerations. From Eq. (4.1), taking the relation $\int d\mathbf{R}\,\Lambda(\mathbf{R})=4\pi l$ into account, we obtain

$$\int_{-\infty}^{0}\tilde{\gamma}_0(0,Z)dZ=l. \tag{4.2}$$

Equation (4.2) is an exact relation. It has the significance of the energy balance of the incident and scattered radiations. We demonstrate this interpretation for point scatterers, $\mu=0$. According to Eq. (2.2), the total intensity of multiple scattering of a plane wave with wave vector $\mathbf{k}_i$ at a point $\mathbf{r}$ is written as follows in the ladder approximation:

$$\langle|\delta E(\mathbf{R})|^2\rangle=\tilde{\gamma}_0(0,Z)|E|^2. \tag{4.3}$$

We note that in the given geometry the average intensity of the scattered field depends only on the distance from the boundary of the medium,

$$\langle|\delta E(\mathbf{R})|^2\rangle=\langle|\delta E(Z)|^2\rangle.$$

Integrating Eq. (4.3) over the region $-\infty<Z<0$, i.e., outside the scattering medium, and taking Eq. (4.2) into account, we obtain

$$l|E|^2=\int_{-\infty}^{0}\langle|\delta E(Z)|^2\rangle dZ. \tag{4.4}$$

The right-hand side represents the total backscattered radiation energy per unit area. The expression on the left-hand side can be written in the form

$$l|E|^2=\int_{-\infty}^{0}E^2\exp[i(\mathbf{k}_i-\mathbf{k}_i^*)\cdot\mathbf{R}]dZ. \tag{4.5}$$

This integral has the significance of the energy of radiation that is incident on the interface $Z=0$ and is effectively damped out in a layer of thickness l. Consequently, Eq. (4.4) actually represents the balance of the incident and scattered radiations, because according to (4.4), the entire radiation energy incident on unit area is returned in the backward direction as diffusing radiation energy in the given situation of elastic scattering.

We use the identity (4.2) to evaluate the parameter z^*. We proceed from the first equation of the system (2.17) as written for the scalar case:

$$\gamma_0(\boldsymbol{\rho}_{21},Z_2)=\frac{1}{2}\,\theta(Z_2)\exp\left(-\frac{Z_2}{l}\right)\delta(\boldsymbol{\rho}_{21})+\frac{1}{4\pi l}\times\int d\mathbf{R}_3\Lambda(R_{23})\Gamma(\boldsymbol{\rho}_{31},Z_3,\mathbf{k}_{23},\mathbf{k}_i). \tag{4.6}$$

We interpret the function Γ in the integrand as the well-known solution given by the method of mirror images. In this method the function $\Gamma(\mathbf{R}_2,\mathbf{R}_1,\mathbf{k}_s,\mathbf{k}_i)$ is replaced by the difference[27]

$$\Gamma(\mathbf{R}_2,\mathbf{R}_1,\mathbf{k}_s,\mathbf{k}_i)\to\Gamma(\mathbf{R}_2,\mathbf{R}_1)=\Gamma_0(|\mathbf{R}_1-\mathbf{R}_2|)-\Gamma_0(|\mathbf{R}_1-\mathbf{R}_2^{(\mathrm{mir})}|), \tag{4.7}$$

where $\mathbf{R}_2^{(\mathrm{mir})}=(X_2,Y_2,-Z_2-2Z^*)$ is the mirror image of the point $\mathbf{R}_2(X_2,Y_2,Z_2)$ about the plane $Z=-z^*l$, and the function $\Gamma_0(R)=3(1-\mu)/4\pi l^3R$ is the solution of the Bethe–Salpeter equation for an unbounded medium. Terms describing the anisotropy of multiple scattering in the unbounded medium are omitted from this solution. As a result, also taking the single-scattering contribution into account, we can write

$$\Gamma(\mathbf{R}_2,\mathbf{R}_1)=\frac{3(1-\mu)}{4\pi l^3}\left[\frac{1}{|\mathbf{R}_1-\mathbf{R}_2|}-\frac{1}{|\mathbf{R}_1-\mathbf{R}_2^{(\mathrm{mir})}|}\right]+\frac{1}{l}\,\delta(\mathbf{R}_1-\mathbf{R}_2). \tag{4.8}$$

Integrating Eq. (4.6) over the transverse variables and over Z_2 in the interval $(-\infty,0)$, we have

$$\int_{-\infty}^{0}\tilde{\gamma}_0(0,Z_2)dZ_2=l=\frac{1}{4\pi l}\int_{-\infty}^{0}dZ_2\times\int_0^{\infty}dZ_3\tilde{\Lambda}(0,Z_{23})\tilde{\gamma}_0^{(\mathrm{mir})}(0,z_3), \tag{4.9}$$

where

$$\tilde{\gamma}_0^{(\mathrm{mir})}(0,z_3)=\int d\boldsymbol{\rho}_{31}\int_0^{\infty}dZ_1e^{-Z_1/l}\Gamma(\mathbf{R}_3,\mathbf{R}_1). \tag{4.10}$$

Substituting Eq. (4.8) into (4.10), we obtain

$$\tilde{\gamma}_0^{(\mathrm{mir})}(0,z)=3(1-\mu)[1+z^*-\exp(-z)]+\exp(-z). \tag{4.11}$$

As a result, from Eq. (4.9) we find

$$z^*=\frac{1}{1-\mu}\left[\frac{1}{3}(5-4\ln 2)+(2\ln 2-1)\mu\right]\approx\frac{0.7425+0.3863\mu}{1-\mu}. \tag{4.12}$$

Equation (4.12) is essentially a consequence of the self-consistency of the diffusion approximation with an identity deduced from the law of conservation of energy.

We have calculated the angular dependence of coherent backscattering for an isotropic scattering diagram, $\mu=0$, according to Eq. (3.11), using the above-determined value $z^*=0.7425$. The results are given in column 3 of Table I. Note the almost perfect agreement of the given self-consistent approach with the exact results (to within $\sim 0.1\%$ error limits), encouraging hope that this approach to the determination of z^* will be equally effective for finite values of μ.

It is important to note that the value of z^* obtained for $\mu=0$ is fairly close to the Milne interpolation length. For large values of μ, however, z^* is appreciably higher than the value $0.71(1-\mu)^{-1}$ given by the pure diffusion approach.

We have used Eq. (4.12) to draw a comparison with experimental data from measurements of the angular dependence of coherent backscattering by particles of finite size.[28,29] We choose the results of measurements in two systems for which the values of the parameter μ are given. In particular, measurements for an aqueous suspension of latex particles of diameter $d=1.091\mu$m are reported in Ref. 28, where the wavelength of the incident radiation is

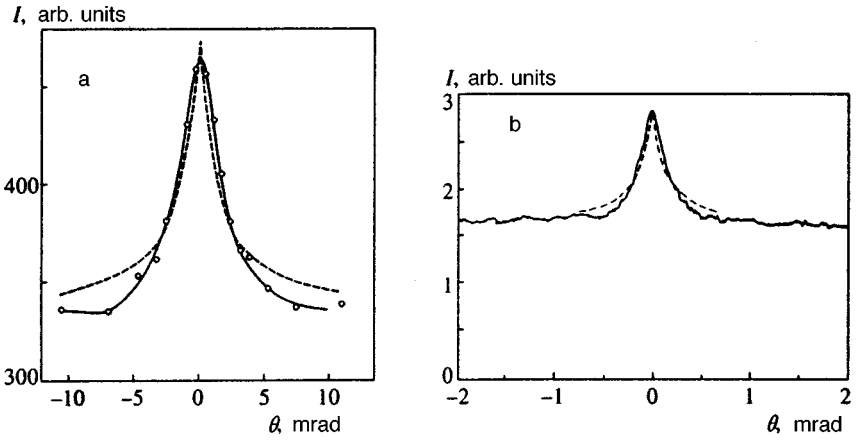

FIG. 1. Comparison of the results of calculations of the coherent backscattering intensity with experimental results for an aqueous suspension of latex particles of diameter 1.091 μm (Ref. 28) (a) and 0.46 μm (Ref. 29) (b). The dashed curves represent the results of calculations for $\mu=0.93$ (a) and $\mu=0.85$ (b).

$\lambda=0.633\,\mu$m, and the refractive index of the medium is $n=1.33$. For this system the authors have measured the extinction length $l=2.6\,\mu$m and from multiple light-scattering data have obtained the value $\mu=0.93$. The measured full width at half maximum (FWHM) of the coherent backscattering peak is $W_{\rm exp}=1.58$ mrad. Our calculations for this value of μ indicate that the height of the peak calculated from Eqs. (3.11) and (4.12) decreases by one half for a value of the parameter $q=kl\theta_s=0.0525$. For the given values of $k=2\pi n/\lambda$ and l we obtain $W_{\rm theor}=1.53$ mrad, which is reasonably consistent with experiment.

In Ref. 29 the FWHM of the peak determined for a suspension of latex particles of diameter $d=0.46\,\mu$m and parameters $\lambda=0.515\,\mu$m, $n=1.33$, and $l=2.8\,\mu$m is $2W_{\rm exp}=4.9$ mrad. The values of the parameters l^* and μ for this system are given in Ref. 27: $l^*=19\,\mu$m, $\mu=0.85$. Our calculations for these values of l^* and μ give $2W_{\rm theor}=4.4$ mrad.

The experimental and calculated data for these two systems are shown in Fig. 1 in a unified scale. It is evident that the theoretical and experimental results are in good agreement. Our comparison with experiment demonstrates the potential interest of measurements over a wider range of angles with a simultaneous determination of the parameters describing the single-scattering diagram. In particular, such an endeavor could reveal how the shape of the coherent backscattering peak is influenced not only by $\overline{\cos\theta}$, but also by higher-order moments $\overline{\cos^n\theta}$ (Ref. 30). These contributions can become significant at large angles, where it is evident from Fig. 1 that the theoretical curves are well above the experimental.

5. ELECTROMAGNETIC FIELD

Here we generalize the above-developed approach to find the analog of the Milne interpolation length z^* in an electromagnetic field. Integrating the equation for the tensor $\hat{\gamma}^{(0)}(\boldsymbol{\rho}_{21},Z_2)$ in the system (2.17) over the volume, we obtain

$$\int_0^\infty \tilde{\gamma}^{(0)}_{ss,11}(0,z)dz+\int_{-\infty}^0 \tilde{\gamma}^{(0)}_{ss,11}(0,z)dz$$
$$=\frac{3}{2}\delta_{s1}l+\frac{3}{8\pi l}\int d\mathbf{R}_1$$
$$\times\int_0^\infty dZ_2\Lambda_{ss,jj}(\mathbf{R}_{12})\tilde{\gamma}^{(0)}_{jj,11}(0,Z_2). \tag{5.1}$$

The integrals of the components of the tensor $\Lambda_{jj,ll}$ over the total volume are readily calculated. We have

$$\int \Lambda_{jj,ii}(\mathbf{R})d\mathbf{R}=\frac{1}{8}\int \Lambda_{jj,jj}(\mathbf{R})d\mathbf{R}=\frac{4\pi l}{15}. \tag{5.2}$$

Substituting Eq. (5.2) into (5.1) and summing over the subscripts, we obtain

$$\sum_s \int_{-\infty}^0 \tilde{\gamma}^{(0)}_{ss,11}(0,Z)dZ=\frac{3}{2}l. \tag{5.3}$$

Equation (5.3) is an extension of the integral identity (4.2) to an electromagnetic field. Here the equivalence of this relation to the balance equation is also easily confirmed for isotropic scattering. In the ladder approximation we can write the following equation for $Z<0$ on the basis of Eq. (2.2):

$$\langle|\delta E_s(Z)|^2\rangle=\frac{1}{lk_0^4\int \tilde{G}(\mathbf{k}_i-\mathbf{k}_s)d\Omega_s}\tilde{\gamma}^{(0)}_{ss,11}(0,Z)|E|^2. \tag{5.4}$$

Integrating this relation in the interval $-\infty<z<0$, summing over the polarizations of the scattered radiation, and making use of the optical theorem (2.18) and Eq. (5.3), we obtain

$$\sum_s \int_{-\infty}^0 \langle|\delta E_s(Z)|^2\rangle dZ=l|E|^2. \tag{5.5}$$

Equation (5.5) signifies that the energy of polarized incident radiation is completely returned from the medium in the form of scattered radiation of all polarizations.

We integrate the first of the three equations (2.17) over the half-space $Z_2<0$ and sum over the exit polarizations. On the basis of Eq. (5.3) we obtain

$$\frac{3}{2}l=\frac{3}{8\pi l}\sum_s \int_{-\infty}^{0}dZ_2\int_{Z_3>0}d\mathbf{R}_3\Lambda_{ss,jj}(\mathbf{R}_{23})\tilde{\gamma}^{(0)}_{jj,11}(0,Z_3). \tag{5.6}$$

As in the scalar case, we replace $\hat{\gamma}^{(0)}$ on the right-hand side with the expression obtained by the method of mirror images

$$\tilde{\gamma}^{(0)\mathrm{mir}}_{jj,11}(0,z)=\frac{3}{2}(1-\mu)(1+z^*-e^{-z})+\frac{3}{2}\delta_{1j}e^{-z}. \tag{5.7}$$

Direct integration gives

$$\frac{3}{4\pi l^2}\int_{-\infty}^{0}dZ_2\int_{Z_3>0}d\mathbf{R}_3\Lambda_{ss,11}(\mathbf{R}_{23})\exp(-z_3)=\begin{cases}0.18394, & s=1,\\ 0.01458, & s=2,\\ 0.025, & s=3,\end{cases} \tag{5.8}$$

$$M_{11,11}=M_{22,22}=3\pi l\left(0.12856+\frac{5}{16}z^*\right),$$

$$M_{11,22}=M_{22,11}=3\pi l\left(0.00625+\frac{1}{48}z^*\right),$$

$$M_{11,33}=M_{33,11}=M_{22,33}=M_{33,22}=3\pi l\left(0.01666+\frac{1}{24}z^*\right),$$

$$M_{33,33}=M_{22,11}=3\pi l\left(0.07648+\frac{1}{6}z^*\right), \tag{5.9}$$

where we have introduced the notation

$$M_{ss,jj}=\frac{3}{2}\int_{-\infty}^{0}dz_2\int_0^{\infty}dz_3\tilde{\Lambda}_{ss,jj}(0,|z_{23}|)\times[1+z^*-\exp(-z)]. \tag{5.10}$$

Substituting Eqs. (5.8)–(5.10) into (5.6), we have

$$z^*=\frac{1}{1-\mu}(0.697+0.4127\mu). \tag{5.11}$$

The value $z^*=0.697$ is very close to the Milne interpolation length $z^*=0.7104$, so that the result (5.11) can be regarded as justification for applying to electromagnetic fields the Milne interpolation length obtained for scalar fields.

We have used Eq. (5.11) to calculate the polarized and depolarized components of the backscattering of light for various values of μ.

When the boundary conditions are taken into account by the method of mirror images, the polarized and depolarized components of coherent backscattering have the form[4,31]

$$I^{(c)}_{VV}(q)\sim\int_{-\infty}^{\infty}dq_z f(q_z)\gamma_{11,11}(\nu), \tag{5.12}$$

$$I^{(c)}_{VH}(q)\sim\int_{-\infty}^{\infty}dq_z f(q_z)\left(\gamma_{12,21}(\nu)\frac{q_z^2}{\nu^2}+\gamma_{13,31}(\nu)\frac{q^2}{\nu^2}\right),$$

where

$$\gamma_{11,11}=\frac{1}{2}\left\{\frac{\tilde{\Lambda}_{11,11}-\tilde{\Lambda}_{11,22}}{1-\xi[\tilde{\Lambda}_{11,11}+\tilde{\Lambda}_{11,22}]}+\frac{(1-\xi\tilde{\Lambda}_{33,33})(\tilde{\Lambda}_{11,11}+\tilde{\Lambda}_{11,22})+2\xi\tilde{\Lambda}^2_{11,33}}{(1-\xi\tilde{\Lambda}_{33,33})[1-\xi(\tilde{\Lambda}_{11,11}+\tilde{\Lambda}_{11,22})]-2\xi^2\tilde{\Lambda}^2_{11,33}}\right\},$$

$$\gamma_{1j,1j}=\frac{\tilde{\Lambda}_{11,jj}}{(1-\xi\tilde{\Lambda}_{1j,1j})^2-\xi^2\tilde{\Lambda}^2_{11,jj}},\quad j=2,3, \tag{5.13}$$

$$\xi=\frac{3}{8\pi l}.$$

The functions $\tilde{\Lambda}_{ij,kl}$ are the Fourier transforms of the components of the tensor (2.9) and have the form

$$\tilde{\Lambda}_{11,11}=\frac{\pi l}{2}(3m_0+2m_1+3m_2),$$

$$\tilde{\Lambda}_{11,22}=\frac{1}{8}\tilde{\Lambda}_{33,33}=\frac{\pi l}{2}(m_0-2m_1+m_2),$$

$$\tilde{\Lambda}_{12,12}=\frac{\pi l}{2}(m_0+6m_1+m_2), \tag{5.14}$$

$$\tilde{\Lambda}_{11,33}=\frac{\pi l}{2}(m_1-m_2),$$

$$\tilde{\Lambda}_{13,13}=\frac{\pi l}{2}(m_0-m_2),$$

where

$$m_0(\nu)=\frac{1}{\nu}\arctan\nu,\quad m_1(\nu)=\frac{1}{\nu^2}[1-m_0(\nu)],$$

$$m_2(\nu)=\frac{1}{\nu^2}\left[\frac{1}{3}-m_1(\nu)\right]. \tag{5.15}$$

The quantities $\hat{\gamma}(\nu)$ result from the Fourier transformation of the tensor $\hat{\Gamma}(\mathbf{R}_1,\mathbf{R}_2|\mathbf{k}_s,\mathbf{k}_i)$ (2.6) in the unbounded medium.

The results of calculations of the polarized and depolarized components of coherent backscattering for various values of the parameter μ are shown in Fig. 2. It is a well-known fact that only the polarized component has a triangular peak. It is evident from Fig. 2a that, as in the scalar case, the steepness of the peak increases sharply as the single-scattering anisotropy increases, consistent with experiment and theoretical predictions. The angular dependence for a scalar field is shown in the same figure.

The linear slope of the peak is known to be attributable to the presence of a diffusion pole. For an electromagnetic field it is evident from Fig. 2a that the slope of the normalized polarized components is somewhat smaller than the slope calculated for a scalar field with the same values of μ. This disparity indicates that the relative contribution of nondiffusion terms increases for a vector field.

The diffusion pole contributes absolutely nothing to the depolarized component, and it is evident from Fig. 2b that this fact leads to a peak of the Lorentzian type.

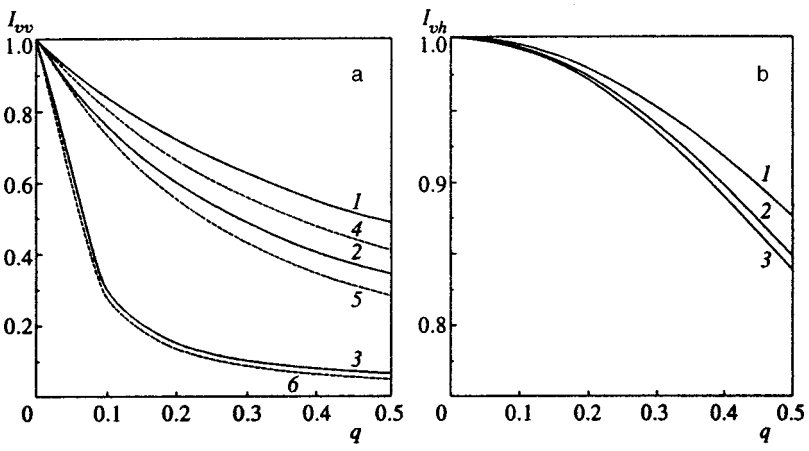

FIG. 2. (a) Coherent backscattering intensity of the polarized component, normalized to the height of the peak, as a function of the angle $q=kl\sin\theta_s$, calculated for various values of μ according to Eqs. (5.11) and (5.12): (*1*) $\mu=0$; (*2*) 0.5; (*3*) 0.93. Also shown for comparison are the results of calculations (dashed curves) for a scalar field according to Eqs. (3.11) and (4.12) using the same values of the scattering parameter: (*4*) $\mu=0$; (*5*) 0.5; (*6*) 0.93. (b) Angular dependence of the depolarized component: (*1*) $\mu=0$; (*2*) 0.5; (*3*) 0.93.

We have not compared the results of the calculations according to Eqs. (5.11) and (5.12) with experiment, because the values of the numerical parameters for z^* in the scalar problem (4.12) and the vector problem (5.11) are close and, taking the experimental error limits into account, essentially yield identical results for the polarized component.

6. CONCLUSION

We have investigated the Bethe–Salpeter equation for a strongly inhomogeneous medium with an anisotropic single-scattering diagram. For a medium occupying a half-space this equation reduces to a system of equations for the coefficients of an expansion in Legendre polynomials. By convolution on the spatial variables we have obtained an integral equation that can be interpreted as the energy balance of the incident and scattered radiations. The application of the method of mirror images in this integral identity yields a simple expression for the parameter characterizing the position of the effective boundary of the mirror image in the general case of anisotropic single scattering.

We have compared these results quantitatively, on the one hand, with the exact theoretical results of Milne for isotropic scattering and, on the other, with existing experimental data on the angular dependence of coherent backscattering in systems with a large parameter μ, which accounts for the anisotropy of single scattering.

In our description of the peak for isotropic scattering we have obtained agreement with the predictions of Milne theory to within fractions of one percent. A rigorous theory does not yet exist for anisotropic scattering, and we have drawn on the good agreement with experiment as a criterion of the final results.

We have generalized the newly developed approach to an electromagnetic field. We have calculated and angular and polarization dependences of coherent backscattering for various values of μ. We have compared the predictions of the theory for the backscattering peak in the model of scalar waves and for an electromagnetic field in order to discern the influence of the vectorial nature of the latter field on coherence effects in multiple scattering. We have found that allowance for the vectorial nature of an electromagnetic field produces a relative decrease in the contribution of the diffusion pole responsible for coherent backscattering effects.

The foregoing discussion is valid in the weak scattering approximation, $\lambda \ll l$. However, the parameter l also has an upper bound imposed by currently existing experimental accuracy. The width of the coherent backscattering peak θ_w is determined from the relation $kl^*\theta_w \sim 1$. Assuming that the error of the angular measurements is of the order of 10^{-3}, we obtain $l^* < (10^2-10^3)\lambda$, i.e., $l < (1-\mu)(10^2-10^3)\lambda$. This condition significantly restricts the choice of systems in which the backscattering peak is accessible to observation and, in particular, excludes systems in which the multiple scattering regime is achieved only at the expense of large sample thicknesses.

The results obtained in this study can be generalized to more complex scattering system geometries. The method developed here is also applicable for calculations of time correlation functions and for the solution of problems of visualizing the structure of opaque systems in diffusing wave spectroscopy.

This work has received financial support from the Russian Fund for Fundamental Research (RFFI Grant No. 98-02-18201).

*)E-mail: vadim.romanov@pobox.spbu.ru

[1] P. J. Pine, D. A. Weitz, G. Maret, P. E. Wolf, E. Herbolzheimer, and P. M. Chaikin, in *Scattering and Localization of Classical Waves in Random Media*, P. Sheng (Ed.), World Scientific, Singapore (1989).
[2] P. E. Wolf and G. Maret, in *Scattering in Volumes and Surfaces*, M. Nieto-Vesperanos and J. C. Dainty (eds.), Elsevier, Elsevier (1990), p. 37.
[3] Yu. N. Barabanenkov, Yu. A. Kravtsov, V. D. Ozrin, and A. I. Saichev, in *Progress in Optics*, E. Wolf (ed.), Vol. 29, 247 (1991).
[4] V. L. Kuz'min and V. P. Romanov, Usp. Fiz. Nauk **166**, 247 (1996).
[5] B. A. van Tiggelen and R. Maynard, in *Waves in Random and Other Complex Media*, G. Papanicolaou, R. Burridge, and L. Pastor (Eds.), Springer-Verlag, Berlin-New York (1997).
[6] D. B. Rogozkin, Zh. Éksp. Teor. Fiz. **111**, 1674 (1997) JETP **84**, 916 (1997).
[7] E. E. Gorodnichev and D. B. Rogozkin, Waves Random Media **4**, 51 (1994)
[8] D. S. Wiersma, M. P. van Albada, B. A. van Tiggelen, and A. Lagendijk, Phys. Rev. Lett. **74**, 4193 (1995).
[9] M. U. Vera, P.-A. Lemieux, and D. J. Durian, J. Opt. Soc. Am. A **14**, 2800 (1997).
[10] F. Scheffold, W. Hartl, G. Maret, and E. Matijevic, Phys. Rev. B **56**, 10942 (1997).

[11] L. V. Korolev and D. B. Rogozkin, Zh. Éksp. Teor. Fiz. **113**, 291 (1998) [JETP **86**, 164 (1998)].
[12] F. Scheffold and G. Maret, Phys. Rev. Lett. **81**, 5800 (1998).
[13] A. Yodh and B. Chance, Phys. Today **10**, 34 (1995).
[14] D. Bicout and G. Maret, Physica A **210**, 87 (1994).
[15] M. Heckmeier and G. Maret, Europhys. Lett. **34**, 257 (1996).
[16] M. Heckmeier, S. E. Skipetrov, G. Maret, and R. Maynard, J. Opt. Soc. Am. A **14**, 185 (1997).
[17] S. E. Skipetrov and I. V. Meglinskiĭ, Zh. Éksp. Teor. Fiz. **113**, 1213 (1998) [JETP **86**, 661 (1998)].
[18] V. V. Tuchin, Usp. Fiz. Nauk **167**, 517 (1997).
[19] Yu. N. Barabanenkov and V. D. Ozrin, Zh. Éksp. Teor. Fiz. **94**(6), 56 (1988) [Sov. Phys. JETP **67**, 1117 (1988)].
[20] M. B. van der Mark, M. P. van Albada, and A. Lagendijk, Phys. Rev. B **37**, 3575 (1988).
[21] T. M. Nieuwenhuizen and J. M. Luck, Phys. Rev. E **48**, '569 (1993).
[22] R. Aronson, J. Opt. Soc. Am. A **12**, 32532 (1995).
[23] A. A. Golubentsev, Zh. Éksp. Teor. Fiz. **86**, 47 (1984) [Sov. Phys. JETP **59**, 26 (1984)].
[24] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, 2nd ed. (rev. and enl., with L. P. Pitaevskiĭ), Pergamon Press, Oxford-New York (1984).
[25] E. Amic, J. M. Luck, and T. M. Nieuwenhuizen, J. Phys. A **29**, 4915 (1996).
[26] B. Davison and J. B. Sykes, *Neutron Transport Theory*, Clarendon Press, Oxford (1957).
[27] E. Akkermans, P. E. Wolf, and R. Maynard, Phys. Rev. Lett. **56**, 1471 (1986).
[28] M. P. van Albada and A. Lagendijk, Phys. Rev. Lett. **55**, 2692 (1985).
[29] P. E. Wolf and G. Maret, Phys. Rev. Lett. **55**, 2696 (1985).
[30] V. L. Kuz'min and V. P. Romanov, Zh. Éksp. Teor. Fiz. **113**, 2022 (1998) [JETP **86**, 1107 (1998)].
[31] V. L. Kuzmin and V. P. Romanov, Phys. Rev. E **56**, 6008 (1997).

Translated by James S. Wood

Stabilization of circular states of the hydrogen atom in a strong field

E. A. Volkova, A. M. Popov, and O. V. Tikhonova*)

D. V. Skobel'tsyn Scientific-Research Institute of Nuclear Physics at M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

(Submitted 16 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 1929–1940 (December 1999)

A solution of the three-dimensional time-dependent Schrödinger equation, describing the ionization dynamics of the hydrogen atom in a circular state in an electromagnetic field, is obtained by direct numerical integration. It is shown that the observed stabilization effect can be interpreted on the basis of the Kramers–Henneberger approach. A simple analytical model is proposed, which qualitatively describes the basic laws of the ionization process under the conditions of the reported calculations and laboratory experiments on ionization of the circular hydrogenlike $5g$, $m=4$ state of the Ne atom. © *1999 American Institute of Physics.* [S1063-7761(99)00412-6]

1. INTRODUCTION

The stabilization of atomic systems in the presence of ionization by an electromagnetic field, predicted earlier[1,2] and observed recently in experimental work,[3,4] is one of the most fascinating phenomena in the physics of interaction between a high-intensity laser beam and matter. The crux of the phenomenon is that when the laser intensity is increased above a certain critical value, called the stabilization threshold, the ionization probability does not increase (or it even decreases). From the theoretical point of view two principal stabilization mechanisms are discerned at the present time: interference stabilization[2,5–7] and stabilization in the Kramers–Henneberger (KH) regime.[1,8–11] However, major difficulties are encountered in proving experimentally the prevalence of one mechanism or the other in a specific situation.

The swift progress of computer technology in recent times has set the stage for computer experiments based on direct numerical integration of the time-dependent Schrödinger equation for a three-dimensional quantum system in an electromagnetic field without any simplifying assumptions.[12–22] Calculations of this kind in application to the dynamics of the hydrogen atom in a strong electromagnetic field have made it possible to observe stabilization in the KH regime,[13] to corroborate the basic postulates of the theory of interference stabilization,[19,20] and to investigate salient characteristics of resonance multiphoton ionization of the ground state of the hydrogen atom.[21]

In this paper we discuss the results of calculations of the dynamics of ionization of a Rydberg hydrogen atom existing initially in a circular state. We determine the probabilities of ionization and population of various atomic states toward the end of the laser pulse. We disclose saturation of the ionization probability at a level much lower than unity. We construct an analytical model, which integrates the resulting data based on the KH approach and can be used to describe with qualitative correctness the results of the calculations, along with those of laboratory experiments on the ionization of the hydrogenlike neon atom in the circular $5g$ state. We discuss the interpretation of the observed dynamics of the atomic system in a wave field in the basis of free-atom states.

2. NUMERICAL MODEL

When a laser pulse with linear polarization of the field is incident on an atom, electromagnetic transitions take place without any change in the projection of the electron orbital momentum onto the direction of the electric field vector ε. If the quantization axis (z axis) of the initial state is directed along the field vector $\varepsilon(t)$ (as is the case in laboratory experiments[4]), the electron wave function is written as follows in cylindrical coordinates:

$$\Psi(\rho,z,\phi,t)=\psi(\rho,z,t)\exp(im\phi),$$

where m is the magnetic quantum number governing the projection of the orbital momentum onto the quantization axis. The function $\psi(\rho,z,t)$ satisfies the two-dimensional time-dependent Schrödinger equation

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2\mu}\left(\frac{1}{\rho}\frac{\partial}{\partial\rho}\rho\frac{\partial\psi}{\partial\rho}+\frac{\partial^2\psi}{\partial z^2}\right)+V_{\text{eff}}^{(m)}(\rho,z)\psi(\rho,z,t)-ez\varepsilon(t)\cos(\omega t)\psi(\rho,z,t) \tag{1}$$

with the effective potential

$$V_{\text{eff}}^{(m)}(\rho,z)=V_C(\rho,z)+\frac{\hbar^2m^2}{2\mu\rho^2}, \tag{2}$$

which takes into account the presence of a centrifugal barrier for an electron for which the projection of its momentum onto the z axis is equal to $m\hbar$. Here V_C is the Coulomb potential.

As in Refs. 19 and 20, in the calculations we use the smoothed Coulomb potential

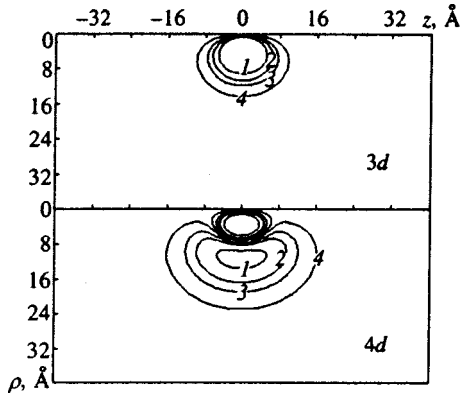

FIG. 1. Distribution of the probability density function in the $3d$ and $4d$ states for a magnetic quantum number $m=2$. Level lines: (*1*) 10^{-3}; (*2*) 10^{-4}; (*3*) 10^{-4}; (*4*) 10^{-5}.

$$V_C=-\frac{e^2}{\sqrt{a^2+\rho^2+z^2}} \tag{3}$$

with the smoothing parameter $a=0.5a_0$, where a_0 is the Bohr radius. It has been shown[19] that this potential preserves the structure of the energy spectrum and the wave functions of the hydrogen atom, so that for all states with a smooth quantum number $n\geqslant 3$ the difference in the energies of stationary states in the real and smoothed Coulomb potentials does not exceed one percent.

We note that the circular state characterized by an orbital quantum number $l=m$ and a principal quantum number $n=m+1$ is the ground state in the potential (2). All lower states of the hydrogen spectrum are eigenstates of a Hamiltonian with the effective potential (2) for smaller values of the magnetic quantum number; transitions to these states are forbidden in the presence of a linearly polarized radiation field.

Many of the calculations discussed in this paper are carried out for an initial circular $3d$, $m=2$ state. The structures of this state and of the lowest excited $4d$, $m=2$ state are shown in Fig. 1. The procedure used to calculate the wave functions of stationary states of the hydrogen atom are similar to those in Ref. 19.

The ionization dynamics calculations are carried out for radiation with a photon energy $\hbar\omega=5.0$ eV in the range $P=5\times10^{11}-1\times10^{16}$ W/cm^2. Here the initial $3d$ state is associated with a one-photon transition continuum. The envelope of the pulse is assigned a trapezoidal profile with smoothed edges of duration $\tau_f=5T$ and a plateau $\tau_{\rm pl}=10T$ ($T=2\pi/\omega$ is the duration of the optical cycle):

$$\varepsilon(t)=\begin{cases}\varepsilon_0\sin^2\dfrac{\pi t}{2t_f}, & t\leqslant t_f,\\ \varepsilon_0, & t_f\leqslant t\leqslant t_f+t_{\rm pl},\\ \varepsilon_0\sin^2\dfrac{\pi[t-(2t_f+t_{\rm pl})]}{2t_f}, & t_f+t_{\rm pl}\leqslant t\leqslant 2t_f+t_{\rm pl}.\end{cases} \tag{4}$$

Some of the calculations are also carried out for an initial $5g$, $m=4$ state and a photon energy $\hbar\omega=2.0$ eV to model conditions similar to those encountered in experiment.[4]

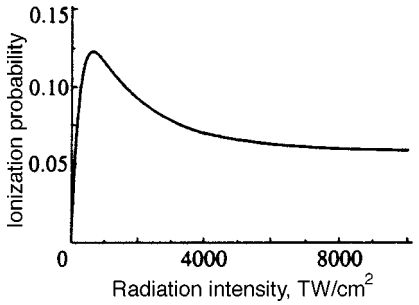

FIG. 2. Dependence of the ionization probability of a hydrogen atom existing initially in the circular $3d$, $m=2$ state on the radiation intensity within the limits of a trapezoidal pulse with smoothed leading edges of duration $\tau_f=5T$ and a plateau $\tau_{\rm pl}=10T$.

The procedure used to solve the time-dependent Schrödinger equation numerically and to calculate the probabilities of ionization and population of bound atomic $|nlm\rangle$ states is described in Ref. 19. For a given m all possible $nl(l\geqslant m)$ states are taken into account for $(l+1)\leqslant n\leqslant n_{\max}=13$.

3. DISCUSSION OF THE RESULTS

We begin with an analysis of the ionization dynamics of the circular $3d$ state of the hydrogen atom. The main numerical simulation result is the dependence of the ionization probability of the hydrogen atom in the $3d$, $m=2$ state on the radiation intensity within the limits of a smoothed trapezoidal pulse; this dependence is shown in Fig. 2. Clearly, the ionization probability as a function of the intensity, $W_i(P)$, is linear in the weak-field range, attains a maximum at $P^*=6\times10^{14}$ W/cm^2, and becomes a decreasing function of the intensity up to values of the latter $\sim10^{16}$ W/cm^2. The distribution of the electron density in the continuum, calculated at the end of the laser pulse and shown in Fig. 3 for an intensity $P=10^{15}$ W/cm^2, indicates that ionization is a one-photon process in the given range of the parameters.

We have obtained a similar $W_i(P)$ curve in calculations for the initial $5g$, $m=4$ state. Here the qualitative form of the curve and the value of the stabilization threshold agree with the data of laboratory experiments[4] and computer calculations.[22] However, detailed investigations of the causes of stabilization are not reported in these papers. Another issue that has not been fully explained is the sizable deviation, even in comparatively weak fields, of the $W_i(P)$ curve from the result of perturbation theory.

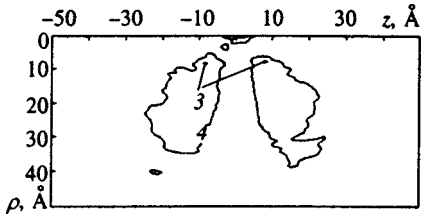

FIG. 3. Distribution of the electron density in the continuum at termination of a laser pulse having an intensity on the lasing ''shelf'' 10^{15} W/cm^2. The level lines correspond to Fig. 1.

A distinctive feature of circular states is their ''squeezing out'' from the Coulomb center by the centrifugal potential, a phenomenon that diminishes the matrix element characterizing the ionization probability, $d_{nE}=\langle n,l=n-1, m=l|d_z|E,l+1\rangle$. In effect, for pulse durations τ such that

$$\Gamma_i\tau\ll 1 \tag{5}$$

(where Γ_i is the ionization width of the initial state, calculated by Fermi's golden rule) and

$$\Gamma_i/\Delta E_n\ll 1 \tag{6}$$

(where $\Delta E_n=E_{n+1}-E_n$ is the energy distance between consecutive levels), perturbation theory should be valid for calculating the ionization probability and linear dependence of the ionization probability on the intensity and duration of the laser pulse:

$$W_i\propto|V_{nE}|^2\tau\propto|d_{nE}|^2P\tau. \tag{7}$$

The above-stated criteria of the validity of perturbation theory yield estimates of the radiation intensity taking into account the known value of the matrix element for the $3d$, $m=2$ state, $\hbar\omega=5$ eV ($d_{nE}\simeq 0.32$ a.u.; Ref. 23): $P\ll 4\times10^{14}$ W/cm^2 and $P\ll 5\times10^{15}$ W/cm^2, respectively. In the calculations, however, saturation of the ionization is observed for $P>P^*=6\times10^{14}$ W/cm^2, and the linear dependence (7) is observed only in the range of intensities $P<3\times10^{13}$ W/cm^2. In this case the slope of the $W_i(P)$ curve in the weak-field range can be used to calculate the matrix element d_{nE} and yields a value $d_{nE}\simeq 0.27$ a.u., which is in good agreement with previously reported computational data.[23] Consequently, the data obtained here begin to deviate from the results of perturbation theory in the range of comparatively weak laser intensities $P>3\times10^{13}$ W/cm^2. This conclusion is further corroborated by the results of calculations of the dynamics of population of the ground state in the given effective potential $V_{\text{eff}}^{(m=2)}$ during the laser pulse $W_0(t)$ for various laser intensities.

The validity of perturbation theory rests on the assumption that during the active period of the laser pulse the system resides predominantly in the $|n_0l_0\rangle$ state, i.e.,

$$W_0(t)=|\langle\psi(t)|n_0l_0\rangle|^2\simeq 1,$$

and the population probabilities of all other states of the discrete and continuous spectra of the atomic Hamiltonian are small. For intensities $P\leqslant 10^{13}$ W/cm^2 the total population of all states except the ground $3d$ state is indeed small, and perturbation theory is valid. For $P\geqslant 10^{14}$ W/cm^2 the population of the ground state during the entire period of the pulse is a rapidly oscillating function of time with a period equal to half the duration of the optical cycle. During the entire optical cycle the population of the ground state is close to zero, and only at times corresponding to the classical turning points of a free electron oscillating in an electromagnetic field does it increase to values close to unity. This dynamics of the process clearly rejects perturbation theory and indicates the occurrence of strong electromagnetic transitions in the system, even though the ionization probability is small in the final state after termination of the incident laser pulse. We note that the observed ionization probability toward the end of the laser pulse is substantially lower than predicted by perturbation theory, evincing stabilization of the investigated system.

One possible explanation for the observed stabilization could be an interference mechanism.[2,5–7] However, significant interference of the amplitudes of transition to the continuum, describing transitions of the form

$$n\to E,$$

$$n\to E'\to n'\to E,$$

occurs only under the condition

$$\Gamma_i/\Delta E_n\sim 1,$$

which corresponds to intensities $P\geqslant 5\times10^{15}$ W/cm^2. Consequently, interference of the amplitudes of the given transitions to the continuum is insignificant over essentially the entire range of our calculations.

According to various estimates, on the other hand, the threshold wave field corresponding to the overbarrier ionization regime for the $3d$ state does not exceed an intensity $\sim10^{11}$ W/cm^2 (Refs. 24 and 25), i.e., over the entire investigated range an atomic electron does in fact execute almost free oscillations. In terms of the populations of states of the atomic basis these oscillations are described as transitions between the initial $3d$ state and the continuum with allowance for free-free transitions between different states of the continuum and, under the stated conditions, cannot be described in the first order of perturbation theory.

The presence of almost free oscillations of an electron wave packet, corresponding to the overbarrier ionization regime, suggests that the computational data could be interpreted within the framework of the Kramers-Henneberger method.[16] Indeed, it has been shown[26] that, given the condition $\hbar\omega>I_{\text{at}}$ (I_{at} is the zero-field ionization potential of an atom), the basis of KH states is preferable from the physical standpoint over the basis of field-unperturbed states for any values of the radiation intensity. Under the stated conditions, therefore, the dynamics of population of states of the discrete spectrum and the continuum within the period of the laser pulse should in fact be investigated in the basis of KH states. In this case stationary states of the system can be sought in the KH potential

$$V_{\text{KH}}=\frac{1}{2\pi}\int_0^{2\pi}V_C(\rho,z-\alpha_e\cos\xi)d\xi, \tag{8}$$

where $\alpha_e=e\varepsilon_0/m\omega^2$ is the amplitude of the free-electron oscillations in an electromagnetic field of strength ε_0.

Over the entire investigated range of intensities up to $P=10^{16}$ W/cm^2 the amplitude α_e of the free-electron oscillations in the electromagnetic field is found to be smaller than the characteristic localization length of the wave function $a_0n^2=9a_0$ for $3d$, $m=2$ states in the potential $V_C(\rho,z)$. In this case we can form a series expansion of the integrand in Eq. (8):

$$V_C(\rho,z-\alpha_e\cos\xi)=V_C(\rho,z)-\alpha_e\cos\xi\frac{\partial V_C}{\partial z}+\frac{\alpha_e^2}{2}\cos^2\xi\frac{\partial^2 V_C}{\partial z^2}+\ldots$$

Then, evaluating the integral (8), we obtain[1)]

$$V_{\rm KH}(\rho,z)\simeq-\frac{e^2}{r}-\frac{\alpha_e^2}{2}\frac{e^2}{r^3}P_2(\cos\theta), \tag{9}$$

where $r=\sqrt{\rho^2+z^2}$, $P_2(\cos\theta)=(1/2)(3\cos^2\theta-1)$ is a Legendre polynomial, and $\cos\theta=z/r$.

Regarding the second term in Eq. (9) as a perturbation, we find the difference between the energy of the KH state and the energy of the initial atomic state:

$$\Delta E_{nl}^{\rm KH}=\langle nl,m=2|-\frac{\alpha_e^2}{2}\frac{e^2}{r^3}P_2(\cos\theta)|nl,m=2\rangle=\frac{\alpha_e^2}{2}e^2\left\langle\frac{1}{r^3}\right\rangle\frac{l}{2l+3}. \tag{10}$$

Making use of the relation

$$\left\langle\frac{1}{r^3}\right\rangle=\frac{1}{a_0^3n^3l(l+1/2)(l+1)},$$

we obtain the following expression for the energy levels in the KH potential:

$$E_{nl}^{\rm KH}=-\frac{e^2}{2a_0}\frac{1}{n^2}-\frac{\alpha_{\rm KH}\varepsilon_0^2}{4}, \tag{11}$$

where

$$\alpha_{\rm KH}=-\frac{e^4}{m^2\omega^4a_0^3}\frac{1}{n^3(l+1/2)(l+1)(l+3/2)}.$$

It is instructive to compare the resulting expression for the energy of stationary states in the KH potential in the weak-field range ($\alpha_e\ll a_0n^2$) with the Stark shift of the levels of a hydrogenlike atom

$$\Delta E_{nl}=-\beta\varepsilon_0^2/4. \tag{12}$$

Calculations of this kind have been carried out previously.[27] The following equation (in atomic units) was obtained for states with $n\simeq l\gg1$:

$$\beta=-\frac{1}{\omega^2}-\frac{2l}{2l+3}\frac{l}{\omega^4}\left\langle\frac{1}{r^3}\right\rangle+\ldots \tag{13}$$

Here the first term of the expansion ($-1/\omega^2$) corresponds to the vibrational energy of a free electron in an electromagnetic field. This term vanishes upon transition to a Kramers oscillating coordinate system. The next term of the expansion gives the deviation of the Stark shift of Rydberg levels from the shift of the continuum boundary. It is evident from a comparison of Eqs. (10), (11) and (12), (13) that under the conditions of small-amplitude oscillations of the wave packet the electron energy spectrum in the KH potential coincides with the atomic spectrum adjusted for the Stark shift of levels. This result is important from the general physical point of view and encourages a new look at the modification of the atomic spectrum in the presence of a radiation field: In weak fields the Kramers–Henneberger approach is merely a more practical method for calculating the shift of the atomic levels. Moreover, this approach provides a means for visualizing directly how the atomic potential is distorted under the influence of an electromagnetic field.

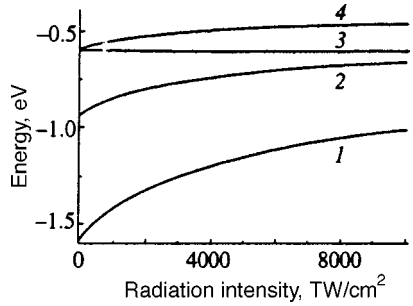

FIG. 4. Dependence of the energy of stationary $(3d)_{\rm KH}$ (*1*), $(4d)_{\rm KH}$ (*2*), $(5g)_{\rm KH}$ (*3*), and $(5d)_{\rm KH}$ (*4*) states in the KH potential on the laser intensity.

Direct numerical calculations of the energy levels (Fig. 4) and the wave functions of stationary states in the KH potential in the investigated range of intensities confirm the stated conclusions. The wave functions of the ground stationary state in the atomic potential, $3d$, $m=2$, and the corresponding KH states with the same projection of the orbital momentum onto the direction of the electric field, $(3d)_{\rm KH}$, $m=2$, for various intensities are close to one another, and the energies of stationary KH states in the range of intensities corresponding to the stabilization threshold differ at most by 10% from the corresponding free-atom values. In addition, for $P\leqslant10^{14}$ W/cm^2 the dependence of $E_{nl}^{\rm KH}$ on the intensity is linear, i.e., agrees with Eq. 11. The determination of the polarizability $\alpha_{\rm KH}$ from the slope of the $E_{nl}^{\rm KH}(P)$ curve for the $(3d)_{\rm KH}$ state (Fig. 4, curve *1*) yields a value that agrees, within 5% error limits, with the value of $\alpha_{\rm KH}$ calculated from Eq. (11).

Again we note that the stabilization regime emerges even in weak fields when $\alpha_e\ll a_0n^2$, the KH potential has a single well and differs only somewhat from the atomic potential, and the energy levels in it coincide with those in the atomic potential adjusted for the Stark shift. In this situation the causes underlying the onset of stabilization require further careful study.

In the KH approach the transition of an electron from a bound state to the continuum takes place under the influence of harmonics of the KH potential:

$$V_{\rm KH}^{(j)}(\rho,z)=\frac{1}{2\pi}\int_0^{2\pi}V_C(\rho,z-\alpha_e\cos\xi)\exp(-ij\xi)d\xi, \tag{14}$$

where the jth harmonic describes transitions in the system with an energy increment equal to $j\hbar\omega$.

From the physical standpoint the onset of stabilization in the KH regime is attributable to a decrease in the probability of transition from a bound KH state to the KH continuum

TABLE I. Squares of the overlap integrals $|\langle nl,m=2|(3d)_{\mathrm{KH}},m=2\rangle|^2$ of the wave functions of stationary Kramers–Henneberger states and the corresponding states of the unperturbed atom.

P,W/cm^2	$3d$	$4d$	$5g$	$5d$
0	1	0	0	0
10^{13}	0.995	2.3×10^{-3}	1.4×10^{-4}	6.1×10^{-4}
10^{14}	0.990	4.5×10^{-3}	1.1×10^{-4}	1.2×10^{-3}
6×10^{14}	0.957	2.02×10^{-2}	4.0×10^{-6}	5.5×10^{-3}
10^{15}	0.926	3.31×10^{-2}	2.0×10^{-5}	9.3×10^{-3}
3×10^{15}	0.786	8.25×10^{-2}	8.3×10^{-4}	2.68×10^{-2}
10^{16}	0.500	0.147	6.90×10^{-3}	6.00×10^{-2}

under the influence of harmonics of the potential $V_{\mathrm{KH}}^{(j)}(\rho,z)$. This transition probability is described by matrix elements of the form

$$\langle V_{\mathrm{KH}}^{(j)}\rangle=\langle(3d)_{\mathrm{KH}}|V_{\mathrm{KH}}^{(j)}|E_{\mathrm{KH}}\rangle,$$

where $|(nl)_{\mathrm{KH}}\rangle$ and $|E_{\mathrm{KH}}\rangle$ are the wave functions of the discrete and continuous spectra of the KH potential. In weak fields it is sufficient to restrict the discussion to transitions occurring only under the influence of the lowest harmonic $V_{\mathrm{KH}}^{(1)}(\rho,z)$. Allowing for the fact that stabilization is observed in our calculations for $\alpha_e\ll a_0n^2$, we write $V_{\mathrm{KH}}^{(1)}$ in the form

$$V_{\mathrm{KH}}^{(1)}(\rho,z)=-\frac{1}{2}\frac{\alpha_e e^2}{r^2}P_1(\cos\theta)-\frac{3}{8}\frac{\alpha_e^3e^2}{r^4}P_3(\cos\theta)+\ldots \tag{15}$$

Rewriting Eq. (15) in the form

$$V_{\mathrm{KH}}^{(1)}(\rho,z)\simeq-e\varepsilon_0zf(\rho,z), \tag{16}$$

where

$$f(\rho,z)=\frac{e^2}{2m\omega^2}\left\{\frac{1}{r^3}-\frac{9}{8}\frac{\alpha_e^2}{r^5}+\frac{15}{8}\frac{\alpha_e^2z^2}{r^7}+\ldots\right\}, \tag{17}$$

we note that the interaction operator $V_{\mathrm{KH}}^{(1)}$ can be represented by a product of two factors, one of which, $-ez\varepsilon_0$, corresponds to the operator of interaction with an electromagnetic field in $d\varepsilon$ gauge in the atomic basis. Stabilization in the KH regime sets in because the matrix element $\langle V_{\mathrm{KH}}^{(1)}\rangle$ does not increase as the field increases. This behavior is possible when the increase in the ionization probability with increasing intensity, characterized by the operator $-ez\varepsilon_0$, is compensated by the contribution of the second factor $f(\rho,z)$. The onset of stabilization is also conducive to an increase in the region of localization of a bound KH state in comparison with the corresponding atomic state and (or) an increase in the photoelectron energy in the continuum as a result of the decrease in the electron binding energy in the KH potential as the intensity is increased. However, the overlap integrals of the wave functions of the atomic state and the KH $|\langle(3d)_{\mathrm{KH}}|nl\rangle|^2$ states, shown in Table I for various intensities, and the curves representing the energies of stationary states of the KH potential in the interval of onset of stabilization $P=10^{14}-10^{15}$ W/cm^2 (see Fig. 4) show that these two factors can be disregarded in our situation. Estimating the amplitude of the probability of transition from the discrete spectrum to the continuum as

$$C_i\sim\langle E_{\mathrm{KH}}|V_{\mathrm{KH}}^{(1)}|(nl)_{\mathrm{KH}}\rangle\simeq-e\varepsilon_0\sum_{n',l'}\langle E_{\mathrm{KH}}|z|(n'l')_{\mathrm{KH}}\rangle\times\langle(n'l')_{\mathrm{KH}}|f|(nl)_{\mathrm{KH}}\rangle$$

and ignoring the difference between the wave functions of stationary states in the atomic and KH potentials, we obtain

$$C_i\sim-e\varepsilon_0\langle E|z|nl\rangle\langle f\rangle, \tag{18}$$

where

$$\langle f\rangle=\langle nl|f(\rho,z)|nl\rangle$$

is the average value of f over the initial quantum $|nl\rangle$ state. Invoking the explicit expression for the radial wave function of the circular $|nl\rangle=|n,n-1\rangle$ state,

$$R_{n,n-1}(r)=\sqrt{\frac{8}{(2n)!n^3a_0^3}}\left(\frac{2r}{na_0}\right)^{n-1}\exp\left(-\frac{r}{na_0}\right), \tag{19}$$

for $l\geqslant2$ we obtain

$$\langle V_{\mathrm{KH}}^{(1)}\rangle\approx-e\varepsilon_0z_{nE}\frac{e^2}{2m\omega^2a_0^3}\frac{1}{n^4(n-1/2)(n-1)}\times\left\{1-\frac{9}{8}\frac{\alpha_e^2}{a_0^2n^2(n-2)(n-3/2)}\right\}. \tag{20}$$

Making use of the fact that the ionization probability $W_1\sim|\langle V_{\mathrm{KH}}^{(1)}\rangle|^2$, from Eq. (20) we estimate the circular state stabilization threshold:

$$\alpha_e^2\approx(4/9)a_0^2n^2(n-2)(n-3/2). \tag{21}$$

Estimates of the stabilization threshold by means of Eq. (21) for the circular $3d$, $m=2$ state with $\hbar\omega=5$ eV give $\alpha_e^2\approx6a_0^2$, so that $P^*\cong3.5\times10^{14}$ W/cm^2, in good agreement with the results of our calculations. We also note that under the conditions of experiments[4] on the ionization of the hydrogenlike Ne atom in the circular $5g$, $m=4$ state by laser radiation with $\hbar\omega=2$ eV, from Eq. (21) we obtain $\alpha_e^2\approx100a_0^2$ and $P^*\simeq8\times10^{13}$ W/cm^2, which are also close to the data of laboratory[4] and computer[22] experiments.

The most conclusive proof of the validity of the Kramers–Hennebergen picture under the conditions of our calculations lies in the data shown in Fig. 5. This figure shows the total population of all bound KH states W_{KH}^{Σ} (only the two KH states corresponding to restructured $3d$ and $4d$ states of the atomic potential contribute perceptibly to this sum in the investigated range of the parameters) during the active period of the laser pulse for various radiation intensities on the lasing ''shelf.'' The plotted curves are smooth time functions and do not undergo the typical sharp oscillations, with a period equal to half the duration of the optical cycle, of the kind exhibited by the projection of the wave function $\psi(\rho,z,t)$ onto states of the unperturbed atom. Here the time rate of decay of KH states is a monotonic function of the laser intensity. The rate of decay of KH states is ob-

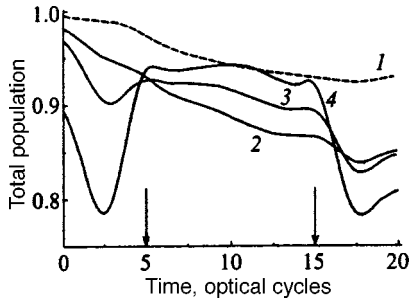

FIG. 5. Dynamics of the total population of bound KH states within the time limits of the laser pulse for various laser intensities: (*1*) 10^{14} W/cm^2; (*2*) 6×10^{14} W/cm^2; (*3*) 10^{15} W/cm^2; (*4*) 3×10^{15} W/cm^2. The arrows indicate the beginning and end of the ''shelf'' of the laser pulse. The KH states correspond to the laser intensity on the lasing ''shelf.''

served to increase as the intensity is increased to $P^*=6\times10^{14}$ W/cm^2, and then with a further increase in the intensity the KH states become increasingly stable against ionization, i.e., the stabilization effect is observed. This dynamics of the total population of KH states is completely consistent with our analytical estimates.

4. CONCLUSION

In summary, the stabilization of circular states observed in our computer calculations and in the experiments reported in Ref. 4 can be interpreted on the basis of the Kramers-Hennebergen approach. An essential consideration here is the saturation of ionization when the amplitude of the free-electron oscillations is much smaller than the characteristic length of the atomic potential, and the stationary states in the KH potential are close to those of the free atom. The interpretation of the results in the basis of free-atom states is interesting in this situation. Under conditions such that Λ-type two-photon transitions linking different states of the discrete spectrum cannot inaugurate the interference mechanism,[2,5–7] the only possible cause of saturation of the ionization probability at a level much lower than unity is transitions between states of the continuous spectrum of the system, such transitions being manifested in almost free oscillations of the electron wave packet. This interpretation of the process justifies the contemplation of a new stabilization mechanism, which would probably have an interference character, viz.: stabilization sets in as a result of coherent repopulating of the continuum by free-free transitions and interference of the amplitudes of one-photon and three-photon transitions to the continuum:

$$|n\rangle\rightarrow|E\rangle,$$

$$|n\rangle\rightarrow|E'\rangle\rightarrow|E''\rangle\rightarrow|E\rangle.$$

We also note that a further increase in the radiation intensity, on the one hand, will increase the probability of multiphoton transitions, resulting in interference stabilization, and, on the other hand, will form a KH potential, which differs significantly from the atomic potential. In this situation we are confronted with the conceptual issues of whether it is possible for the known stabilization mechanisms to exist and whether there is competition between them.

This work has received financial support from the Russian Fund for Fundamental Research (RFFI Grants No. 96-15-96447 and No. 98-02-16670). The authors are grateful to N. B. Delon and M. V. Fedorov for their interest, as well as for helpful discussions and comments.

*)E-mail: OVT@mics.msu.su

[1)]In our analytical estimates, from now on, we disregard the correction for using the smoothed Coulomb potential in the calculations.

[1] M. Gavrila and J. Kaminski, Phys. Rev. Lett. **52**, 613 (1984).
[2] M. V. Fedorov and A. M. Movsesian, J. Phys. B **21**, L155 (1988).
[3] M. P. de Boer, J. H. Hoogenraad, R. B. Vrijen *et al.*, Phys. Rev. A **50**, 4085 (1994).
[4] N. J. van Druten, R. C. Constantinescu, J. M. Schins *et al.*, Phys. Rev. A **55**, 622 (1997).
[5] M. V. Fedorov, J. Phys. B **27**, 4145 (1994).
[6] M. V. Fedorov, M.-M. Tehranchi, and S. M. Fedorov, J. Phys. B **29**, 2907 (1996).
[7] O. V. Pont, N. R. Walet, M. Gavrila, and C. W. McCurdy, Phys. Rev. Lett. **61**, 939 (1988).
[8] M. Pont and M. Gavrila, Phys. Rev. Lett. **65**, 2362 (1990).
[9] R. Grobe and M. V. Fedorov, Laser Phys. **3**, 265 (1993).
[10] Q. Su, A. Sanpera, and L. Roso-Franco, Int. J. Modern Phys. **8**, 1655 (1994).
[11] K. C. Kulander, Phys. Rev. A **35**, 445 (1987).
[12] K. C. Kulander, K. J. Shafer, and J. L. Krause, Phys. Rev. Lett. **20**, 2601 (1991).
[13] M. Dorr, R. M. Potvliege, D. Proulx, and R. Shakeshaft, Phys. Rev. A **43**, 3729 (1991).
[14] M. Gajda, B. Piraux, and K. Rzazewski, Phys. Rev. A **50**, 2528 (1994).
[15] K. Im, R. Grobe, and J. H. Eberly, Phys. Rev. A **49**, 2853 (1994).
[16] A. A. Andreev and V. E. Semenov, Opt. Spektrosk. **78**, 594 (1995) [Opt. Spectrosc. **78**, 533 (1995)].
[17] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **108**, 436 (1995) [JETP **81**, 235 (1995)].
[18] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Éksp. Teor. Fiz. **113**, 593 (1995) [JETP **86**, 328 (1995)].
[19] O. V. Tikhonova, E. A. Volkova, A. M. Popov, and M. V. Fedorov, Laser Phys. **8**, 85 (1998).
[20] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Opt. Spektrosk. (2000) [Opt. Spectrosc. (Russia) (2000)] (in press).
[21] H. Barry Bebb, J. Math. Phys. **7**, 955 (1966).
[22] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic Press, New York-London, 1958), reprinted from: *Handbuch der Physik*, edited by S. Flügge (Springer-Verlag, Berlin, 1957), Vol. 35, pp. 88–346.
[23] B. M. Smirnov, *Excited Atoms* [in Russian], Énergoatomizdat, Moscow (1982).
[24] A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **9**, 201 (1999).
[25] A. A. Krylovetsky, N. L. Manakov, and S. I. Marmo, Laser Phys. **7**, 781 (1997); A. A. Krylovetskiĭ, N. L. Manakov, and S. I. Marmo, in *Digests of the 16th Conference on Fundamental Atomic Spectroscopy* [in Russian], Zvenigorod (1998), p. 148.

Translated by James S. Wood

Stimulated Brillouin scattering of CO_2 radiation in compressed xenon

S. A. Buĭko, S. M. Kulikov, V. N. Novikov,*) and S. A. Sukharev

Russian Federal Nuclear Center Research Institute of Experimental Physics, 607190 Sarov, Nizhniĭ Novgorod Region, Russia

(Submitted 17 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 1941–1946 (December 1999)

Stimulated Brillouin scattering (STBS) and phase conjugation of CO_2 laser radiation have been demonstrated experimentally for the first time in compressed xenon (59 atm at 21 °C) located inside the low-Q cavity of this laser. The nonlinear medium was exposed to the action of counterpropagating focused multimode radiation beams. The difference between the frequencies of the longitudinal cavity modes was set at the frequency of the acoustic wave ($\nu_s = 32.2 \pm 0.3$ MHz) excited as a result of STBS by 9.584 μm radiation. The duration of the radiation pulse τ_L was close to the acoustic phonon lifetime ($\tau_L < \tau_{ph} \approx 3 \times 10^{-6}$ s). The excitation of STBS was manifested experimentally as the locking of longitudinal modes, an increase in power and energy, and also an increase in the duration of the lasing pulse and a reduction in the divergence to the diffraction limit. © *1999 American Institute of Physics.* [S1063-7761(99)00512-0]

Stimulated Brillouin scattering (STBS) is presently used for the phase conjugation of laser radiation over almost the entire wavelength range of high-power modern lasers, ranging from excimer ($\lambda_L = 0.2-0.3\,\mu$m) to chemical lasers ($\lambda_L = 2.5-4.5\,\mu$m) (Refs. 1–4). In the range of CO_2 laser radiation the main and most comprehensively studied method of obtaining phase conjugation is still four-wave mixing in nonlinear media. Compared with four-wave mixing, the use of low-threshold STBS would substantially simplify the technical problem of developing phase-conjugation devices for CO_2 laser radiation and would extend the possibilities for forming the spatial structure of laser beams. However, although STBS has been thoroughly studied in this particular spectral range, STBS in the ten micron range has only been achieved experimentally in a hydrogen plasma at high intensity $10^{11}-10^{13}$ W/cm^2 (Refs. 5 and 6). The possibilities for using this STBS in a plasma for phase conjugation have not yet been clarified. No reliable information is available on the excitation of STBS in condensed or gaseous media. This may be attributable to competition from other nonlinear phenomena, such as optical breakdown and absorption of radiation.[7,8] In the present paper we report results of an experimental investigation of the STBS of CO_2 laser radiation in compressed xenon (59 atm at $T = 21$ °C).

Compressed xenon is a fairly universal STBS medium because of its transparency over a wide wavelength range, high Stokes radiation gain, and small Stokes shift. At the selected pressure the velocity of sound in xenon has a minimum ($v_s = 1.44 \times 10^4$ cm/s, acoustic wave frequency $\nu_s = 32.2$ MHz), whereas the steady-state Stokes radiation gain has a maximum ($g \approx 1.15 \times 10^{-7}$ cm/W) (Ref. 9). However, on transition to the middle infrared and high gas pressures, the acoustic phonon lifetime increases proportionately with the density and the square of the wavelength ($\tau_{ph} \propto \rho\lambda^2$). As a result, for pulsed CO_2 lasers at the selected xenon pressure the acoustic phonon lifetime is longer than the radiation pulse duration τ_L. For $\tau_{ph} > \tau_L$ the excitation of STBS is essentially a nonsteady-state process and in order to excite STBS, the pump intensity must be increased compared with the steady-state regime which causes some deterioration in the optical breakdown situation. The threshold intensity for the excitation of essentially nonsteady-state STBS as a result of pulsed pumping of a nonlinear medium by a focused Gaussian beam (cf. Ref. 10) is given by:

$$I_{thr} \geqslant \frac{M^2}{8\pi g l_d}\frac{\tau_{ph}}{\tau_L}, \quad M \approx \ln\frac{I_L}{I_n}, \quad l_d = \frac{\lambda}{\pi}\left(\frac{F}{r_0}\right)^2,$$

where I_L and $I_n \propto 1/\lambda^4$ are the intensities of the pump and the noise Stokes radiation, respectively, λ is the pump wavelength, F is the focal length of the focusing lens, and r_0 is the radius of the Gaussian beam at the lens. In this case for $M = 36$, $\tau_L \approx \tau_{ph}$, $F = 146$ cm, $r_0 = 2.5$ cm, and $g = 1.15 \times 10^{-7}$ cm/W, the threshold intensity is $I_{thr} \approx 4 \times 10^8$ W/cm^2. The optical breakdown intensity of xenon estimated using the results of Refs. 11 and 12 is 10^7-10^8 W/cm^2. When the nonlinear medium is pumped by a focused beam, STBS of CO_2 laser radiation cannot be achieved. Quite clearly, the condition for which the threshold STBS excitation intensity is lower than the optical breakdown intensity of compressed xenon cannot be satisfied in experiments using conventional systems. We succeeded in lowering the probability of optical breakdown by significantly reducing the STBS excitation threshold. This was accomplished by increasing the input Stokes signal (reducing M) by pumping the nonlinear medium with focused counterpropagating polychromatic beams, with the frequency difference between the monochromatic components being equal to the acoustic wave frequency, achieved by placing the nonlinear medium inside the laser cavity.[13] This method of exciting STBS can reduce the threshold intensity by almost two orders of magnitude to $I_{thr} \approx 5 \times 10^6$ W/cm^2.

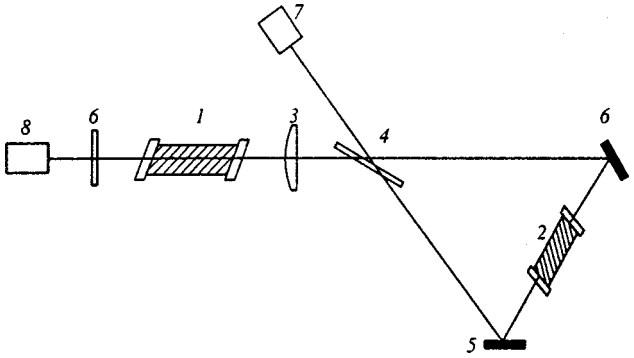

FIG. 1. Optical diagram of CO_2 laser with STBS intracavity nonlinearity: *1* — active medium, *2* — STBS cell filled with xenon to 59 atm at 21 °C, *3* — lens of 146 cm focal length, *4* — 2 cm thick plane-parallel BaF_2 plate mounted at Brewster angle, *5* — copper mirrors, *6* — plane cavity mirror, *7* — system to record radiation power and energy , and *8* — system to record radiation power, energy, divergence, and spectral composition.

The experiment is shown schematically in Fig. 1. The laser radiation source was a chemical CO_2 laser where the excitation was transferred from DF* molecules to CO_2 molecules. The laser operated on a single vibrational–rotational transition [$P(24)$, (001–020) vibrational transition]. The half-height width of the active-medium gain profile for our laser is ≈ 700 MHz. The duration of the radiation pulse τ_L was close to the phonon lifetime ($\tau_L < \tau_{ph} \approx 3\times 10^{-6}$ s). The STBS cell was a 760 mm long metal cylinder having an inner diameter of 40 mm and 40 mm thick BaF_2 windows. The windows were sealed with indium. The cell was filled with xenon via a filter with a 99.99999% filtration efficiency for impurity particles larger than 0.1 μm.

The CO_2 laser cavity was formed by a mirror *6* and a Sagnac ring interferometer. The interferometer was formed by a beam splitter *4*, in our case a plane-parallel BaF_2 plate mounted at the Brewster angle, and copper mirrors *5*. The STBS cell was installed in the focal plane of the lens *3*. The initial Q factor of the cavity was set by varying the reflection coefficient of the mirror *6*. This mirror was formed by plane-parallel BaF_2, ZnSe, and NaCl plates with a dielectric coating having a reflection coefficient of 0.28. In order to achieve efficient excitation of STBS, the optical length of the cavity L_R (length of a circular round trip) was determined from the condition that the frequency difference between the cavity longitudinal modes should be equal to the acoustic wave frequency, i.e., $L_R = L_{SBS} = c/\nu_s \approx (931 \pm 1)$ cm. Any difference between the frequency difference of the longitudinal modes and the acoustic wave frequency should be less than the half-width of the spontaneous scattering frequency profile, i.e., the cavity length should have an accuracy better than $\Delta L_{SBS} \approx L_{SBS}^2/2c\tau_{ph} = c/2\tau_{ph}\nu_s^2 \approx 5$ cm. It should be noted that by varying the cavity length, the value of τ_{ph} can be estimated from this last relationship by noting whether STBS is excited or not. When STBS is excited, the coefficient of reflection of radiation in the direction of the Sagnac interferometer will increase and consequently the cavity Q factor will increase, which should be evident from the character of the CO_2 laser radiation.

The energy, power, spectral composition, and divergence of the radiation were recorded experimentally. The

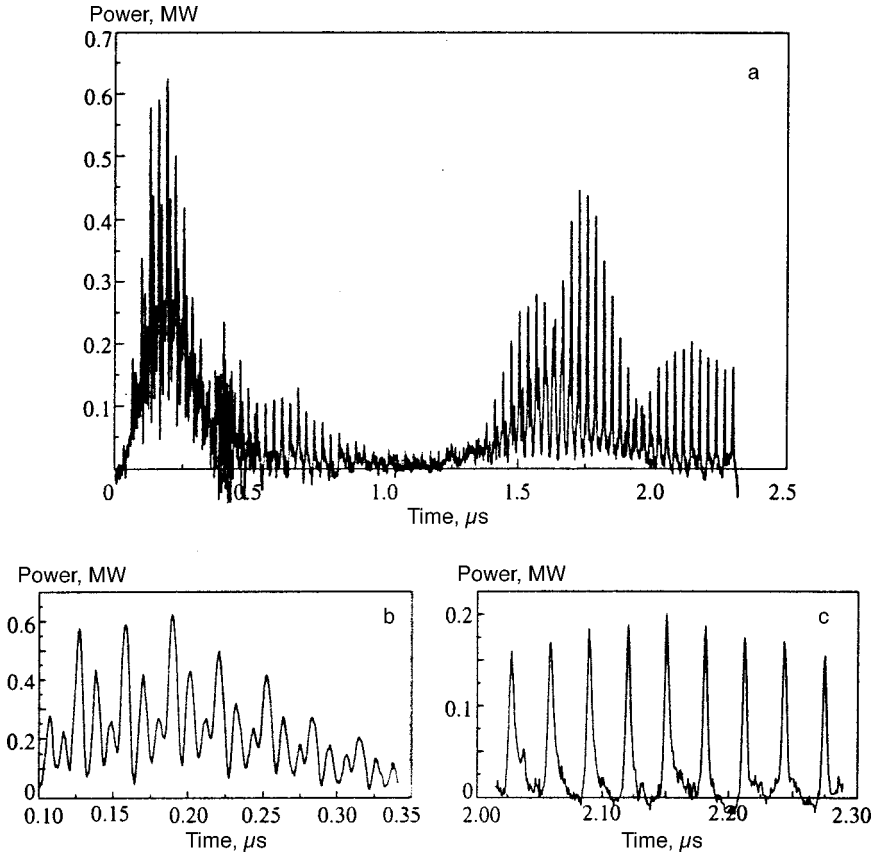

FIG. 2. Time profile of CO_2 laser radiation power with STBS excited (a) and fine structure of pulse at the beginning (b) and end of lasing (c).

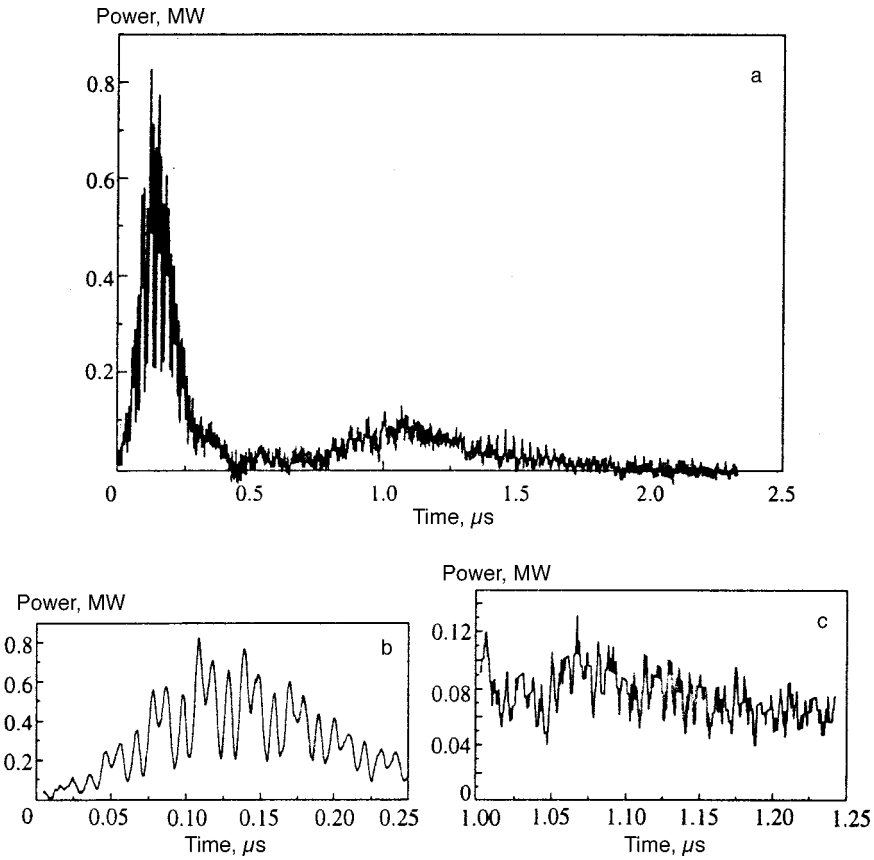

FIG. 3. Time profile of CO_2 laser radiation power in the absence of STBS (a) ($L_R = L_{SBS} - 18$ cm) and fine structure of pulse at the beginning (b) and end of lasing (c).

energy was recorded using IMO-2 calorimeters and the power was recorded using an FP-1 photodetector (≈ 0.7 ns time resolution) (Ref. 14). The photodetector signal was fed to a 500 MHz Tektronix TDS-540B oscilloscope. The divergence of the radiation was recorded using a 50% mirror wedge at the focus of the lens ($F = 146$ cm).

In experiments satisfying the condition for resonance of the Stokes radiation with the longitudinal cavity modes $L_R = L_{SBS}$ we always observed the excitation of STBS regardless of the initial cavity Q-factor. No STBS was excited in the absence of resonance when the cavity length was $L_R = L_{SBS} - 18$ cm, or $L_R = L_{SBS} + 80$ cm. When the xenon was removed from the STBS cell, no effects associated with the excitation of STBS were recorded regardless of the cavity length.

When STBS is excited, an increase in the number of modes (from four to ten) is observed during the lasing pulse and these gradually become phase-locked until complete locking is achieved. Whereas at the beginning of the pulse (Fig. 2) four-mode beats with arbitrary phases are observed, by the end of the pulse we observe a regular series of high-power 6 ns peaks with a repetition period $1/\nu_s = 31$ ns, i.e., we observe the locking of approximately ten modes. Self-mode locking at STBS nonlinearity is energetically favorable[15] since in this case all the modes are involved in creating a single acoustic grating which enhances the cavity Q factor. As a result, the increase in the Q factor with time increases the power and duration of the radiation (compare Figs. 2 and 3). The radiation energy of the CO_2 laser was doubled as a result of the excitation of STBS.

In the absence of STBS, the beginning of the pulse was also modulated by four modes with arbitrary phases but subsequently the beats diminish and the pulse profile is

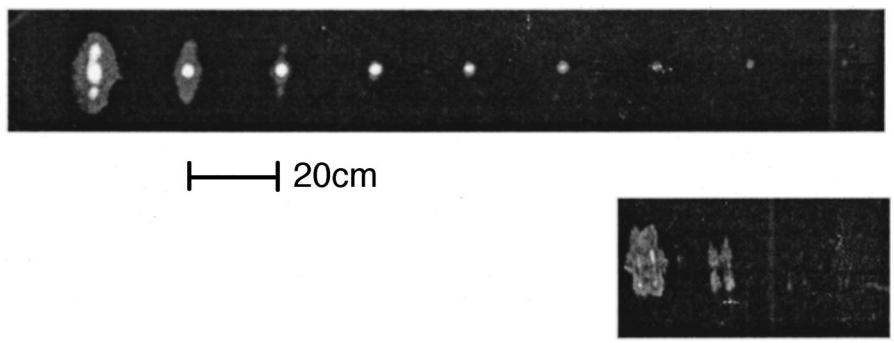

FIG. 4. Traces of far-field zone at the focus of $F = 146$ cm lens obtained using a 50% wedge: upper trace — in the presence of STBS ($L_R = L_{SBS}$), radiation energy 1.6 J; lower trace — in the absence of STBS ($L_R = L_{SBS} + 80$ cm), radiation energy 0.6 J.

smoothed as a result of an increase in the number of modes and the absence of locking, i.e., the pulse has the normal profile for a free-running CO_2 laser (Fig. 3).

The excitation of STBS and the onset of phase conjugation can be seen particularly clearly in the change in the divergence of the laser radiation (Fig. 4). Whereas in the absence of STBS the divergence of the radiation was equal to eight diffraction limits, in the presence of STBS and phase conjugation the divergence decreased almost to the diffraction limit and the brightness of the laser radiation increased 64 times. It is also important to note that when STBS is excited, the selected optical cavity system could deliver diffraction-quality radiation with a high Fresnel number ($N \approx 10$).

In this study we have demonstrated that a low-threshold STBS mirror can be developed for long-wavelength radiation where the influence of competing processes is particularly strong. However, the range of application of these results is not confined to CO_2 lasers. In particular, this laser system may prove useful for CO and oxygen–iodine lasers, and for studying various types of stimulated Rayleigh scattering processes. For lasers having a uniformly broadened gain profile the mode locking dynamics can be used to obtain additional information on the temperature of the inverted medium, the kinetics of the processes leading to the lasing effect, and also to achieve longitudinal mode selection using STBS nonlinearity.

*)E-mail: novikov@otd13.vniief.ru

[1] S. S. Alimpiev, V. S. Bukreev, S. K. Vartapetov *et al.*, Kvantovaya Élektron. (Moscow) **18**, 89 (1991) [Sov. J. Quantum Electron. **21**, 80 (1991)].
[2] M. T. Duignan, B. J. Feldman, and W. T. Whitney, Opt. Lett. **12**, 111 (1987).
[3] S. D. Velikanov, Yu. V. Dolgopolov, V. V. Egorov *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **52**, 553 (1988).
[4] E. N. Vorontsov, Yu. V. Dolgopolov, G. A. Kirillov *et al.*, Kvantovaya Élektron. (Moscow) **17**, 317 (1990) [Sov. J. Quantum Electron. **20**, 256 (1990)].
[5] A. Ng, L. Pitt, D. Salzmann, and A. A. Offenberger, Phys. Rev. Lett. **42**, 307 (1979).
[6] R. Fedosejevs, W. Tighe, D. C. D. McKen, and A. A. Offenberger, Opt. Commun. **40**, 35 (1981).
[7] N. G. Basov, V. I. Kovalev, and F. S. Faĭzulov, in *Phase Conjugation of Radiation in Nonlinear Media*, edited by V. I. Bespalov [in Russian], Institute of Applied Physics Press, Academy of Sciences of the USSR, Gorky (1982).
[8] V. V. Ragul'skiĭ, *Phase Conjugation Accompanying Stimulated Light Scattering*, [in Russian], Nauka, Moscow (1990).
[9] L. I. Zykov, S. A. Buyko, Yu. V. Dolgopolov *et al.*, in *Optical Resonators-Science and Engineering*, edited by R. Kossovsky *et al.* (Kluwer, Dordrecht, 1998), p. 233.
[10] B. Ya. Zel'dovich, N. F. Pilipetskiĭ, and V. V. Shkunov, *Principles of Phase Conjugation* (Springer-Verlag, Berlin, 1985 [Russ. original, Nauka, Moscow, 1985].
[11] Yu. P. Raĭzer, *Laser-Induced Discharge Phenomena* (Consultants Bureau, New York, 1977) [Russ. original, Nauka, Moscow, 1974].
[12] Yu. V. Dolgopolov, S. M. Kulikov, M. N. Solov'eva *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **52**, 549 (1988).
[13] A. Z. Grasyuk, V. V. Ragul'skiĭ, and F. S. Faĭzulov, JETP Lett. **9**, 6 (1969).
[14] P. M. Valov, K. V. Goncherenko *et al.*, Kvantovaya Élektron. (Moscow) **4**, 95 (1977) [Sov. J. Quantum Electron. **7**, 50 (1977).
[15] F. A. Korolev, O. M. Vokhnik, and V. I. Odintsov, JETP Lett. **18**, 32 (1973).

Translated by R. M. Durham

Angular scattering diagrams of linearly polarized relativistic-intensity electromagnetic radiation in a plasma

A. V. Borovskiĭ, A. L. Galkin, and O. B. Shiryaev*)

Institute of General Physics, Russian Academy of Sciences, 117942, Moscow, Russia
(Submitted 17 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 1947–1962 (December 1999)

Instability of the propagation of nonlinear nonmonochromatic relativistic-intensity electromagnetic waves in a cold subcritical-density plasma is analyzed in three-dimensional geometry. Angular diagrams of their scattering are presented. The calculations show that forward and backward scattering may occur. The radiation in a specific direction is a set of harmonics, propagating against a continuum background, whose frequencies depend on the angle. Radiation at a specific frequency propagates in a set of scattering cones. The azimuthal cone angles depend on frequency. © *1999 American Institute of Physics.* [S1063-7761(99)00612-5]

1. INTRODUCTION

Recently an increasing number of studies have been devoted to plasma electrodynamics[1] and in particular to the propagation of relativistic-intensity laser radiation in matter (see, for example, Refs. 2–8).

The focusing of ultrashort laser pulses gives rise to ultrahigh electromagnetic radiation intensities $I \geqslant 10^{18}\,\mathrm{W/cm^2}$. A characteristic feature of these fields is the appearance of relativistic effects in the electron motion caused by an increase in their mass. In addition, in such strong fields the material is frequently converted to a completely ionized state. Under these conditions nonlinear currents of free electrons make the main contribution to the polarization of the material[2] (a wider range of phenomena associated with the interaction between high-power electromagnetic radiation and matter was considered in Ref. 3).

Below, we consider the scattering of laser radiation in a plasma caused by the formation of nonlinear currents of free electrons. Corresponding instabilities of the electromagnetic field are already manifest in a spatially one-dimensional geometry. The scattering of radiation at relativistic intensities was analyzed in Refs. 9–13 and also by the present authors.[14–16]

The scattering of laser radiation in matter has been studied in broad terms for more than forty years but nevertheless it is extremely difficult to achieve quantitative agreement between theory and experiment, evidently because of the following circumstances. First, in an incompletely ionized material several mechanisms contribute simultaneously to the scattering process. Second, a spectral device receives radiation from different points in the scattering volume, i.e., we are dealing with an integral effect. Third, the laser radiation entering the plasma is transformed. The monochromatic photon flux emits and absorbs plasmons with the result that it is converted into a sum of monochromatic fluxes at frequencies shifted by multiples of the plasma frequency ω_p. The plasma frequency then depends on the intensity of the strong electromagnetic wave. This transformation process takes a certain time. The transformed wave is scattered which introduces an additional element of indeterminacy into the theoretical interpretation of the scattering. In the present study we consider the scattering of a transformed relativistic-intensity electromagnetic wave (Akhiezer–Polovin wave) in a completely ionized material by nonlinear currents of the plasma electron current. As in Ref. 14, we shall study the most interesting case of the scattering of a linearly polarized electromagnetic wave.

The local characteristics of the field-plasma system are studied in detail by determining the temporal instability growth rates of the electromagnetic radiation and constructing angular diagrams of its scattering. In some cases, for instance in the focal spot where the integral and local characteristics differ only slightly, the angular diagrams can be used to interpret the experimental data.

The present paper is also a continuation of Ref. 15 and 16, in which the authors analyzed a three-dimensional theory of the scattering of a circularly polarized, relativistic-intensity, monochromatic electromagnetic wave in a cold subcritical-density plasma. Below, we consider the scattering of a linearly polarized strong electromagnetic wave which is the case most frequently encountered experimentally. First, a rigorous linear analysis of this problem is made in three-dimensional geometry. Until recently the description of the scattering of relativistic-intensity electromagnetic radiation was confined to using various approximations. These include: (1) the one-dimensional approximation (see, for example, Ref. 9), (2) an approximation using a monochromatic linearly polarized reference wave which does not strictly satisfy a system of relativistic Maxwell and hydrodynamic equations; (3) searching for the growth rates assuming that one of the transverse components of the wave vector is zero; (4) resonance approximations which involve using exact phase-matching conditions for various wave processes (see, for instance, Ref. 17).

In experiments, nearly monochromatic laser radiation forms a plasma and is transformed in this plasma to give a nonlinear electromagnetic wave (nonmonochromatic Akhiezer–Polovin wave[18,16]). This nonlinear electromag-

netic wave is naturally considered as the reference solution when analyzing scattering. The results presented below show that a nonlinear Akhiezer–Polovin wave describes a regime with the establishment of a doubly-periodic solution.

Hence, this analysis is complex first, because it involves selecting a rigorous reference solution which is nonmonochromatic for a linearly polarized wave (as opposed to monochromatic for a circularly polarized wave) and second, because of the extreme strictness of the linear analysis of the growth rates performed for a system of partial differential equations with oscillating coefficients, unlike previous approximate analyses. The use of a universal approach allowed us (as for a circularly polarized wave, see Ref. 15) to describe the generation of stimulated Raman scattering (STRS) harmonics, the hydrodynamic analog of Compton scattering, continuum generation, and also the interaction of these processes for the case of a linearly polarized wave.

We note some characteristic features of this theoretical method of analysis. The temporal growth rate of this instability is defined as the maximum eigenvalue of the matrix of the linear system of ordinary differential equations being solved. This system is obtained by introducing a traveling variable along the propagation axis, taking the Fourier transform with respect to the spatial coordinates, and converting from an infinite to a finite number of coupled ordinary differential equations. In particular, this approach can avoid the conventional procedure of writing out and analyzing cumbersome dispersion equations. The periodicity of the growth rate with respect to the longitudinal component of the wave vector is substantiated below for periodic reference solutions. In an earlier study we merely drew attention to the quasiperiodicity of the numerically determined growth rates[18].

Thus, we present results of a rigorous linear analysis of a system of Maxwell and relativistic electron hydrodynamic equations for a linearly polarized reference wave of arbitrary amplitude in three-dimensional geometry. It is shown that the periodicity of the reference solution generates periodicity of the growth rate in the corresponding perturbation wave vector space and that the maximum period of the doubly-periodic rigorous reference solution determines the period of the growth rate.

2. INITIAL EQUATIONS

The propagation of relativistic-intensity laser radiation in a plasma is described by a system of Maxwell and relativistic electron hydrodynamic equations:

$$(\Delta-\partial_t^2)\mathbf{A}=\nabla\varphi_t+\frac{n}{\gamma}(\mathbf{A}+\nabla\psi), \tag{1}$$

$$\Delta\varphi=n-1, \tag{2}$$

$$\nabla\mathbf{A}=0, \tag{3}$$

$$\psi_t=\varphi-\gamma, \tag{4}$$

$$n_t+\left(\nabla\frac{n}{\gamma}(\mathbf{A}+\nabla\psi)\right)=0, \tag{5}$$

$$\gamma=\sqrt{1+|\mathbf{A}+\nabla\psi|^2}. \tag{6}$$

Here $\mathbf{A}$ and φ are the vector and scalar potentials of the electromagnetic field, ψ is the potential of the generalized momentum of the field-plasma system, and n is the electron density. The expression (6) determines the relativistic mass factor γ. The subscript t indicates a partial time derivative. The system (1)–(6) is normalized as follows: $\mathbf{A}$ and φ are normalized to mc^2/e, n is normalized to the unperturbed value of the electron density n_0, the momentum of the electron liquid is normalized to mc, the time is normalized to ω_p^{-1} (where ω_p is the unperturbed plasma frequency), and the spatial coordinates are normalized to c/ω_p.

Note that when deriving this system of equations we assumed that the generalized electron momentum has no vortex component ($\nabla\times(\mathbf{p}-\mathbf{A})\equiv 0$) (the law of frozen curl of the generalized momentum applies: if this condition is satisfied initially, it is then satisfied at all subsequent times).

The initial system of equations (1)–(6) satisfies the conservation laws for the material and the field (see, for instance, Ref. 6).

3. LINEARLY POLARIZED RELATIVISTIC-INTENSITY REFERENCE WAVE

3.1. Relativistic theory

We shall analyze the Maxwell and relativistic hydrodynamic equations for the plasma electron component (1)–(6) for $x_1=x_2=0$. As was shown by Akhiezer and Polovin,[19] the search for solutions of these equations

$$\mathbf{A}(\mathbf{x},t)=\mathbf{e}_1A_1(\xi)+\mathbf{e}_2A_2(\xi),\quad \varphi=\varphi(\xi),$$

$$n=n(\xi),\quad \psi=\psi(\xi),$$

which depend on the single variable $\xi=x_3-qt$, $q=\sqrt{1+\epsilon^2}$, where q is the phase velocity and ε is the parameter, can be reduced to solving the following problem:

$$\epsilon^2\mathbf{A}_{\xi\xi}+F(\mathbf{A},\varphi,\epsilon)\mathbf{A}=0, \tag{7}$$

$$\epsilon^2\varphi_{\xi\xi}+F(\mathbf{A},\varphi,\epsilon)\varphi-1=0, \tag{8}$$

$$F(\mathbf{A},\varphi,\epsilon)=\sqrt{\frac{1+\epsilon^2}{\varphi^2+\epsilon^2(1+|\mathbf{A}|^2)}}. \tag{9}$$

In this case, the momentum and electron density are expressed in terms of the vector and scalar potentials of the electromagnetic field[19] (see also Refs. 18 and 16).

In general, solutions of this problem can be obtained numerically. Analytic approximations were obtained in Ref. 19 for specific cases such as small-amplitude oscillations, purely longitudinal oscillations, monochromatic circularly polarized electromagnetic waves, and monochromatic oscillations having a small longitudinal component. In addition, approximate solutions were constructed in Ref. 19 for the case of greatest practical interest, when the phase velocity of the propagating electromagnetic radiation is close to the velocity of light. However, in that study the authors only analyzed the particular case when the transverse component of the electromagnetic wave is monochromatic. In this case, no nonlinear amplitude-phase modulation occurs. Claire and Perkins[20] report an analytic investigation of plane linearly polarized electromagnetic waves whose phase velocity con-

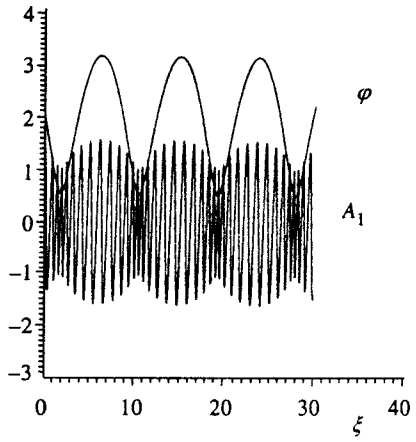

FIG. 1. Results of a numerical solution of the Akhiezer–Polovin problem (7)–(9). Initial conditions: $A_1(0)=1.2$, $A_{1_\xi}(0)=5$, $\varphi(0)=2$, $\varphi_\xi(0)=-1$, and $\epsilon=0.1$.

siderably exceeds the velocity of light. Kaw *et al.*[21] describe results of a detailed numerical investigation of solutions of the Akhiezer–Polovin equations on the phase plane. Chian and Clemmow[22] also consider small-amplitude waves corresponding to solutions of the Akhiezer–Polovin equations and in addition construct solutions of these equations in the high amplitude limit.

In previous papers[18,16] we studied the nonlinear amplitude-phase self-modulation of nonmonochromatic Akhiezer–Polovin waves. Figure 1 gives results of a numerical solution of the Akhiezer–Polovin problem[18,16] (7)–(9) using the following conditions: $A_1(0)=1.2$, $A_{1_\xi}(0)=5$, $\varphi(0)=2$, $A_2(0)=0$, $A_{2_\xi}(0)=0$ (linear polarization), $\varphi_\xi(0)=0$, and $\epsilon=0.1$. Concentration of electromagnetic radiation is observed between the crests of the electrostatic potential (amplitude modulation) and the entire pattern is shifted with the phase velocity q. The oscillation frequency of the vector potential also varies between the crests and the minimum of the scalar potential (phase modulation).

3.2. Asymptotic theory of nonlinear self-modulation of linearly polarized rf electromagnetic waves in plasma

An analytic theory of nonmonochromatic, rf, linearly polarized, relativistic-intensity electromagnetic waves in plasma (in terms of the Akhiezer–Polovin problem) was developed in Refs. 18 and 16. When the frequency of the propagating laser radiation is much higher than the plasma frequency, the phase velocity of the electromagnetic waves q is close to the velocity of light and the parameter ϵ is small. In Refs. 18 and 16 the authors used the following representation to obtain approximate asymptotic solutions for a particular parameter:

$$A_1=\sum_{m=0}^{\infty}\epsilon^m U_m(\xi,\Theta(\xi)),\quad \varphi=\sum_{m=0}^{\infty}\epsilon^m\phi_m(\xi,\Theta(\xi)),$$

$$\Theta_\xi=\epsilon^{-1}\mu(\xi).$$

Here $\mu(\xi)$ is an additional unknown function. As was shown in these studies, the function ϕ_0 only depends on ξ and is a solution of the equation

$$\phi_{0_{\xi\xi}}=\frac{1}{2}\left(\frac{1+(1/2)g^2\phi_0^{1/2}}{\phi_0^2}-1\right), \tag{10}$$

and the following relations are also satisfied:

$$U_0(\xi)=g\phi_0^{1/4}(\xi)\sin\Theta, \tag{11}$$

$$\mu(\xi)=\phi_0^{-1/2}. \tag{12}$$

The constant g is the coupling parameter between the electromagnetic wave and the Langmuir response of the plasma.

For Eq. (10) the following conservation law is satisfied:

$$\phi_{0_\xi}^2+V(\phi_0)=E, \tag{13}$$

$$V(\phi_0)=V_p(\phi_0)+V_l(\phi_0), \tag{14}$$

$$V_p(\phi_0)=\phi_0+\phi_0^{-1}, \tag{15}$$

$$V_l(\phi_0)=g^2\phi_0^{-1/2}. \tag{16}$$

In Refs. 18 and 16 the authors also give the corresponding principal terms of the asymptotic forms for the density and the longitudinal component of the electron momentum.

Calculations of higher-order approximations in terms of the small parameter yield the following results:

$$U_1(\xi,\Theta)=a_1(\xi)\cos\Theta,$$

$$a_1(\xi)=\frac{\phi_0^{1/4}}{2}\int\frac{g(g^2\phi_0^{1/2}-3(E\phi_0-3\phi_0^2+1))}{16\phi_0^{5/2}}d\xi,$$

$$\phi_1\equiv 0,$$

$$U_2(\xi,\Theta)=\frac{g^3}{128\phi_0^{5/4}(\xi)}\sin 3\Theta+a_2(\xi)\cos(\Theta+\text{const}),$$

$$\phi_2(\xi,\Theta)=\frac{g^2}{16\phi_0^{1/2}(\xi)}\cos 2\Theta,$$

where the last term in the equation for U_2, which corresponds to a small correction to the solution ''at the fundamental frequency'' should be determined from the following approximations for ϵ. Thus, the second-order corrections describe second and third harmonic generation at the local frequency of the propagating electromagnetic wave.

An integral expression for the period of the slow oscillations generated by the propagating electromagnetic radiation follows explicitly from the relationships (13)–(16):

$$T(E,g)=\oint\frac{d\phi}{\sqrt{E-V(\phi)}}.$$

We can also introduce the concept of the average wave number for the rf electromagnetic wave::

$$k_0=(T\epsilon)^{-1}\oint\phi_0^{-1/2}d\xi.$$

In this case, as was shown earlier,[18,16] we obtain the simple dispersion relationship:

$$k_0 = Nk_p, \quad k_p = \frac{2\pi}{T}.$$

The normalized intensity of an Akhiezer–Polovin wave averaged over the rf oscillations of the electromagnetic field and the plasmon period, is given by:[18,16]

$$\overline{I} = \frac{g^2 k_0}{2\epsilon}.$$

3.3. Quasirelativistic approximation

It is possible to have a situation where the electromagnetic waves are such that their propagation in a plasma may be described using a quasirelativistic approximation: when relativistic corrections to the masses of the oscillating electrons are taken into account, we can confine our analysis to an expansion of the gamma factor in powers of the ratio of the electron momentum to mc (in equations which have not been reduced to dimensionless form) and we can neglect corrections of the fourth order or higher (it should then be borne in mind that in this particular case, the phase velocity of the electromagnetic radiation is higher than the velocity of light). The mathematically most rigorous implementation of this approach can be found in Ref. 23, where the authors introduce a small parameter associated with the deviation of the electromagnetic field and plasma parameters from the equilibrium values and by means of expansions in terms of this small parameter, they derive a corresponding nonsteady-state nonlinear problem (see also Ref. 24). Below we consider the steady-state particular case, the quasirelativistic analog of the Akhiezer–Polovin problem, which is obtained assuming that the partial time derivatives in the problem described in Refs. 23 and 24 are zero.

Mori *et al.*[25] used a quasirelativistic approximation to study the instabilities of circularly polarized monochromatic electromagnetic radiation in a plasma.

A quasirelativistic analog of the Akhiezer–Polovin problem (considered below for the case of a linearly polarized electromagnetic wave) is also obtained from the ''complete'' relativistic problem (7)–(9) if the nonlinearities in these equations are expanded in terms of ε, which corresponds to the propagation of rf electromagnetic waves, the scalar potential is expressed in the form

$$\varphi = 1 + f, \tag{17}$$

it is assumed that $f \sim A_1^2$, and corrections of the order A_1^4 are neglected. The corresponding problem has the form

$$\epsilon^2 A_{1_{\xi\xi}} + (1-f)A_1 = 0,$$

$$f_{\xi\xi} + f = \frac{1}{2}A_1^2.$$

For a given external field the second of these equations becomes a linear problem to determine the scalar potential and is solved[4] using the Green function (provided that the integral converges):

$$f(\xi) = \int_{-\infty}^{\infty} \sin|\xi - \xi'| \frac{A_1^2(\xi')}{2} d\xi'.$$

An expansion in terms of the small parameter ε, performed as for the relativistic problem considered in the previous section, yields the following results:

$$A_1 = g(1-f)^{-1/4} \sin\left(\epsilon^{-1} \int (1-f)^{1/2} d\xi\right) + O(\epsilon^2), \tag{18}$$

$$f_{\xi\xi} + f = \frac{1}{4} g^2 (1-f)^{-1/2}.$$

The second of these equations corresponds to the conservation law

$$f_\xi^2 + V_{qr}(f) = E', \tag{19}$$

$$V_{qr}(f) = f^2 + g^2(1-f)^{1/2}. \tag{20}$$

These approximate relationships can be obtained from the corresponding expressions in relativistic theory. Then, substituting (17) into the expressions (11), and (12) for the vector-potential obtained in Sec. 3.2 of this study we obtain, after making an elementary transformation:

$$A_1 = g\left(\frac{1-f^2}{1-f}\right)^{1/4} \sin\left(\epsilon^{-1} \int \left(\frac{1-f^2}{1-f}\right)^{-1/2} d\xi\right) + O(\epsilon).$$

Neglecting the value of f^2 in accordance with the quasirelativistic approximation, we quite clearly arrive at the relationship (18). We then expand

$$V_p(\varphi) = V_p(1+f) \approx 2 + f^2$$

and use the obvious representation

$$V_l(\varphi) = V_l(1+f) = g^2 \sqrt{\frac{1-f}{1-f^2}}.$$

Again neglecting f^2 in this last equation, from the relativistic conservation law (13)–(16) for the average vector-potential we obtain the quasirelativistic conservation law (19) and (20).

4. SCATTERING EQUATIONS FOR A LINEARLY POLARIZED WAVE

As the reference solution we take an arbitrary linearly polarized Akhiezer–Polovin wave (see Fig. 1). This is an exact numerical solution of the system (1)–(6) described by the functions $\mathbf{A}_0$, φ_0, n_0, $\mathbf{p}_0 = \mathbf{e}_3 p_{3,0}$ (we give these the subscript ''0''). Let us assume that the small perturbations $\delta\mathbf{A}$, $\delta\varphi$, δn, and $\delta\mathbf{p} = \nabla\delta\psi$ propagate against the background of this reference wave. The system of linearized equations for the perturbations has the form

$$\mathbf{A} = \mathbf{A}_0 + \delta\mathbf{A}, \quad \varphi = \varphi_0 + \delta\varphi,$$

$$n = n_0 + \delta n, \quad \mathbf{p} = \mathbf{p}_0 + \delta\mathbf{p},$$

$$(\Delta - \partial_t^2)\delta\mathbf{A} = \nabla\delta\varphi_t + \frac{n_0}{\gamma_0}(\delta\mathbf{A} + \delta\mathbf{p}) + \frac{1}{\gamma_0}(\mathbf{A}_0 + \mathbf{p}_0)\delta n$$

$$- \frac{n_0}{\gamma_0^3}(\mathbf{A}_0 + \mathbf{p}_0)((\mathbf{A}_0 + \mathbf{p}_0)\cdot(\delta\mathbf{A} + \delta\mathbf{p})),$$

$$\nabla\mathbf{A} = 0, \quad \Delta\delta\varphi = \delta n,$$

$$\delta\mathbf{p}_t=\nabla(\delta\varphi-\gamma_0^{-1}(\mathbf{A}_0+\mathbf{p}_0)\cdot(\delta\mathbf{A}+\delta\mathbf{p})),$$

$$\delta n_t+\nabla\left(\frac{n_0}{\gamma_0}(\delta\mathbf{A}+\delta\mathbf{p})+\frac{1}{\gamma_0}(\mathbf{A}_0+\mathbf{p}_0)\delta n-\frac{n_0}{\gamma_0^3}(\mathbf{A}_0+\mathbf{p}_0)\cdot((\mathbf{A}_0+\mathbf{p}_0)\cdot(\delta\mathbf{A}+\delta\mathbf{p}))\right)=0.$$

We introduce the notation for the functions appearing in these equations:

$$f^1(\xi)=\frac{n_0(\xi)}{\gamma_0(\xi)},\quad f^2(\xi)=\frac{A_{1,0}(\xi)}{\gamma_0(\xi)},\quad f^3(\xi)=\frac{p_{3,0}(\xi)}{\gamma_0(\xi)},$$

$$f^4(\xi)=\frac{n_0(\xi)A_{1,0}(\xi)}{\gamma_0^2(\xi)},\quad f^5(\xi)=\frac{n_0(\xi)p_{3,0}(\xi)}{\gamma_0^2(\xi)}.$$

These functions are periodic and can be expanded as Fourier series:

$$f^j(\xi)=\sum_m f_m^j\exp(imk_p\xi). \tag{21}$$

$$f_m^4=\sum_j f_j^1 f_{m-j}^2,\quad f_m^5=\sum_j f_j^1 f_{m-j}^3.$$

Substituting Eq. (21) into the linearized equations given above and using the Fourier transformation:

$$(\delta\mathbf{A},\delta\varphi,\delta n,\delta\mathbf{p})^T=(2\pi)^{-3/2}\int\exp(-i\mathbf{k},(x_1,x_2,\xi))\times(\delta\mathbf{A},\delta\varphi,\delta n,\delta\mathbf{p})_{\mathbf{k}}^T d^3k,$$

we obtain the following system of equations to describe the instability: $(\mathbf{k}==(k_1,k_2,\chi))$:

$$-(|\mathbf{k}|^2+(\partial_t+iq\chi)^2)\delta A_{1_0}=\sum_m(a_m^1\delta A_{1_{-m}}+a_m^2\delta A_{2_{-m}}+a_m^3\delta\psi_{-m}+a_m^4\delta n_{-m}), \tag{22}$$

$$-(|\mathbf{k}|^2+(\partial_t+iq\chi)^2)\delta A_{2_0}=\sum_m(b_m^1\delta A_{1_{-m}}+b_m^2\delta A_{2_{-m}}+b_m^3\delta\Psi_{-m}+b_m^4\delta n_{-m}), \tag{23}$$

$$(\partial_t+iq\chi)\delta n_0=\sum_m(c_m^1\delta A_{1_{-m}}+c_m^2\delta A_{2_{-m}}+c_m^3\delta\Psi_{-m}+c_m^4\delta n_{-m}), \tag{24}$$

$$(\partial_t+iq\chi)\delta\psi_0=\sum_m(d_m^1\delta A_{1_{-m}}+d_m^2\delta A_{2_{-m}}+d_m^3\delta\Psi_{-m}+d_m^4\delta n_{-m}), \tag{25}$$

where

$$a_m^1=\left(1-\frac{k_1^2mk_p}{|\mathbf{k}|^2(\chi+mk_p)}\right)f_m^1-\sum_l\left(\left(1-\frac{k_1^2}{|\mathbf{k}|^2}\right)f_l^4-\frac{k_1\chi}{|\mathbf{k}|^2}f_l^5\right)\left(f_{m-l}^2-\frac{k_1}{\chi+mk_p}f_{m-l}^3\right),$$

$$a_m^2=-\frac{k_1k_2mk_p}{|\mathbf{k}|^2(\chi+mk_p)}f_m^1+\sum_l\left(\left(1-\frac{k_1^2}{|\mathbf{k}|^2}\right)f_l^4-\frac{k_1\chi}{|\mathbf{k}|^2}f_l^5\right)\frac{k_2}{\chi+mk_p}f_{m-l}^3,$$

$$a_m^3=-\frac{ik_1\chi mk_p}{|\mathbf{k}|^2}f_m^1+i\sum_l\left(\left(1-i\frac{k_1^2}{|\mathbf{k}|^2}\right)f_l^4-i\frac{k_1\chi}{|\mathbf{k}|^2}f_l^5\right)\times(k_1f_{m-l}^2-(\chi+mk_p)f_{m-l}^3),$$

$$a_m^4=f_m^2-\frac{k_1}{|\mathbf{k}|^2}(k_1f_m^2+\chi f_m^3),$$

$$b_m^1=-\frac{k_1k_2mk_p}{|\mathbf{k}|^2(\chi+mk_p)}f_m^1+\sum_l\frac{k_2}{|\mathbf{k}|^2}(k_1f_l^4+\chi f_l^5)\times\left(f_{m-l}^2-\frac{k_1}{\chi+mk_p}f_{m-l}^3\right),$$

$$b_m^2=\left(1-\frac{k_2^2mk_p}{|\mathbf{k}|^2(\chi+mk_p)}\right)f_m^1-\sum_l\frac{k_2^2}{|\mathbf{k}|^2(\chi+mk_p)}(k_1f_l^4+\chi f_l^5)f_{m-l}^3,$$

$$b_m^3=i\frac{k_2\chi mk_p}{|\mathbf{k}|^2}f_m^1-i\frac{k_2}{|\mathbf{k}|^2}\sum_l(k_1f_l^4+\chi f_l^5)(k_1f_{m-l}^2+(\chi+mk_p)f_{m-l}^3),$$

$$b_m^4=-\frac{k_2}{|\mathbf{k}|^2}(k_1f_m^2+\chi f_m^3),$$

$$c_m^1=i\frac{k_1mk_p}{\chi+mk_p}f_m^1-i\sum_l(k_1f_l^4+\chi f_l^5)\times\left(f_{m-l}^2-\frac{k_1}{\chi+mk_p}f_{m-l}^3\right),$$

$$c_m^2=i\frac{k_2mk_p}{\chi+mk_p}f_m^1+i\sum_l\frac{k_2}{\chi+mk_p}(k_1f_l^4+\chi f_l^5)f_{m-l}^3,$$

$$c_m^3=(|\mathbf{k}|^2+\chi mk_p)f_m^1-\sum_l(k_1f_l^4+\chi f_l^5)(k_1f_{m-l}^2+(\chi+mk_p)f_{m-l}^3),$$

$$c_m^4=i(k_1f_m^2+\chi f_m^3),$$

$$d_m^1=-f_m^2+\frac{k_1}{\chi+mk_p}f_m^3,\quad d_m^2=\frac{k_2}{\chi+mk_p}f_m^3,$$

$$d_m^3=i(k_1f_m^2+(\chi+mk_p)f_m^3),\quad d_m^4=-|\mathbf{k}|^{-2}\delta_{m,0}.$$

After being converted to the standard form, the equation to determine the eigenvalues of the matrix B of the system of ordinary differential equations obtained from the equations presented above has the form

$$\det|B-GI|=0$$

and remains unchanged when the argument χ shifts by k_p since the unbounded matrix B transforms back to itself. Consequently, all the eigenvalues

$$G_m=\mathrm{Re}\,G_m+i\,\mathrm{Im}\,G_m$$

of the matrix B are periodic functions of the argument χ with the period k_p. The growth rate of the problem is defined as the maximum positive real part of the eigenvalues of the matrix B:

$$G=\max \mathrm{Re}\,G_m.$$

5. SCATTERING OF A LINEARLY POLARIZED WAVE IN THE ONE-DIMENSIONAL CASE

Borovskiĭ *et al.*[14] investigated the instability of a linearly polarized wave using the system given above in the spatially one-dimensional case, i.e., when $k_1=k_2=0$. These authors established that a linearly polarized nonmonochromatic wave is unstable with respect to forward–backward scattering in the relativistic intensity range. Calculations of specific variants showed that the scattering spectrum in the wave vector space consists of a periodic set of lines comprising STRS harmonics shifted relative to one another by multiples of k_p. The profile of a particular growth rate line is also influenced by a relativistic hydrodynamic analog of the Compton scattering of photons by electron propagating in the field of the reference wave at velocities close to the velocity of light. The profile of the line wings is attributed to the generation of Langmuir noise by the propagating reference wave.

Thus, a plane wave in a plasma is unstable even in one-dimensional geometry.

A comparison between the results of a rigorous linear analysis of the propagation instability of plane (nonmonochromatic) linearly polarized and (monochromatic) circularly polarized strong waves in a cold plasma reveals substantial differences in the growth rate lines. Both the period between the neighboring harmonics in **k** space (k_0 for circular and k_p for linear polarizations) and the structure of the lines differ. For circular polarization the hydrodynamic analog of the Compton effect does not have a significant influence in one-dimensional geometry whereas for linear polarization, this effect forms the line center. In the nonrelativistic range the growth rate line splits into the Stokes and anti-Stokes components for the circular polarization whereas for the linear polarization all the STRS harmonics are represented.

6. ANGULAR SCATTERING DIAGRAMS FOR A LINEARLY POLARIZED WAVE

The maximum growth rate for the problem (22)–(25) was investigated as a function of three components of the perturbation wave vector **k**. Only a graph of the growth rate as a function of two variables can give a clear representation. It is interesting to examine the dependence of the growth rate

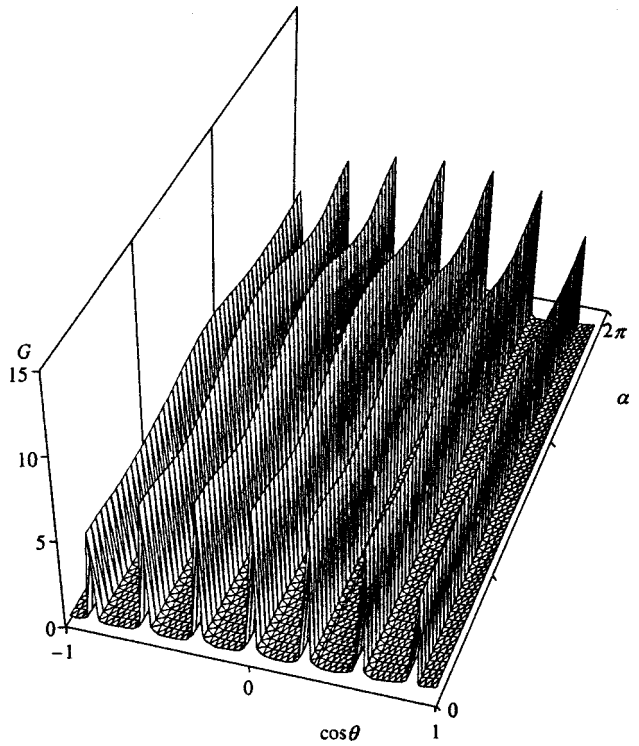

FIG. 2. Distribution of the growth rate G as a function of $(\cos\theta,\alpha)$ for $k=3.5k_p$ and the reference wave shown in Fig. 1.

on the variables in a spherical coordinate system. In Cartesian coordinates $(\mathbf{e}_1,\mathbf{e}_2,\mathbf{e}_3)$ the reference wave propagates along the $\mathbf{e}_3$ axis and is polarized along the $\mathbf{e}_1$ axis. The vector **k** of the scattered wave forms the angle θ with the $\mathbf{e}_3$ axis and has the modulus k. The vector $\mathbf{k}_\perp$, being the projection of the vector **k** on the plane $(\mathbf{e}_1,\mathbf{e}_2)$, forms the angle α with the $\mathbf{e}_1$ axis. The angles α,θ are the polar and azimuthal angles in spherical coordinates. We shall analyze the distributions of the growth rate with a fixed radius k as a function of the angles θ and α. The value of k determines the frequency of the scattered radiation $\omega=k/c$ and the angles θ and α determine the direction of scattering in spherical coordinates. Since the experimental measurements are made at different angles θ and α using spectral devices tuned to a specific frequency ω, the diagrams calculated below can be used directly to interpret the experimental data.

Figure 2 shows the distribution of G as a function of $\cos\theta$ and α for $k=3.5k_p$ and the reference wave shown in Fig. 1. The scattering structure has the following characteristic features. The scattering takes place in a discrete range of nested cones. Within each cone the scattering is anisotropic in terms of the polar angle. The scattering angle is symmetric with respect to $\theta=\pi/2$ (forward–backward) and with respect to the polarization plane $\alpha=0$ (right-left).

Figure 3 gives dependences of G on α for two fixed scattering cones $\cos\theta=mk_p/k$, where $m=1,8$ and $k=8.5k_p$ (the reference wave is the same as for Fig. 2). This diagram gives an idea of the azimuthal scattering anisotropy: the scattering predominates for $\alpha=0,\pi$, i.e., in the direction of polarization of the reference wave.

Figure 4 gives G as a function of $\cos\theta$ for the polar

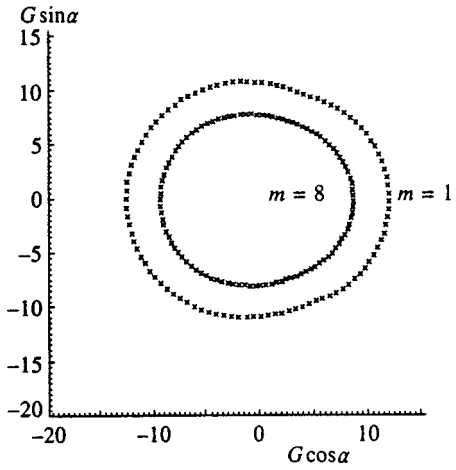

FIG. 3. Dependence of the growth rate G on α for two fixed scattering cones $\cos\theta = mk_p/k$, where $m = 1,8$ and $k = 8.5k_p$ (the reference wave is the same as that for Fig. 2).

angle $\alpha = 0$ corresponding to the scattering maximum for various values of $k/k_p = 0.5$ (a), 1.0 (b), 3.5 (c), and 8.5 (d). This figure shows the change in the structure of the angular scattering diagram as k increases (scattered radiation frequencies). For values of k which are multiples of k_p (Fig. 4b) narrow cones of scattered radiation appear in the direction of the angles $\theta = 0, \pi$ (forward–backward), which correspond to the hydrodynamic analog of Compton scattering. A further increase in k shifts the spikes toward the center (on the scale in Fig. 4) and leads to the formation of a pair of new spikes each time k passes through a value which is a multiple of k_p. As a result, in Fig. 4d we have eight pairs of spikes and one central one. The coordinates of the maxima of all the spikes are given by

$$\cos\theta = mk_p/k, \quad m = 0, \pm 1, \pm 2, \ldots$$

We note that the structure of the growth rate is a continuous background which provides evidence of continuum generation, with a set of discrete spikes. These are attributable to STRS at plasmons and the hydrodynamic analog of Compton scattering.

7. CONCLUSIONS

A spatially three-dimensional geometry has been used to make a rigorous linear analysis of the propagation instability of a plane linearly polarized strong electromagnetic wave in a plasma. The nature of this wave is described: this wave is not monochromatic which significantly complicates the analysis. The wave is a doubly periodic solution of relativistic Maxwell and electron hydrodynamic equations (an Akhiezer–Polovin wave). From the physical point of view, this amplitude-phase modulated wave is a set of photon fluxes at frequencies shifted by multiples of the plasma frequency.

An Akhiezer–Polovin wave was used as the reference solution to analyze the scattering problem. The theory describes the following wave processes: generation of STRS harmonics of the propagating reference radiation in the nonlinear medium, a hydrodynamic analog of Compton scattering as a result of an electron recoil effect, electromagnetic wave interaction in the plasma, and also continuum genera-

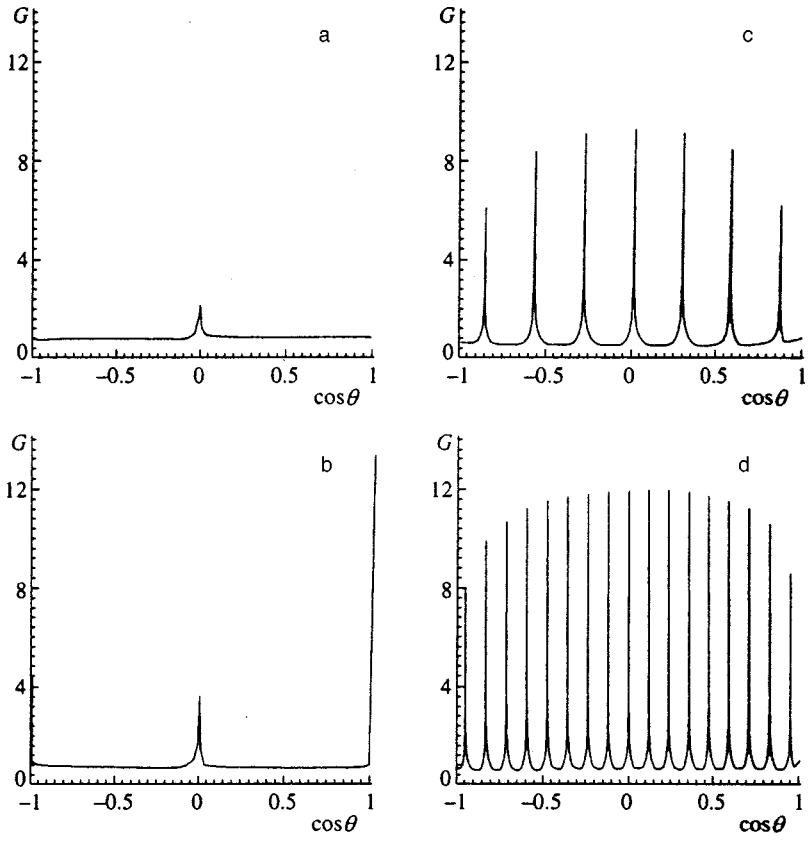

FIG. 4. Dependences of the growth rate G on $\cos\theta$ for the polar angle $\alpha = 0$ corresponding to the scattering maximum, for various values of k/k_p: 0.5 (a), 1.0 (b), 3.5 (c), and 8.5 (d).

tion as a result of the scattering of laser radiation. The combination of these wave processes was studied in the relativistic case.

In the present study we used a relativistic hydrodynamic model of a cold plasma in an electromagnetic field. This model is valid when the energy of the electron oscillations in the field substantially exceeds their thermal energy. We can use $I \geqslant 2\times 10^{16}\,\mathrm{W/cm^2}$ (see Ref. 6) as an estimate of the validity of the model.

The ''model'' nature of the analysis also involves excessive allowance for long-wavelength perturbations, which leads to the appearance of divergence in the spectral lines of the STRS harmonics. The overall scattering pattern (the frequencies and directions of the scattered waves, the position of the scattering cones, and the scattering anisotropy) are almost independent of this circumstance. This model cannot be used to calculate the narrow region of spectral profiles near the maxima. In order to eliminate the contribution of long-wavelength perturbations, we can propose several procedures: allowance for thermal effects, including shielding of the potential, allowance for the bounded length of the laser radiation pulse, and thus the dimensions of the plasma region, and so on. In the numerical investigations the presence of characteristic features was taken into account using special nonuniform grids.

The calculations show that forward and backward scattering may occur. Each radiation flux can be expanded as an angular spectrum. The radiation in a specific direction is a set of harmonics propagating against a continuum background, whose frequencies depend on the angle. However, the radiation flux has a continuous frequency spectrum. Radiation at a specific frequency propagates in a set of scattering cones. The polar angles of the cones depend on frequency.

Calculations of the angular distributions show that the scattering growth rate is symmetrical relative to the polarization plane $(\mathbf{e}_1, \mathbf{e}_3)$ of the reference wave and relative to the plane perpendicular to the direction of propagation of the reference wave. Despite the symmetry of the growth rate under the experimental conditions the backscattering intensity is low because of the short interaction time of the radiation traveling in opposite directions[26]. At the same time, the growth rate exhibits polar anisotropy. The preferential directions of scattering lie in the plane of polarization of the reference wave.

The theory also describes a hydrodynamic analog of Compton scattering. This scattering predominates in the forward–backward directions and is a set of harmonics at frequencies which are multiples of the plasma frequency. An electron rotating along a relativistic figure-eight in the plane of polarization of the reference wave possesses a momentum parallel and anti-parallel to the momentum of the field as it passes through the upper and lower ''tips'' of the figure-eight. In the corresponding time intervals, effective scattering takes place where the electron loses or acquires momentum.

Since linearly polarized radiation is predominantly used in experiments with relativistic-intensity pulses, the angular distributions obtained in this study can be used to interpret the experimental data.

To sum up, the instability of an Akhiezer–Polovin wave has been investigated for the first time in multidimensional geometry. Note that the results presented are naturally converted to the results of the one-dimensional theory.[14] It is shown that the growth rate is periodic with respect to the longitudinal component of the perturbation wave vector. The condition for periodicity of the growth rate can be applied to any other periodic reference wave, including an approximate one.

This work was partly financed by INTAS (Grant No. 94-1937) and by the RFBR (Grants Nos. 96-02-18264 and 99-02-18242).

*)E-mail: obs@kapella.gpi.ru

[1] A. F. Aleksandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Springer-Verlag, Berlin, 1984) [Russ. original, later ed., Vysshaya Shkola, Moscow, 1988].
[2] N. L. Tsintsadze and D. D. Tskhakaya, *Relativistic Nonlinear Effects in Plasmas* [in Russian], Metsniereba, Tbilisi (1989).
[3] B. Luther-Davis, E. G. Gamaliĭ, Y. Wang, A. V. Rode, and V. T. Tikhonchuk, Kvantovaya Élektron. (Moscow) **19**, 317 (1992) [Sov. J. Quantum Electron. **22**, 289 (1992)].
[4] V. M. Gorbunov and V. I. Kirsanov, Tr. Fiz. Inst. Akad. Nauk SSSR **219**, 3 (1992).
[5] S. V. Bulanov, I. N. Inovenkov, V. I. Kirsanov, N. M. Naumova, and A. S. Sakharov, Phys. Fluids B **4**, 1935 (1992).
[6] A. V. Borovskiĭ and A. L. Galkin, *Laser Physics* [in Russian], IzdAT, Moscow (1996).
[7] R. N. Sudan, Y. S. Dimant, and O. B. Shiryaev, Phys. Plasmas **4**, 1489 (1997).
[8] A. V. Borovskiĭ and O. B. Shiryaev, Zh. Éksp. Teor. Fiz. **110**, 865 (1996) [JETP **83**, 475 (1996)].
[9] V. I. Kirsanov and A. S. Sakharov, Fiz. Plazmy **21**, 623 (1995) [Plasma Phys. Rep. **21**, 587 (1995)].
[10] A. S. Sakharov and V. I. Kirsanov, Fiz. Plazmy **21**, 632 (1995) [Plasma Phys. Rep. **21**, 596 (1995)].
[11] A. S. Sakharov and V. I. Kirsanov, Phys. Plasmas **4**, 3382 (1997).
[12] B. Quesnel, P. Mora, J. C. Adam, A. Heron, and G. Laval, Phys. Plasmas **4**, 3358 (1997).
[13] B. Quesnel, P. Mora, J. C. Adam, S. Guerin, A. Heron, and G. Laval, Phys. Rev. Lett. **78**, 2132 (1997).
[14] A. V. Borovskiĭ, A. L. Galkin, V. V. Korobkin, and O. B. Shiryaev, Kvantovaya Élektron. (Moscow) **26**, 54 (1999).
[15] A. V. Borovskiĭ, A. L. Galkin, V. V. Korobkin, and O. B. Shiryaev, Zh. Éksp. Teor. Fiz. **113**, 2034 (1998) [JETP **86**, 1114 (1998)].
[16] A. V. Borovsky, A. L. Galkin, V. V. Korobkin, and O. B. Shiryaev, Phys. Rev. E **59**, 2253 (1999).
[17] C. J. McKinstrie, A. Simon, and E. A. Williams, Phys. Fluids **27**, 2738 (1984).
[18] A. V. Borovskiĭ, A. L. Galkin, and O. B. Shiryaev, Kratk. Soobshch. Fiz. **5**, 33 (1998).
[19] A. I. Akhiezer and R. V. Polovin, Zh. Éksp. Teor. Fiz. **30**, 915 (1956) [Sov. Phys. JETP **3**, 696 (1956)].
[20] M. Claire and F. Perkins, Phys. Rev. Lett. **276**, 1342 (1971).
[21] P. K. Kaw, A. Sen, and E. J. Valeo, Physica D **9**, 96 (1983).
[22] A. C.-L. Chian and P. C. Clemmow, J. Plasma Phys. **14**, 505 (1975).
[23] H. H. Kuehl, C. Y. Zhang, and T. Katsouleas, Phys. Rev. E **47**, 1249 (1993).
[24] H. H. Kuehl and C. Y. Zhang, Phys. Rev. E **48**, 1316 (1993).
[25] W. B. Mori, C. D. Decker, D. E. Hinkel, and T. Katsouleas, Phys. Rev. Lett. **72**, 1482 (1994).
[26] A. V. Borovskiĭ, A. L. Galkin, V. V. Korobkin, and O. B. Shiryaev, Kvantovaya Élektron. (Moscow) **24**, 929 (1997).

Translated by R. M. Durham

Characteristics of the coherent excitation of an impurity atom in a photonic crystal

A. M. Basharov*)

Moscow Engineering-Physics Institute, 115409 Moscow, Russia
(Submitted 6 July 1999)
Zh. Éksp. Teor. Fiz. **116**, 1963–1978 (December 1999)

It is shown that when an optically allowed transition is coherently excited in an impurity atom in a photonic crystal under conditions where one of the normal relaxation channels is suppressed by the spectral characteristics of the photonic crystal, new relaxation mechanisms are activated involving a coherent field quantum. These mechanisms substantially alter the dynamics of the atomic system, leading to filling of levels of the impurity atom which do not belong to the coherently excited atomic transition. Under certain conditions this leads to population inversion as a result of an optically allowed transition which does not interact with the coherent pump and at a frequency where no photonic band gaps can exist. © *1999 American Institute of Physics.* [S1063-7761(99)00712-X]

1. INTRODUCTION

The fabrication technology being developed for photonic crystals,[1–3] i.e., media having forbidden bands in photonic states, has stimulated interest in the dynamics of excited impurity atoms whose radiative transition frequencies lie in forbidden bands. Media have now been obtained in which the width of the photonic band gap is of the order of 30% of the gap-center frequency and the gaps are situated in ranges between the microwave and the visible. Some of the main examples of photonic crystals are periodic dielectric structures. The specific characteristics of the propagation of electromagnetic radiation in periodic dielectric structures find applications in various important devices such as microlasers,[1,4,5] various types of optical fibers,[6] and so on. In addition, from the general physical point of view, photonic crystals provide a unique possibility of controlling the ''switching on'' or ''switching off'' of various fundamental electromagnetic interactions. Thus, photonic crystals not only have various technical applications but are also of enormous interest for fundamental research, including studies of the dynamics of impurity atoms in which photonic crystals can achieve various scenarios where the frequency of a particular atomic transition falls within the photonic band gap. As a result, some interaction channels are suppressed and others are activated.

Some work on the dynamics of impurity atoms in photonic crystals[7–13] has been devoted to mechanisms for the relaxation of impurity atoms whose transition frequencies lie in the photonic band gap. Sohn and Wang[7] studied a quadrupole mechanism for the radiative relaxation of excited impurity atoms. In Refs. 7–9 the authors studied a mechanism for resonant dipole–dipole interaction of impurity atoms and the concept of localized photons. Rupasov and Singh[10] proposed a model for spontaneous emission from impurity atoms which can be solved analytically. In a series of studies, Rupasov and Singh[11,12] consider the influence of dipole–dipole interaction between impurity atoms on the appearance of the photonic band gap and on the superradiance of an ensemble of like impurity atoms. Basharov[13] discusses an electric-dipole diatomic mechanism for the spontaneous emission of an impurity atom.

Other studies have considered the interaction between coherent radiation and impurity atoms at whose transition frequency no photons occur. John and Quang[14] investigated optical bistability and phase transitions when the carrier frequency of the coherent wave and the excited level of the impurity atoms were situated far (in both directions) from the edge of the photonic band gap. Quang *et al.*[15] studied the spontaneous emission of a three-level system in which the frequency of one optically allowed transition was located in the photonic band gap and a neighboring optically forbidden transition was pumped by a coherent wave. In these and other studies the coherent pumping ''operated'' under conditions where particular relaxation processes discussed earlier[7–12] played a dominant role.

In the present study the author discusses the situation where coherent excitation of an impurity atom in a photonic crystal with a wide photonic band gap leads to the appearance of new relaxation channels. It is shown that, as a result of coherent excitation of a transition at whose frequency no photons occur, these new relaxation channels may under certain conditions form a population inversion on a normal transition whose frequency does not fall within the photonic band gap and other characteristics of the doped crystal. In addition, these same relaxation mechanisms may also be responsible for filling an impurity atomic level situated in the photonic band gap as a result of the normal coherent excitation of a neighboring atomic transition lying outside the spectral characteristics of the photonic crystal.

The essential features of the relaxation mechanism forming the basis of the new relaxation channels under coherent excitation of an atomic transition in a photonic crystal are illustrated in Fig. 1. In addition to coherent excitation of an optically allowed transition $E_a \rightarrow E_b$ by an electromagnetic wave of frequency Ω_{coh} (Fig. 1a), spontaneous two-photon processes take place in the field of this wave: the absorption of a coherent field photon $\hbar\Omega_{\mathrm{coh}}$ is accompanied by the

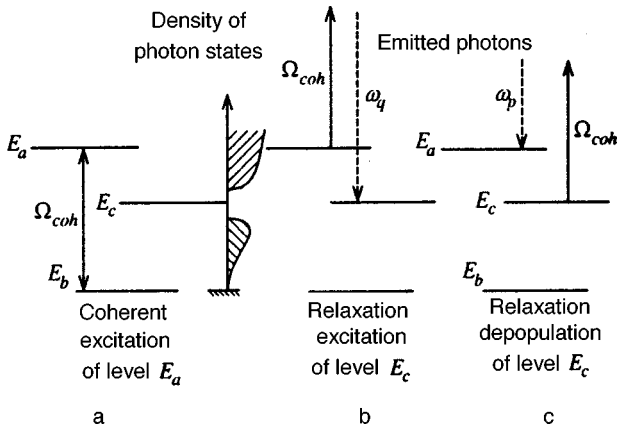

FIG. 1. Two-quantum mechanism for level filling in the photonic band gap.

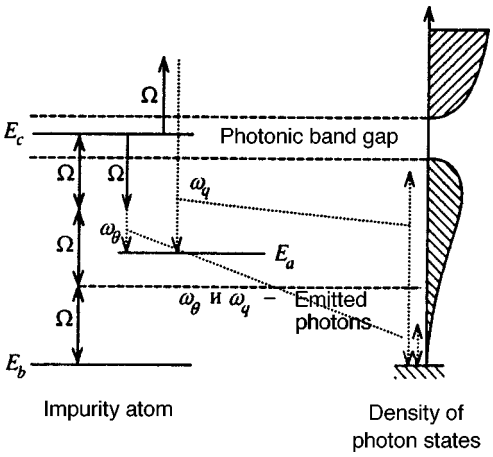

FIG. 2. Relaxation processes as a result of three-photon excitation of an impurity level located in the photonic band gap.

emission of a photon $\hbar\omega_q$, $\omega_q-\Omega_{\rm coh}\approx(E_a-E_c)/\hbar$ (Fig. 1b) such that a two-quantum transition $E_a\to E_c$ takes place.[16] The rate of this transition is proportional to the intensity of the coherent wave although under normal conditions it is still lower than the rate of a spontaneous single-photon optically allowed transition. Hence, this two-quantum relaxation does not generally appear in normal media. The situation is different in photonic crystals. Despite its smallness this two-quantum relaxation process leads to filling of the E_c level, from which it is difficult to depart by the usual relaxation channels since it is situated in the photonic band gap. A similar process involving the absorption of a photon $\hbar\Omega_{\rm coh}$ and the simultaneous emission of a photon $\hbar\omega_p$, $\Omega_{\rm coh}-\omega_p\approx(E_a-E_c)/\hbar$ leads to depopulation of this level (Fig. 1c). The competition between these processes will determine the real population of the E_c level. If the impurity atom is sufficiently isolated, its level E_c is some distance from the edges of the photonic band gap, and the intensity of the coherent wave is sufficiently high, other relaxation channels[7–13] may be neglected. In this case, however, the gradual depopulation of the E_c level accompanied by the emission of a photon $\hbar\omega_p$ is also insignificant because the frequencies of these photons fall within the photonic band gap since

$$\omega_p\approx\Omega_{\rm coh}-(E_a-E_c)/\hbar\approx(E_a-E_b)/\hbar-(E_a-E_c)/\hbar=(E_c-E_b)/\hbar.$$

Nevertheless, this process must be taken into account when going beyond the representation of the energy structure of an impurity atom using the three levels E_a, E_b, and E_c. Note that when selecting the notation for a particular energy level, we adhere to the convention whereby the pair of levels E_a and E_b or E_c and E_b forms optically allowed transitions, whereas the levels E_a and E_c belong to the optically forbidden (two-quantum) transition.

The filling of the E_c level situated in the photonic band gap and the establishment of population inversion for the $E_c\to E_b$ transition is not the only manifestation of two-quantum relaxation in photonic crystals. Another process accompanying the coherent excitation of the $E_c\to E_b$ transition is also extremely interesting in our opinion. In the present paper we consider the coherent excitation of the optically allowed $E_c\to E_b$ transition as a result of three-photon resonant interaction with a coherent wave of frequency $\Omega\approx(E_c-E_b)/3\hbar$ which does not fall within the photonic band gap. Then, the two-quantum mechanism described above leads to filling of the E_a excited level of the neighboring optically allowed $E_a\to E_b$ transition (Fig. 2). In this case, no constraints are imposed on the subsequent relaxation from the E_a level. Under certain conditions, population inversion may occur as a result of the $E_a\to E_b$ transition, but its appearance and magnitude are strictly related to the relaxation parameters of the problem. Quite clearly, other variants of the coherent excitation of an impurity atom and two-quantum relaxation processes leading to population inversion for any transition are also possible.

The effects discussed in the present paper are of a fairly general nature for problems involving the coherent excitation of impurity atoms in photonic crystals since for any other configurations of neighboring transitions, it is always possible to select those frequencies of the exciting coherent field for which the corresponding[16] two-quantum relaxation processes take place efficiently provided that some of the usual relaxation channels are suppressed by the spectral characteristics of photonic crystals.

The present paper is constructed as follows. The formulation of the problem and the effective Hamiltonian are given in Sec. 2. Kinetic equations for a *V*-configuration three-level impurity atom are derived in Sec. 3 for cases when one of the excited levels lies in the photonic band gap and one of the optically allowed transitions interacts resonantly with the coherent wave. The dynamics of the coherent excitation of an impurity atom are studied in Secs. 4 and 5 and population inversion conditions are determined for a neighboring optically allowed transition for various configuration of impurity atomic levels and various positions of the photonic band gap. A bleaching effect induced by two-quantum relaxation and coherent pumping of an optically allowed transition is described in Sec. 5. In the Conclusions we briefly discuss the case of a Λ-configuration of impurity atomic levels where the coherent wave interacts resonantly with an optically allowed transition which is not related to the ground atomic level and a neighboring optically allowed transition is ex-

posed to the action of the spectral characteristics of the photonic crystal. It is shown that in this case, the two-quantum mechanism can be considered as an additional relaxation channel for the level in the photonic band gap.

2. FORMULATION OF THE PROBLEM AND EFFECTIVE HAMILTONIAN

We shall first consider the situation where electromagnetic radiation having the electric field intensity $\mathbf{E}$

$$\mathbf{E}=\mathscr{E}e^{-i\Phi}+\mathscr{E}^{*}e^{i\Phi},\quad \Phi=\Omega t+\varphi_0, \tag{1}$$

is in three-photon resonance ($\Omega\approx(E_c-E_b)/3\hbar$) with an optically allowed transition $E_c\to E_b$ of an impurity atom whose excited level E_c lies in the photonic band gap and E_b is the ground level. For the other atomic levels we shall assume that the field (1) only forms two-quantum resonances with a certain level E_a belonging to the neighboring optically allowed $E_a\to E_b$ transition: the two-photon resonance ($\Omega+\omega_\theta\approx(E_c-E_a)/\hbar$) involving quanta ω_θ of the photon thermostat H_θ and the Raman resonance ($\omega_q-\Omega\approx(E_c-E_a)/\hbar$) involving quanta of the other photon thermostat H_q (Fig. 2).

We write the initial Hamiltonian of the system in the form

$$H=H_0+V_{\mathrm{coh}}+V_v+V_{ca}^{\theta}+V_{ca}^{q}+V_{ab}+H_\Omega+H_l+H_\theta+H_q, \tag{2}$$

where H_0 is the Hamiltonian of an isolated impurity atom, V_{coh}, V_v, V_{ab}, V_{ca}^{θ}, and V_{ca}^{q} are the operators of the interaction between an impurity atom and the coherent field (1), quanta of the photon thermostat with the central frequency Ω, quanta of photon thermostats in resonance with the optically allowed transition $E_a\to E_b$, and quanta of the photon thermostat in two-photon resonance [jointly with the wave (1)] with the optically forbidden transition $E_c\to E_a$, respectively:

$$V_{\mathrm{coh}}=-\mathscr{E}e^{-i\Phi}\sum_{\alpha,\alpha'}d_{\alpha\alpha'}a_\alpha^{+}a_{\alpha'}+\mathrm{H.c.},$$

$$V_v=-i\sum_{\alpha,\alpha'}\int d\omega_\Omega K(\omega_\Omega)d_{\alpha\alpha'}a_\alpha^{+}a_{\alpha'}b_{\omega_\Omega}+\mathrm{H.c.},$$

$$V_{ab}=-i\sum_{\alpha,\alpha'=a,b}\int d\omega_l K(\omega_l)d_{\alpha\alpha'}a_\alpha^{+}a_{\alpha'}b_{\omega_l}+\mathrm{H.c.},$$

$$V_{ca}^{\theta}=-i\sum_{\alpha,\alpha'}\int d\omega_\theta K(\omega_\theta)d_{\alpha\alpha'}a_\alpha^{+}a_{\alpha'}b_{\omega_\theta}+\mathrm{H.c.},$$

$$V_{ca}^{q}=-i\sum_{\alpha,\alpha'}\int d\omega_q K(\omega_q)d_{\alpha\alpha'}a_\alpha^{+}a_{\alpha'}b_{\omega_q}+\mathrm{H.c.}$$

Here H_Ω, H_l, H_θ, and H_q are the Hamiltonians of the corresponding photon thermostats, which, in accordance with Ref. 17, can be considered to be independent and which we shall also assume to be Markovian, assuming that their central frequencies do not fall within the photonic band gap and are located at some distance from the edges of this gap [see also the conditions (7)]. We shall use the index ''th'' to indicate that some quantity belongs to the thermostat characteristics:

$$H_{\mathrm{th}}=\int d\omega_{\mathrm{th}}\hbar\omega_{\mathrm{th}}b_{\omega_{\mathrm{th}}}^{+}b_{\omega_{\mathrm{th}}},\quad \mathrm{th}=\Omega,l,\theta,q.$$

The other quantities are the creation a_α^{+} and annihilation a_α operators of an atom in the quantum state α, the creation $b_{\omega_{\mathrm{th}}}^{+}$ and annihilation operators $b_{\omega_{\mathrm{th}}}$ of thermostat photons, the coupling constants $K(\omega)$, and the matrix elements $d_{\alpha\alpha'}$ of the atomic operator of the dipole moment d. We shall neglect polarization and spatial effects.

We stress that the role of the photonic band gap can be seen in that the initial Hamiltonian (2) contains no interaction operator with a photon thermostat whose central frequency matches the $E_c\to E_b$ transition frequency. The assumption that the E_c level is some distance from the edges of this gap and that the density of impurity atoms is low (in our case, we only consider a single impurity atom) means that we can neglect relaxation mechanisms based on resonant dipole–dipole interaction between like atoms, interaction with a non-Markovian thermostat, and so on.[7–12] Quadrupole[7] and diatomic[13] mechanisms can be neglected because of their smallness compared with those included in Eq. (2).

By means of a unitary transformation,[16,18] we obtain the effective Hamiltonian of the problem from Eq. (2). For this purpose we transform the density matrix of the system using the transformation $\tilde{\rho}=e^{-iS}\rho e^{iS}$. The new density matrix will satisfy the equation

$$i\hbar\frac{\partial}{\partial t}\tilde{\rho}=[\tilde{H},\tilde{\rho}], \tag{3}$$

where the transformed Hamiltonian

$$\tilde{H}=e^{-iS}(H_0+V_{\mathrm{coh}}+V_v+V_{ca}^{\theta}+V_{ca}^{q}+V_{ab}+H_\Omega+H_l+H_\theta+H_q)e^{iS}-i\hbar e^{-iS}\frac{\partial}{\partial t}e^{iS}, \tag{4}$$

and the Hermitian operator S, which determines the transformation, are expanded as a series in terms of the coherent and vacuum fields

$$S=S^{(10)}+S^{(01)}+S^{(11)}+\ldots,$$

$$\tilde{H}=\tilde{H}^{(00)}+\tilde{H}^{(10)}+\tilde{H}^{(01)}+\tilde{H}^{(11)}+\ldots,$$

where the left-hand index in each pair (m,n) refers to the coherent field and the right-hand index refers to the vacuum field. For $n\geqslant 1$ we divide $S^{(m,n)}$ and $\hat{H}^{(m,n)}$ into four terms

$$S^{(m,n)}=S_\Omega^{(m,n)}+S_l^{(m,n)}+S_\theta^{(m,n)}+S_q^{(m,n)},$$

$$\tilde{H}^{(m,n)}=\tilde{H}_\Omega^{(m,n)}+\tilde{H}_l^{(m,n)}+\tilde{H}_\theta^{(m,n)}+\tilde{H}_q^{(m,n)},$$

corresponding to the contributions of each of the photon thermostats taken into account. Since

$$\tilde{H}=H-i[S,H]-\frac{1}{2}[S,[S,H]]$$

$$+\frac{i}{6}[S,[S,[S,H]]]+\ldots-i\hbar e^{-iS}\frac{\partial}{\partial t}e^{iS},$$

we have

$$\tilde{H}^{(0,0)}=H_0+H_l+H_\theta,$$

$$\tilde{H}^{(1,0)}=V_{\rm coh}-i[S^{(1,0)},\tilde{H}^{(0,0)}]+\hbar\frac{\partial}{\partial t}S^{(1,0)},$$

$$\tilde{H}_{\rm th}^{(0,1)}=V_{\rm th}-i[S_{\rm th}^{(0,1)},\tilde{H}^{(0,0)}]+\hbar\frac{\partial}{\partial t}S_{\rm th}^{(0,1)},$$

$$\tilde{H}_{\rm th}^{(1,1)}=-\frac{i}{2}[S_{\rm th}^{(0,1)},V_{\rm coh}]-\frac{i}{2}[S_{\rm th}^{(1,0)},V_{\rm th}]$$

$$-\frac{i}{2}[S_{\rm th}^{(0,1)},\tilde{H}^{(1,0)}]-\frac{i}{2}[S^{(1,0)},\tilde{H}_{\rm th}^{(0,1)}]$$

$$-i[S_{\rm th}^{(1,1)},\tilde{H}^{(0,0)}]+\hbar\frac{\partial}{\partial t}S_{\rm th}^{(1,1)},$$

$$\tilde{H}^{(2,0)}=-\frac{i}{2}[S^{(1,0)},V_{\rm coh}]-\frac{i}{2}[S^{(1,0)},\tilde{H}^{(1,0)}]$$

$$-i[S^{(2,0)},H_0]+\hbar\frac{\partial}{\partial t}S^{(2,0)},$$

$$\tilde{H}^{(3,0)}=-\frac{i}{2}[S^{(2,0)},V_{\rm coh}]-\frac{i}{2}[S^{(1,0)},\tilde{H}^{(2,0)}]$$

$$-\frac{i}{2}[S^{(2,0)},\tilde{H}^{(1,0)}]-\frac{1}{12}[S^{(1,0)},[S^{(1,0)},V_{\rm coh}]]$$

$$+\frac{1}{12}[S^{(1,0)},[S^{(1,0)},\tilde{H}^{(1,0)}]]-i[S^{(3,0)},H_0]$$

$$+\hbar\frac{\partial}{\partial t}S^{(3,0)}.$$

Here the subscript ''th'' indicates one of the photon thermostats.

We impose the constraint that the following conditions corresponding to the resonance processes in the problem and the approximation of slowly varying amplitudes are satisfied:

$$\tilde{H}^{(1,0)}=\tilde{H}_\Omega^{(0,1)}=\tilde{H}_\theta^{(0,1)}=\tilde{H}_q^{(0,1)}=0,$$

$$\tilde{H}_l^{(0,1)}=-i\int d\omega_l K(\omega_l)d_{ab}a_a^+a_b b_{\omega_l}+\text{H.c.},$$

$$\tilde{H}^{(2,0)}=\sum E_\alpha^{(2)}a_\alpha^+a_\alpha,$$

$$\tilde{H}^{(3,0)}=u_{cb}a_c^+a_b e^{-3i\Phi}+\text{H.c.}+\sum E_\alpha^{(3)}a_\alpha^+a_\alpha,$$

where $E_\alpha^{(2)}$, $E_\alpha^{(3)}$, and u_{cb} contain no oscillating exponential functions $e^{is\Phi}$, $s=\pm1,\pm2,\ldots$. The validity of these conditions is confirmed by the absence of any resonant denominators in $S^{(m,n)}$:

$$S^{(10)}=-i\sum\frac{d_{\alpha\alpha'}\mathscr{E}e^{-i\Phi}a_\alpha^+a_{\alpha'}}{\hbar(\omega_{\alpha\alpha'}-\Omega)}+\text{H.c.},$$

$$S_l^{(01)}=-\int d\omega_l\frac{K(\omega_l)d_{ba}a_b^+a_a b_{\omega_l}}{\hbar(\omega_{ab}+\omega)}+\text{H.c.},$$

$$S_{\rm th}^{(01)}=\sum_{\alpha\alpha'}\int d\omega_{\rm th}\frac{K(\omega_{\rm th})d_{\alpha\alpha'}a_\alpha^+a_{\alpha'}b_{\omega_{\rm th}}\theta(E_{\alpha'}-E_\alpha)}{\hbar(\omega_{\alpha\alpha'}-\omega_{\rm th})}+\text{H.c.},\tag{5}$$

$$\text{th}=\Omega,\theta,q,\quad S^{(2,0)}=\frac{i}{4\hbar\Omega}\sum_\alpha\Pi_\alpha(\Omega)$$

$$\times(\mathscr{E}^2e^{-2i\Phi}-\mathscr{E}^{*2}e^{2i\Phi})a_\alpha^+a_\alpha,$$

where $\omega_{\alpha\alpha'}=(E_\alpha-E_{\alpha'})/\hbar$ and the usual notation[18] is used for the parameter determining the Stark level shift:

$$\Pi_\alpha(\Omega)=\sum_{\alpha'}\frac{|d_{\alpha\alpha'}|^2}{\hbar}\left(\frac{1}{\omega_{\alpha\alpha'}+\Omega}+\frac{1}{\omega_{\alpha\alpha'}-\Omega}\right).$$

After various simple calculations, we obtain the effective Hamiltonian of the problem in the following form:

$$H^{\rm eff}=H_0+H^{\rm St}+V^{(3)}+V_\Omega+V_l+V_\theta$$

$$+V_q+H_\Omega+H_l+H_\theta+H_q,\tag{6}$$

where $H^{\rm St}$ describes the dynamic Stark effect, $V^{(3)}$ determines the three-photon coherent excitation of the $E_c\to E_b$ transition, V_l determines the $E_a\leftrightarrow E_b$ relaxation transitions, V_θ and V_q determine the $E_c\leftrightarrow E_a$ relaxation transitions in an impurity atom as a result of interaction with the thermostat, and V_Ω is responsible for the additional broadening of the spectral lines when allowance is made for quantum fluctuations in the Stark effect:

$$H^{\rm St}=\sum|\mathscr{E}|^2\Pi_\alpha(\omega)a_\alpha^+a_\alpha,$$

$$V^{(3)}=u_{cb}a_c^+a_b\mathscr{E}^3e^{-3i\Phi}+\text{H.c.},$$

$$u_{cb}=\frac{d_{cb}}{8\hbar\Omega}(\Pi_b(\Omega)-\Pi_c(\Omega))+\frac{1}{12\hbar^2}$$

$$\times\sum_{\alpha\beta}d_{c\alpha}d_{\alpha\beta}d_{\beta b}\left(\frac{2}{(\omega_{c\alpha}-\Omega)(\omega_{\beta b}-\Omega)}\right.$$

$$\left.-\frac{1}{(\omega_{c\alpha}-\Omega)(\omega_{\alpha\beta}-\Omega)}-\frac{1}{(\omega_{\alpha\beta}-\Omega)(\omega_{\beta b}-\Omega)}\right),$$

$$V_{\rm th}=-i\hbar\int d\omega_{\rm th}K(\omega_{\rm th})[f_{\rm th}(t)R_{\rm th}^+b_{\omega_{\rm th}}-f_{\rm th}^*(t)R_{\rm th}b_{\omega_{\rm th}}^+],$$

$$R_l=a_b^+a_a,\quad f_l(t)=d_{ba}/\hbar,$$

$$R_\Omega=\sum a_\alpha^+a_\alpha\Pi_\alpha(\Omega),\quad f_\Omega(t)=-\mathscr{E}^*e^{i\Phi}/\hbar,$$

$$R_\theta=a_a^+a_c,\quad f_\theta(t)=\mathscr{E}e^{-i\Phi}\Pi_{ac}(\Omega)/\hbar,$$

$$R_q=a_a^+a_c,\quad f_q(t)=\mathscr{E}^*e^{i\Phi}\Pi_{ca}(-\Omega)/\hbar.$$

The parameter $\Pi_{ac}(\Omega)$ introduced determines the two-photon interaction operator:[18]

$$\Pi_{ac}(\Omega)=\sum_{\beta}\frac{d_{a\beta}d_{\beta c}}{\hbar}\left(\frac{1}{\omega_{\beta a}+\Omega}+\frac{1}{\omega_{\beta c}-\Omega}\right).$$

Note that the subscripts α, α', and β number all the energy levels of the impurity atom.

The effective Hamiltonian (6) can be used to isolate the relaxation operator by standard methods and obtain definitive kinetic equations describing the dynamics of an impurity atom in a photonic crystal during its coherent excitation process.

3. KINETIC EQUATION FOR AN IMPURITY ATOM

We shall list the main initial assumptions for the photon thermostats H_Ω, H_l, and H_θ introduced in the analysis and the coupling constants:

$$\langle b_{\rm th}^{+}(t)b_{\rm th}(t')\rangle=N_{\rm th}\delta(t-t'),$$

$$\langle b_{\rm th}(t)b_{\rm th}^{+}(t')\rangle=(1+N_{\rm th})\delta(t-t'),$$

$$\langle b_{\rm th}^{+}(t)b_{\rm th}^{+}(t')\rangle=\langle b_{\rm th}(t)b_{\rm th}(t')\rangle=0,$$

$$b_{\rm th}(t)=\frac{1}{\sqrt{2\pi}}\int d\omega_{\rm th}\exp[-i\omega_{\rm th}(t-t_0)]b_{\omega_{\rm th}},\qquad(7)$$

$$K(\omega_{\rm th})=\sqrt{\kappa_{\rm th}/2\pi},$$

where N is the photon density of the corresponding thermostat and t_0 is a certain initial time to which the photon operators $b_{\omega_{\rm th}}$ and $b_{\omega_{\rm th}}^{+}$ refer. Thus, the photon thermostats of the problem are considered to be wide-band and the coupling constants satisfy the Markov approximation, i.e., they do not depend on frequency and are determined by the central frequency of the corresponding photon thermostat. We shall subsequently assume that the operators of the photon thermostats evolve as the operators of noninteracting systems and determine the Wiener processes

$$B_{\rm th}(t,t_0)=\int_{t_0}^{t}dt'\,b_{\rm th}(t'),$$

$$[B_{\rm th}(t,t_0),B_{\rm th'}^{+}(t,t_0)]=(t-t_0)\,\delta_{\rm th\,th'},$$

whose growth rates satisfy ordinary Ito algebra:

$$dB_{\rm th}^{+}(t)dB_{\rm th'}(t)=N_{\rm th}dt\,\delta_{\rm th\,th'},$$

$$dB_{\rm th}(t)dB_{\rm th'}^{+}(t)=(1+N_{\rm th})dt\,\delta_{\rm th\,th'},$$

$$dB_{\rm th}(t)dB_{\rm th'}(t)=dB_{\rm th}^{+}(t)dB_{\rm th'}^{+}(t)=dB_{\rm th}(t)dt$$
$$=dtdB_{\rm th}(t)=dB_{\rm th}^{+}(t)dt=dtdB_{\rm th}^{+}(t)$$
$$=dtdt=0,$$

where the subscripts ''th'' and ''th'' denote the photon thermostats, i.e., have values of Ω, l, q, and θ.

Subject to these assumptions we can use the usual methods[16,19,20] to write the kinetic equation for the density matrix of an impurity atom in the form

$$\frac{d}{dt}\rho+\hat{\mathscr{T}}\rho=\frac{i}{\hbar}[\rho,H_0+H^{\rm St}+V^{(3)}]\qquad(8)$$

with the relaxation operator in the standard Lindblad form

$$\hat{\mathscr{T}}\rho=-\sum_{{\rm th}=\Omega,l,\theta}\frac{\kappa_{\rm th}}{2}|f_{\rm th}(t)|^2(1+N_{\rm th})(2R_{\rm th}\rho R_{\rm th}^{+}-\rho R_{\rm th}^{+}R_{\rm th}$$
$$-R_{\rm th}^{+}R_{\rm th}\rho)-\sum_{{\rm th}=\Omega,l,\theta}\frac{\kappa_{\rm th}}{2}|f_{\rm th}(t)|^2N_{\rm th}(2R_{\rm th}^{+}\rho R_{\rm th}$$
$$-\rho R_{\rm th}R_{\rm th}^{+}-R_{\rm th}R_{\rm th}^{+}\rho).\qquad(9)$$

Here and subsequently we neglect the tilde over the transformed density matrix. It should be stressed that the unitary transformation method gives a far clearer and more complete physical picture of the various types of relaxation processes, including two-quantum ones,[16] compared with derivations of the relaxation operator from the Bloch equations[21,22] or using the quantum jump technique.[23]

The off-diagonal matrix elements of the density matrix of an impurity atom can be conveniently expressed in the form

$$\rho_{cb}=r_{cb}\exp(-3i\Phi),\quad \rho_{ca}=r_{ca}\exp(-i\omega_{ca}t),$$

$$\rho_{ab}=r_{ab}\exp(-i\omega_{ab}t),$$

and then the following system of equations is obtained from Eqs. (8) and (9):

$$\frac{d}{dt}r_{cb}-i(\Delta-\Delta_{cb}|\mathscr{E}|^2)r_{cb}$$
$$=\frac{i}{\hbar}(\rho_{cc}-\rho_{bb})u_{cb}\mathscr{E}^3-\gamma_{cb}|\mathscr{E}|^2r_{cb},$$

$$\frac{d}{dt}\rho_{bb}=\gamma_a^{(0)}\rho_{aa}+\frac{i}{\hbar}(r_{cb}^{*}u_{cb}\mathscr{E}^3-r_{cb}u_{cb}^{*}\mathscr{E}^{*3}),$$

$$\frac{d}{dt}\rho_{cc}=-\gamma_c|\mathscr{E}|^2\rho_{cc}-\frac{i}{\hbar}(r_{cb}^{*}u_{cb}\mathscr{E}^3-r_{cb}u_{cb}^{*}\mathscr{E}^{*3}),$$

$$\frac{d}{dt}\rho_{aa}=-\gamma_a^{(0)}\rho_{aa}+\gamma_c|\mathscr{E}|^2\rho_{cc},$$

$$\frac{d}{dt}r_{ca}+i\Delta_{ca}|\mathscr{E}|^2r_{ca}=-(\gamma_{ca}^{(0)}+\gamma_{ca}|\mathscr{E}|^2)r_{ca},$$

$$\frac{d}{dt}r_{ab}+i\Delta_{ab}|\mathscr{E}|^2r_{ab}=-(\gamma_{ab}^{(0)}+\gamma_{ab}|\mathscr{E}|^2)r_{ab}.\qquad(10)$$

Here we use the notation

$$\Delta=3\Omega-\omega_{cb},\quad \Delta_{cb}=\Pi_c(\Omega)-\Pi_b(\Omega),$$

$$\Delta_{ca}=\Pi_c(\Omega)-\Pi_a(\Omega),\quad \Delta_{ab}=\Pi_a(\Omega)-\Pi_b(\Omega),$$

$$\gamma_{cb}=\kappa_\Omega(\Pi_c(\Omega)-\Pi_b(\Omega))^2/2\hbar^2+\gamma_c/2,$$

$$\gamma_{ab}^{(0)}=\gamma_{ca}^{(0)}=\gamma_a^{(0)}/2=\kappa_l|d_{ab}|^2/2\hbar^2,$$

$$\gamma_{ca}=\kappa_\Omega(\Pi_c(\Omega)-\Pi_a(\Omega))^2/2\hbar^2+\gamma_c/2,$$

$$\gamma_{ab}=\kappa_\Omega(\Pi_a(\Omega)-\Pi_b(\Omega))^2/2\hbar^2,$$

$$\gamma_c=\kappa_q|\Pi_{ca}(-\Omega)|^2/\hbar^2+\kappa_\theta|\Pi_{ca}(\Omega)|^2/\hbar^2.\qquad(11)$$

We recall that for a three-dimensional spontaneous emission geometry

$$\gamma_a^{(0)}=\frac{4|d_{ab}|^2\omega_{ab}^3}{3\hbar c^3(2J_a+1)},$$

where d_{ab} is the reduced dipole moment of the $E_a\to E_b$ transition and J_a is the angular momentum of the E_a level.

When writing the equations (10), we assumed that the photon thermostats are vacuum ones containing no photons with $N_{\rm th}=0$.

We now write the equations for the density matrix for the case where the wave (1) is in one-photon resonance with the $E_a\to E_b$ transition, $\Omega\approx(E_a-E_b)/\hbar$, and the spectral characteristics of the photonic crystal impose no constraints on this transition (Fig. 1). Let us assume that the E_c energy level lies in a photonic band gap some distance from its edges so that the relaxation mechanisms[7–13] are small compared with the two-quantum mechanism being discussed. For $r_{ab}=\rho_{ab}\exp(i\Phi)$, ρ_{aa}, ρ_{bb}, and ρ_{cc} we then have the equations

$$\frac{d}{dt}r_{ab}-i(\Delta-\Delta_{ab}|\mathscr{E}|^2)r_{ab}=-\frac{i}{\hbar}(\rho_{aa}-\rho_{bb})d_{ab}\mathscr{E}-(\gamma_a^{(0)}/2+\gamma_{ab}|\mathscr{E}|^2)r_{ab},$$

$$\frac{d}{dt}\rho_{bb}=\gamma_a^{(0)}\rho_{aa}-\frac{i}{\hbar}(r_{ab}^*d_{ab}\mathscr{E}-r_{ab}d_{ab}^*\mathscr{E}^*),$$

$$\frac{d}{dt}\rho_{aa}=-(\gamma_a^{(0)}+\gamma_a|\mathscr{E}|^2)\rho_{aa}+\frac{i}{\hbar}(r_{ab}^*d_{ab}\mathscr{E}-r_{ab}d_{ab}^*\mathscr{E}^*),$$

$$\frac{d}{dt}\rho_{cc}=\gamma_a|\mathscr{E}|^2\rho_{aa}. \qquad (12)$$

Here the quantities introduced have different values [unlike Eq. (11)]:

$$\Delta=\Omega-\omega_{ab},\quad \Delta_{ab}=\Pi_a(\Omega)-\Pi_b(\Omega),$$

$$\gamma_{ab}=\kappa_\Omega(\Pi_a(\Omega)-\Pi_b(\Omega))^2/2\hbar^2+\gamma_a/2,$$

$$\gamma_a=\kappa_q|\Pi_{ac}(-\Omega)|^2/\hbar^2,\quad \gamma_a^{(0)}=\kappa_l|d_{ab}|^2/\hbar^2. \qquad (13)$$

4. FILLING DYNAMICS OF THE E_a LEVEL. POPULATION INVERSION

In this section we consider the situation shown schematically in Fig. 2. The system of equations (10) and (11) describing the three-photon excitation of the $E_c\to E_b$transition under conditions where two-quantum relaxation channels play a role, can be used to analyze various effects such as optical nutations, inductions, echo phenomena,[24] accompanying the coherent pumping of the $E_c\to E_b$ transition. At this point we merely discuss the excitation of the E_a level.

The system of equations (10) and (11) has a steady-state solution from which we write the quantities relating to the $E_a\to E_b$: transition:

$$\rho_{aa}=\frac{\gamma_c|\mathscr{E}|^2\Gamma}{\gamma_a^{(0)}\gamma_c|\mathscr{E}|^2+\gamma_c|\mathscr{E}|^2\Gamma+2\gamma_a^{(0)}\Gamma},$$

$$\rho_{bb}=\frac{\gamma_a^{(0)}\Gamma+\gamma_a^{(0)}\gamma_c|\mathscr{E}|^2}{\gamma_a^{(0)}\gamma_c|\mathscr{E}|^2+\gamma_c|\mathscr{E}|^2\Gamma+2\gamma_a^{(0)}\Gamma},$$

$$\rho_{ab}=0,\quad \Gamma=\frac{2\gamma_{cb}|u_{cb}|^2|\mathscr{E}|^8/\hbar^2}{\gamma_{cb}^2|\mathscr{E}|^4+(\Delta-\Delta_{cb}|\mathscr{E}|^2)^2}.$$

It should be noted that pumping of the E_a level as a result of coherent excitation of the $E_c\to E_b$ transition takes place without any coherence appearing for the $E_a\to E_b$ transition. This distinguishes this situation from any other schemes for excitation of the $E_a\to E_b$ transition by coherent fields.

We shall assess the possibility of a steady-state inversion for the $E_a\to E_b$ transition. It can be seen that this is mainly determined by the relationship between the parameters $\gamma_c|\mathscr{E}|^2$ and $\gamma_a^{(0)}$. In the most favorable case we have the estimate

$$\frac{\gamma_c|\mathscr{E}|^2}{\gamma_a^{(0)}}\sim\left(\frac{\omega_q}{\omega_{ab}}\right)^3\frac{\Lambda^2}{\delta^2},$$

where $\Lambda\sim\mathscr{E}d/\hbar$ is the characteristic interaction energy and δ is the smallest detuning from the nonresonant level. This analysis of resonance processes imposes the constraint $\Lambda^2\ll\delta^2$. By selecting the frequencies and dipole moments of the impurity atom this constraint can be reduced but we must assume that $\gamma_c|\mathscr{E}|^2<\gamma_a^{(0)}$. From this it follows that in the steady-state regime no inversion is established for the $E_a\to E_b$ transition although for an ensemble of impurity atoms the fraction of atoms excited to the E_a level is fairly high, of the order of

$$\frac{\gamma_c|\mathscr{E}|^2}{\gamma_c|\mathscr{E}|^2+2\gamma_a^{(0)}},$$

which may account for 10% of the total number of impurity atoms.

We consider the nonsteady-state excitation of the E_a level for the simplest case where the field (1) is switched on at time $t=0$ assuming that $\mathscr{E}(t)=0$ for $t<0$ and $\mathscr{E}(t)=\mathscr{E}=\text{const}$ for $t\geqslant 0$. As the initial conditions we take those corresponding to the equilibrium state of the impurity atom:

$$\rho_{aa}|_{t=0}=\rho_{cc}|_{t=0}=\rho_{cb}|_{t=0}=\rho_{ca}|_{t=0}=\rho_{ab}|_{t=0}=0,$$

$$\rho_{bb}|_{t=0}=1.$$

It follows from Eqs. (10) and (11) that the population of the E_a level behaves as shown in Fig. 3.

The fluctuations of the E_a level population reflect the Rabi oscillations of the pumped $E_c\to E_b$ transition. It can be seen that for typical values of the parameters the level population reaches a steady-state regime with an E_a level population of the order of 10%.

Unlike the steady-state regime, the overpopulation $\rho_{aa}-\rho_{bb}$ of the levels during a certain time interval may have a population inversion, as shown in Fig. 4.

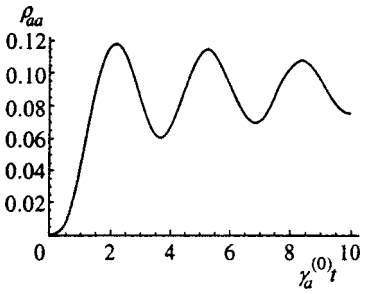

FIG. 3. Filling of the E_a level in the nonsteady-state regime assuming that $\gamma_c|\mathscr{E}|^2=2\gamma_{cb}|\mathscr{E}|^2=0.2\gamma_a^{(0)}$, $u_{cb}|\mathscr{E}|^3/\hbar=\gamma_a^{(0)}$, $\Delta=\Delta_{cb}|\mathscr{E}|^2$.

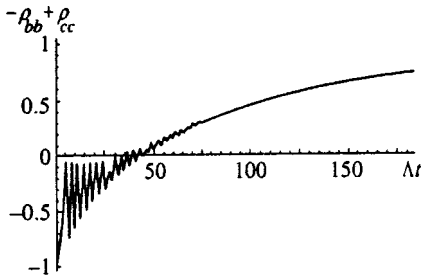

FIG. 5. Dynamics of overpopulation of the $E_c\rightarrow E_b$ transition, assuming $\Lambda=|\mathscr{E}d_{ab}/\hbar|==10\gamma_a^{(0)}=50\gamma_a=100\gamma_{ab}$, $\Delta=\Delta_{ab}|\mathscr{E}|^2$.

It can be seen that inversion of the $E_a\rightarrow E_b$ transition occurs at the first oscillation maximum. As the amplitude of the electric field intensity of the wave increases, the number of ''spikes,'' during which inversion of the $E_a\rightarrow E_b$ transition occurs, increases. The effect becomes clearer if the ratio $\gamma_c|\mathscr{E}|^2/\gamma_a^{(0)}$ increases for some reason. We stress that in the absence of two-quantum relaxation, the inequality $\rho_{aa}-\rho_{bb}\leqslant 0$. is always satisfied. Moreover, the population of the E_a level is always zero (for these initial conditions).

Thus, during short time intervals under conditions of nonsteady-state coherent pumping of the $E_c\rightarrow E_b$ transition at the frequency of the neighboring $E_a\rightarrow E_b$ transition a medium consisting of impurity atoms in a photonic crystal becomes active (amplifying). No coherence is observed for this transition.

5. POPULATION DYNAMICS OF THE E_c LEVEL. BLEACHING EFFECT

We shall consider the situation when an electromagnetic wave interacts coherently with the $E_a\rightarrow E_b$ transition: $\Omega\approx(E_a-E_b)/\hbar$, and the E_c level ascribed to a neighboring optically allowed transition is located in the photonic band gap. In the absence of two-quantum relaxation the system of equations (12) and (13) yields the ordinary Bloch equations for one-photon resonance. However, the presence of the E_c level whose normal gradual depopulation is impeded because of the spectral characteristics of a photonic crystal makes it necessary to allow even for weak two-quantum relaxation processes since this level acts as a particular trap for these processes, i.e., on reaching this level an impurity atom remains in this state for a fairly long time. Equations (12) and (13) can be used to analyze all the characteristic features of the dynamics of ordinary optically allowed transitions associated with the presence of these ''traps.'' At this point we shall discuss some of these features.

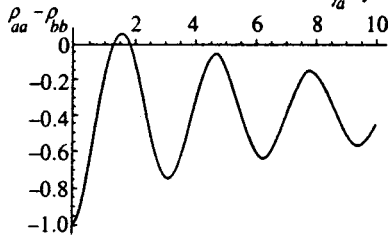

FIG. 4. Behavior of the population inversion for the $E_a\rightarrow E_b$ transition in the nonsteady-state regime. The values of the parameters are the same as in Fig. 3.

The steady-state solution of the system (12) and (13) has the extremely simple form

$$\rho_{cc}=1,\quad \rho_{bb}=0,\quad \rho_{aa}=0,\quad \rho_{ab}=0,\quad \rho_{bc}=0,$$

which confirms the term ''trap'' for the level of an impurity atom located in the photonic band gap. It can be seen that the $E_c\rightarrow E_b$ transition becomes completely inverted. However, the emission as a result of this transition is completely suppressed by the spectral characteristics of the photonic crystal. Figure 5 illustrates the attainment of this steady state when the field (1) is switched on at time $t=0$: $\mathscr{E}(t)=0$ for $t<0$ and $\mathscr{E}(t)=\mathscr{E}=\text{const}$ for $t\geqslant 0$ The oscillations reflect the Rabi oscillations of the $E_a\rightarrow E_b$ transition. The time taken to reach the steady-state regime in a fairly high-intensity field (1) is of the order of

$$\frac{\hbar^2}{\kappa_q|\mathscr{E}\Pi_{ac}(-\Omega)|^2}.$$

Thus, if an impurity atom in a photonic crystal has at least one level coupled to the ground level by an optically allowed transition and falling within the photonic band gap, coherent pumping of any optically allowed transition containing the ground energy level will be accompanied by ''bleaching'' of this transition, i.e., depletion of the energy levels of this transition and filling of the level situated in the photonic band gap. This directly influences the propagation of the coherent wave (1) in the medium, i.e., after a certain time required to transfer the excitation to the E_c level, the impurity atoms become ''transparent'' for the wave (1) since they cease to interact with it. However, after a certain time interval, relaxation mechanisms come into play[7–13] which we neglected compared with two-quantum relaxation channel but which must be taken into account over a fairly long time interval. Such an analysis should be made separately. For the simplest quadrupole[7] or diatomic[13] mechanisms for relaxation of the $E_c\rightarrow E_b$ transition these can be taken into account by adding to the right-hand sides of the equations (12) terms proportional to the rate constant $\gamma_c^{(2)}$ of these processes:

$$\frac{d}{dt}\rho_{bb}\propto\gamma_c^{(2)}\rho_{cc},\quad \frac{d}{dt}\rho_{cc}\propto-\gamma_c^{(2)}\rho_{cc}. \tag{14}$$

The steady-state solution of the system (12)–(14) is then altered slightly:

$$\rho_{cc}\simeq 1-\varepsilon(2+\gamma_a^{(0)}/\Gamma),\quad \rho_{bb}\simeq\varepsilon(1+\gamma_a^{(0)}/\Gamma),\quad \rho_{aa}\simeq\varepsilon,$$

where

$$\varepsilon=\frac{\gamma_c^{(2)}}{\gamma_a|\mathscr{E}|^2}\ll 1,$$

$$\Gamma=2\frac{|d_{ab}\mathscr{E}/\hbar|^2(\gamma_a^{(0)}/2+\gamma_{ab}|\mathscr{E}|^2)}{(\gamma_a^{(0)}/2+\gamma_{ab}|\mathscr{E}|^2)^2+(\Delta-\Delta_{ab}|\mathscr{E}|^2)^2}.$$

Hence, even when the relaxation of the $E_c\to E_b$ transition is taken into account, the neighboring transition has a negligibly small overpopulation:

$$\rho_{aa}-\rho_{bb}=-\varepsilon\gamma_a^{(0)}/\Gamma\sim\varepsilon\gamma_a^{(0)2}/4\Lambda^2,$$

$$\Lambda=|d_{ab}\mathscr{E}/\hbar|^2\gg\gamma_a^{(0)},$$

and the bleaching effect predominates. We stress that in the proposed system for bleaching of the medium the electromagnetic wave interacts resonantly with impurity atoms in the photonic crystal and propagates outside the photonic band gap, unlike in Ref. 25 where a different bleaching mechanism was discussed when Λ-configuration atoms of the photonic crystal interact with a high-power electromagnetic field whose frequency fell within the band gap. The differences between the bleaching mechanism described and the familiar bleaching effect caused by saturation of a resonant transition[26] are also obvious and do not require separate discussion.

6. CONCLUSIONS

In the present paper we have demonstrated how the suppression of the ordinary relaxation channels in impurity atoms in photonic crystals brings new relaxation channels to the forefront, based on two-quantum relaxation mechanisms accompanying the action of coherent fields on impurity atoms. In these examples two-quantum relaxation resulted in the filling of levels inside and outside the photonic band gap. We confined our analysis to the *V* configuration of impurity-atom energy levels involved in the resonant pumping and two-quantum relaxation processes, although all the energy levels of the impurity atom were formally taken into account in the parameters of the resonant processes. However, the dominant role of two-quantum relaxation should also appear in other situations. Depending on the specific conditions, two-quantum relaxation channels can either lead to filling or deactivation of a level in the photonic band gap.

An example may be the situation considered by Quang *et al.*[15] where the frequency of the $E_b\to E_a$ transition assigned to the Λ configuration of impurity levels $E_a<E_c<E_b$, fell within the zone of action of the spectral characteristics of the photonic crystal and the $E_b\to E_c$ transition was pumped by a coherent wave. These authors[15] suggested that the frequency of the $E_b\to E_a$ transition should lie near the edge of the photonic band gap so that the rate of relaxation of the $E_b\to E_a$ transition is not too suppressed. The action of the coherent field in Ref. 15 is considered as coherent control of the relaxation mechanism for the $E_b\to E_a$ transition. Nevertheless, in this situation we need to allow for two-quantum relaxation channels which, unlike those considered in the present article, lead to additional deactivation of the E_b level. Depending on the parameters in the case considered in Ref. 15, the role of two-quantum relaxation may be very substantial and may compete with the processes taken into account in the present study. However, this situation requires separate analysis. We shall illustrate the importance of allowing for two-quantum relaxation in a situation similar to that considered in Ref. 15 when the E_b level falls far from the edges of a wide photonic band gap. This assumption means that relaxation of the $E_b\to E_a$. transition can generally be neglected. Then, for the initially excited level E_b the ground level E_a would remain empty in this approach[15] and the coherent-wave excitation would be transferred from level E_b to level E_c and back. However, when allowance is made for two-quantum relaxation in the case $E_b-E_c<E_c-E_a$, the impurity atom will be transferred from the E_c state to the ground level E_a. This has the result that as a result of the depletion of the E_c level, the population of the E_b level will also decrease and within a time of the order of

$$\tau_{\rm rel}=\frac{\hbar^2}{\kappa_\theta|\mathscr{E}\Pi_{ca}(\omega_{bc})|^2}$$

an impurity atom from the excited state E_b located in the photonic band gap will be transferred to the E_a ground state. Here $\mathscr{E}$ is the amplitude of the coherent field in one-photon resonance with the $E_b\to E_c$ transition and κ_θ is the coupling constant with the photon thermostat having the central frequency $\omega_\theta=\omega_{ca}-\omega_{bc}$. For estimates we can use the formulas

$$\kappa_\theta=\frac{2\hbar\omega_\theta^3}{c^3},\quad \tau_{\rm rel}=\frac{\hbar c^4}{4\pi I\omega_\theta^3|\Pi_{ca}(\omega_{bc})|^2},$$

where I is the intensity of the coherent wave averaged over the period $2\pi/\omega_{bc}$ of the fast oscillations.

This work was supported by INTAS (Grant No. 96-0339).

*)E-mail: ashat@ashat.mephi.su, ashat@amicom.ru

[1] *Photonic Gap Materials*, edited by C. M. Soukoulis, NATO ASI Series E, Vol. 315 (Kluwer, Dordrecht, 1996).
[2] C. T. Chan, S. Datta, K. M. Ho, and C. M. Soukoulis, Phys. Rev. B **49**, 1988 (1994).
[3] U. Gruning, V. Lehmann, and C. M. Engelhardt, Appl. Phys. Lett. **66**, 3254 (1995).
[4] S. John and T. Quang, Phys. Rev. Lett. **74**, 3419 (1995).
[5] M. Salora, J. P. Dowling, C. M. Bowden, and M. J. Bloemer, Phys. Rev. Lett. **73**, 1368 (1994).
[6] J. C. Knight, T. A. Birks, R. F. Cregan, P. St. J. Russell, and J.-P. de Sandro, Opt. Mater. **11**, 143 (1998).
[7] S. John and J. Wang, Phys. Rev. B **43**, 12772 (1991).
[8] R. F. Nabiev, P. Yeh, and J. J. Sanchez-Mondragon, Phys. Rev. A **47**, 3380 (1993).
[9] G. Kweon and N. M. Lawandy, J. Mod. Opt. **41**, 311 (1994).
[10] V. I. Rupasov and M. Singh, J. Phys. A **29**, L205 (1996).
[11] V. I. Rupasov and M. Singh, Phys. Rev. Lett. **77**, 338 (1996).

[12] V. I. Rupasov and M. Singh, Phys. Lett. A **222**, 258 (1996).
[13] A. M. Basharov, A. I. Maĭmistov, and S. O. Elyutin, Zh. Éksp. Teor. Fiz. **115**, 30 (1999) [JETP **88**, 16 (1999)].
[14] S. John and T. Quang, Phys. Rev. A **54**, 4479 (1996).
[15] T. Quang, M. Woldeyohannes, S. John, and G. S. Agarwal, Phys. Rev. Lett. **79**, 5238 (1997).
[16] A. M. Basharov, Zh. Éksp. Teor. Fiz. **102**, 1126 (1992) [Sov. Phys. JETP **75**, 611 (1992)].
[17] M. Lax, Phys. Rev. **145**, 110 (1966).
[18] A. M. Basharov, *Photonics. The Unitary Transformation Method in Nonlinear Optics* [in Russian], Moscow Engineering Physics Institute, Moscow (1990).
[19] C. W. Gardiner, *Quantum Noise*, (Springer-Verlag, Berlin, 1991).
[20] A. Barchielli, Phys. Rev. A **34**, 1642 (1986).
[21] O. Kocharovskaya, Shi-Yao Zhu, M. O. Scully, P. Mandel, and Y. V. Radeonychev, Phys. Rev. A **49**, 4928 (1994).
[22] O. Kocharovskaya, P. Mandel, and M. O. Scully, Phys. Rev. Lett. **74**, 2451 (1995).
[23] H. J. Carmichael, Phys. Rev. A **56**, 5065 (1997).
[24] É. A. Manykin and V. V. Samartsev, *Optical Spectroscopy* [in Russian], Nauka, Moscow (1984).
[25] Yu. V. Rostovtsev, A. B. Matsko, and M. O. Scully, Phys. Rev. A **57**, 4919 (1998).
[26] A. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975) [Russ. transl., Atomizdat, Moscow, 1978].

Translated by R. M. Durham

Magnetic and electric dipole absorption in the dispersed plasma of a laser jet

N. E. Kask,*) S. V. Michurin, and G. M. Fedorov

D. V. Skobel'tsyn Institute of Nuclear Physics, Moscow State University, 119899, Moscow, Russia
(Submitted 17 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 1979–1989 (December 1999)

Experimental results are presented of a study of the magnetic and electric dipole absorption of microwave radiation by a laser plasma which accompanying the evaporation of various metals. An analysis is made of the influence of the finely dispersed fraction on the absorption and it is shown that the model of a percolation cluster provides a good explanation for the entire set of experimental results. © *1999 American Institute of Physics.* [S1063-7761(99)00812-4]

1. INTRODUCTION

In a pioneering study of a laser jet using microwave methods, Ageev *et al.*[1] observed the total reflection (cutoff) of a microwave from the plasma region in which the electron density exceeds a critical value. This cutoff occurred above the breakdown threshold of air and the time taken for the transparency to recover did not depend on the target material. It was shown in Ref. 2 that at moderate optical flux densities (which do not cause breakdown of the vapor-gas medium) it is possible to measure the microwave conductivity of a laser plasma. The conductivity is determined by the target material and is fairly high when both metals and dielectrics evaporate. By placing the jet in the antinode of the electric or magnetic field of a standing microwave, it is fairly easy to separate and measure the conduction current and eddy current losses.[3]

When various materials are evaporated by laser radiation, the material flux almost always contains a finely dispersed condensed phase: small droplets of melt expelled from the crater,[4] condensate microclusters,[5] and fractal structures[6] formed as a result of the cooling of the plasma as it expands. In addition, as the density of the metal vapor or its oxide increases in the plasma of the laser jet, a percolation transition takes place.[2] In the classical percolation model a critical cluster is formed by contacting conducting microparticles. In the microscopic model a percolation cluster appears as a result of the overlap of the wave functions of atoms and molecules for which the density of the material must be close to the critical point.[7] In a plasma containing a finely dispersed phase a percolation cluster may be formed on a chaotic lattice of noncontacting compact microclusters. Being a heat source in a laser radiation field, a microcluster is a center of spatial inhomogeneity in the temperature distribution and the electron density. Additional conductivity appears in the structure when the average distance between the lattice points is comparable with the characteristic spatial inhomogeneity scale of the plasma, for instance, the thickness of the electric double layer surrounding an isolated microcluster. According to our estimates,[8] this situation is achieved for a laser jet.

In the present paper we report results of an experimental investigation of the electric-dipole and magnetic-dipole absorption of a microwave in a low-temperature laser plasma and we analyze the role of the dispersed phase in the microwave absorption. This analysis is based on the conclusions drawn in theoretical studies[9,10] which investigated the optical properties of conducting particles in a dielectric matrix.

2. EXPERIMENT

2.1. Experimental method

Figure 1 shows an experimental apparatus which can be used to study microwave absorption at two frequencies and the optical flash from the laser jet. Heating 1.06 μm laser radiation, having a quasi-cw pulse duration of 10 ms and a power density in the range 10^5-10^6 W/cm^2, was focused onto the surface of the target in a 0.5 mm diameter spot. The laser radiation and the counterpropagating flux of evaporated material from the surface of the target (laser jet) were passed through 5 mm diameter apertures positioned at the center of the wide walls of two waveguide sections having cross sections of 8×17 and 10×23 mm^2. The sections were in contact with one another, having a single common wide wall. The directions of propagation of the 2 cm and 3 cm waves were orthogonal. By measuring the microwave absorption at two frequencies, we were able to compare the signals obtained in different experiments more accurately than in Ref. 3 by using one of the signals as a reference.

In order to study the microwave absorption signals, we used a reflection scheme in which the antinode of the electric or magnetic field of a standing microwave could be positioned in the plane of the jet by changing the position of a short-circuiting plunger in the corresponding waveguide cross section. Typical profiles of the laser pulse, the optical flash, and the microwave absorption signals in the antinodes of the electric and magnetic fields were given in Ref. 3 for a 3 cm wave. The coefficients of microwave absorption in the different antinodes cannot be compared directly since the signals, particularly the initial evaporation time, vary from one pulse to another. We used the $\lambda\approx2$ cm microwave absorption signal as the reference signal.

The reference signal can be used to normalize the absorption signal observed when the jet is located in the anti-

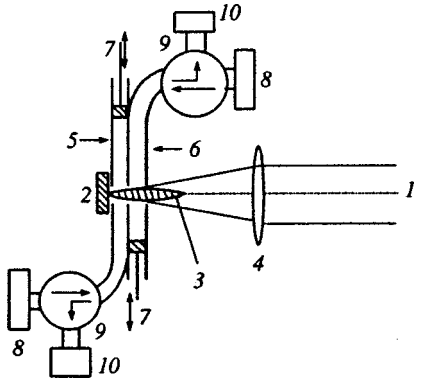

FIG. 1. Schematic of experiment: *1*—laser beam, *2*—target, *3*—jet, *4*—focusing lens ($f=17$ cm), *5*—2 cm waveguide section, *6*—3 cm waveguide section (turned through 90° about the axis of the laser beam relative to section *5*), *7*—short-circuiting plungers, *8*—microwave radiation generators, *9*—circulators, and *10*—microwave radiation detectors.

node of the electric or magnetic fields of the 3 cm wave relative to the absorption signal in the 2 cm waveguide where the position of the jet coincided with the antinode of the magnetic field and remained constant.

The shape and size of the jet are determined by the self-interaction effect in the laser-beam–plasma system and depends on the target material.[3] The time taken for the material to pass through the waveguide was much shorter than the duration of the laser pulse and did not exceed 0.5 ms. After a quasisteady-state evaporation regime had been established ($\leqslant 1$ ms) neither the transit time nor the transverse dimension of the jet depended on the irradiation time. According to Anisimov *et al.*,[4] in the quasisteady-state regime the density of the evaporated material is determined by the density of the heating laser radiation and the buffer gas pressure. It should be noted that at fairly high optical fluxes the dispersed phase in the plasma may be destroyed. In our experiments, when the laser pulse power was increased to 10^7 W/cm^2, a reduction in the luminescence intensity and the microwave conductivity of the jet is observed for some media (e.g. copper) and this can naturally be attributed to the breakup of the absorbing micro- and macroformations in the jet.

When we switched off the 3 cm klystron during the laser pulse, we did not detect any significant transfer of microwave energy from the 2 cm to the 3 cm waveguide channel. The fraction of the microwave radiation scattered by the jet material and escaping through the apertures in the waveguides was also negligible. Replacing the short-circuited 3 cm waveguide section with a transfer section having suitable facilities for detecting the transmitted radiation yields results in good agreement with those obtained for the standing wave. All the control experiments were carried out for the laser jet plasma and for various absorbing (conducting) samples having the same shape and size as the jet. Aqueous solutions of sodium chloride and binary powder mixtures were selected as the calibration media. The latter contained graphite or metal impurities to provide conductivity.

2.2. Plasma parameters

When material is evaporated by moderate-intensity laser pulses, the vapor pressure is equal to the buffer gas pressure. In earlier studies we measured the absolute values of the microwave conductivity and the brightness temperature in the visible part of the spectrum of a laser jet near the surface of various single-component media (including metals)[3] and their binary mixtures.[8] At normal pressure the microwave conductivity of single-component media is comparatively high ($>10^2$ S/cm) and the brightness temperature is close to the boiling point of the material. According to the Drude formula, this conductivity could be provided by an electron density in excess of 10^{15} cm^{-3}. Since the concentration of neutral particles in the jet is $\sim 10^{18}$ cm^{-3}, under conditions of ionization equilibrium at the boiling point of metals the electron density in the plasma is between 1.5 and two orders of magnitude lower than that required by the Drude formula.

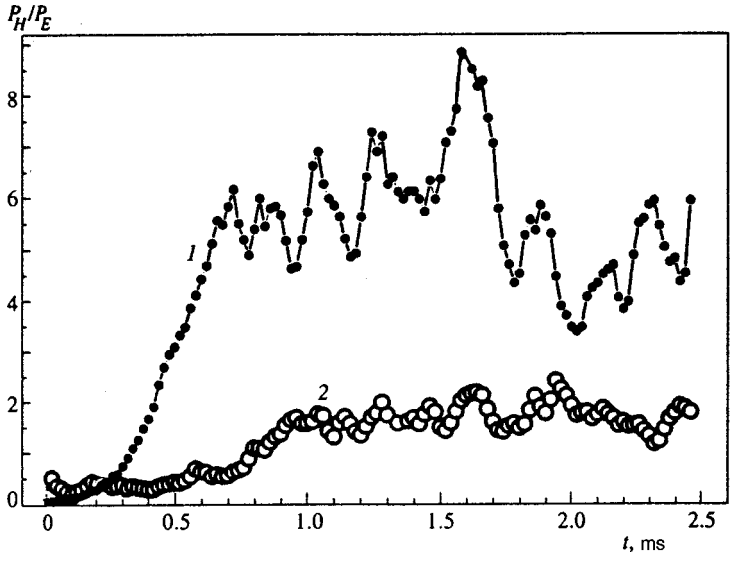

FIG. 2. Evolution of the ratio P_H/P_E during laser irradiation of metals: *1*—dysprosium and *2*—vanadium. Buffer gas—air.

Estimates[8] show that at normal pressure in the low-temperature plasma of a laser jet the concentration of compact microclusters may reach $\sim 10^{15}\,\mathrm{cm}^{-3}$.

In the plasma of a jet having transverse dimensions of the order of 1 mm, such high values of the conductivity lead to effective skinning of the microwave radiation and substantial mismatch of the waveguide channel. Thus, in the present study the measurements were made at gas pressures and optical flux densities for which the microwave power absorbed and reflected by the plasma is much lower than the power in the waveguide and distortions of the microwave field by the jet can be neglected. In the present experiments this condition was satisfied as long as the conductivity of the jet did not exceed 1 S/cm. The level of conductivity (absorption) for each material was selected by varying the buffer gas pressure in the range 10^{-4}–1 atm. These measurements show that as the pressure is reduced, the brightness temperature decreases in accordance with the saturation curve. Whereas the concentration of neutral particles can be assessed from the change in pressure and temperature, at reduced pressures the other plasma parameters of the laser jet require further study which we did not undertake.

Note that at reduced pressures simples estimates of the electron density using the Drude and Saha formulas yield an even larger mismatch than that obtained at normal pressure.

TABLE I. Values of the ratio P_H/P_E.

Target	P, Torr.	Buffer gas: Ar	Buffer gas: Air	r_{air}, mm
Graphite	30	1.5±0.5	2±1	1.8
Mg	70	0.5±0.1	0.5±0.1	3.2
Al	55	0.85±0.15	0.8±0.3	3.9
Si	55	0.6±0.3	0.6±0.6	3.2
Ti	25	1.0±0.5	1.0±0.6	2.2
V	22	2.0±0.5	1.7±0.5	1.1
Cr	33	1.0±0.2	0.45±0.25	1.0
Fe	45	0.20±0.05	0.23±0.05	1.1
Co	80	0.3±0.1	0.3±0.1	2.1
Ni	80	0.05±0.02	0.05±0.02	2.1
Cu	160	0.3±0.2	0.1±0.1	4.7
Y	30	0.50±0.25	1.5±0.5	1.2
Zr	30	0.75±0.25	1.00±0.25	1.3
Nb	100	2.0±0.5	2.5±0.5	1.0
Mo	70	1.75±0.50	1.5±0.5	1.2
La	5	1.0±0.2	10±1	1.2
Sm	5	4.0±0.5	13±2	1.1
Gd	5	0.75±0.25	0.2±0.1	1.1
Dy	5	2.5±0.5	6±1	1.2
Ho	5	6±2	12±2	1.4
Hf	30	0.2–2	0.2–1.0	1.45
Ta	55	2.0±0.2	2.0±0.5	0.9
W	80	1.2±0.6	1.0±0.5	1.2
Re	22	0.7±0.1	1.0±0.1	1.2
Al_2O_3	15	4±2	4±1	1.65

2.3. The ratio P_H/P_E for a laser jet plasma

Figure 2 shows typical behavior of the ratio of the magnetic-dipole to the electric-dipole absorption (subsequently denoted as P_H/P_E) during the establishment of advanced evaporation. For each metal the kinetics are obtained from the ratio of the 3 cm wave absorption signals in two successive laser pulses. The initial evaporation times were matched by using the absorption signals recorded in the 2 cm waveguide. After the beginning of evaporation ($t=0$), the ratio P_H/P_E usually increases monotonically and stabilizes at a certain level after quasisteady-state evaporation has been established. The initial stage of the kinetics clearly reflects an increase in the vapor density. Table I gives the average values of this level together with the error arising from the fact that the fluctuations of the magnetic-dipole and electric-dipole signals recorded in two successive shots differ. These fluctuations are several times the spread of the average level in different series of measurements of the microwave absorption. The error associated with the fluctuations can be minimized by calculating the ratio, not for the instantaneous time, but for the integral absorption in the steady-state section of the kinetics normalized to the absorption in the 2 cm waveguide. Figure 3 gives the ratio P_H/P_E as a function of the composition of a vanadium–iron mixture obtained by this method. A small change in the conductivity (approximately twofold) is accompanied by threshold behavior of the ratio P_H/P_E which may be attributed to percolation in the vanadium subsystem. It should be noted that the size of the jet varies negligibly with the mixture composition. Table I gives the transverse dimensions of the jet (r_{air}) for the various materials studied using air as the buffer gas. The radius was determined by a densitometric method using photographic images of the jet integrated over the visible part of the spectrum and the irradiation time, in the cross section corresponding to the center of the 3 cm waveguide. The measurement error did not exceed 10%.

2.4. Model experiment

Since percolation occurs in an optical discharge plasma, it is interesting to study the behavior of the ratio P_H/P_E near the percolation threshold for condensed composites conventionally used to study the metal–nonmetal transition. We present results of an investigation of microwave absorption in mixtures of graphite and magnesium oxide having a characteristic particle size of $\sim 1\ \mu$m. Cylindrical samples with $r=1$ mm (or 1.5 mm) and $l=3-30$ mm were drawn through the waveguides in the experimental system used to study the plasma jet.

Figure 4 gives the results of an experimental study of microwave absorption when the sample is placed in the antinodes of the electric and magnetic fields. As we predicted, the percolation threshold in the powder mixture corresponds to a volume fraction of the percolating component (p_v) of approximately 0.3. Experiments carried out using the same mixtures at dc current give the same value for the threshold. It can be seen from Fig. 4 that as the concentration of the percolating component (graphite) increases above the percolation threshold, the ratio P_H/P_E increases, approaching 0.3. In the range of concentrations where absorption takes place at isolated graphite particles, the ratio P_H/P_E is approximately 0.01. For concentrations greater than 0.5 this ratio is distorted by the skin effect which reduces the microwave

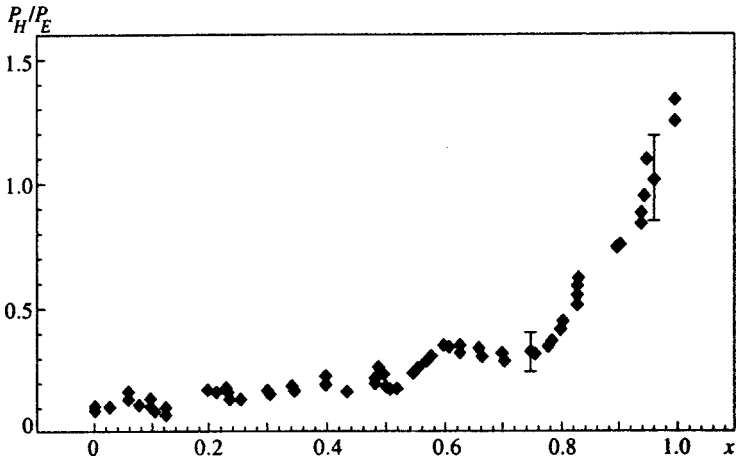

FIG. 3. Dependence of the ratio P_H/P_E on the composition of a vanadium–iron mixture. Buffer gas—air and x is the weight fraction of vanadium in the sample.

absorption in the electric and magnetic fields of the standing wave. Note that the skin effect weakly influences the ratio P_H/P_E.

3. DISCUSSION OF RESULTS

The powers dissipated by the material in the antinodes of the electric E and magnetic H fields of the standing microwave are respectively given by

$$P_E=\int \sigma(r)|E(r)|^2 d^3r, \quad P_H=\omega\int \chi''(r)|H(r)|^2 d^3r. \tag{1}$$

Here σ is the real part of the conductivity, χ'' is the imaginary part of the polarizability of the sample, and E and H are the intensities of the electric and magnetic fields of a microwave at frequency ω, respectively.

We shall analyze these results using the model of a homogeneous plasma and a plasma containing a dispersed phase.

3.1. Homogeneous electron–ion plasma flux

When the skin effect can be neglected, the ratio of the eddy current losses P_H (see Ref. 11) to the power P_E dissipated by the conduction currents in the plasma flux crossing the waveguide between the wide walls can be expressed in the following form:

$$\frac{P_H}{P_E}=\frac{\omega\chi''|H|^2V}{\sigma'|E|^2V}=\varepsilon_0\mu_0\omega^2\rho^2k, \tag{2}$$

where V is the volume occupied by a plasma of cross section $\pi\rho^2$ in the waveguide, ε_0 is the dielectric constant, and μ_0 is the magnetic constant. The coefficient $k=2.5$ allows for the change in the magnetic polarizability on changing from a sphere to a cylinder whose axis is perpendicular to the field.[12] It follows from Eq. (2) that the eddy current losses predominate over the conduction current losses if the transverse dimension of the jet is $\rho>0.3$ cm. In our case, the radius of the jet in the waveguide is approximately 1 mm and the ratio P_H/P_E should be less than 0.1. Control experiments using a column of aqueous electrolyte (NaCl) of the same size yielded $P_H/P_E\approx 0.03$. A similar value was obtained for the thermionic flux from copper and aluminum targets heated by an optical beam before the onset of evaporation and also at the initial stage of the evaporation process (see Fig. 2). It can be seen from Table I and Fig. 3, which give the ratios P_H/P_E, that in the advanced evaporation regime this ratio

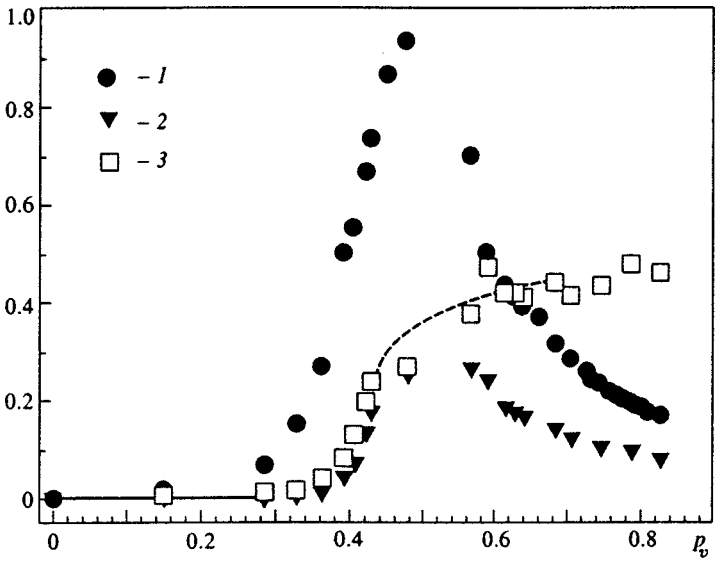

FIG. 4. Power absorbed by sample as a function of the volume fraction of graphite (p_v) in a graphite–MgO composite: *1*—electric-dipole and *2*—magnetic-dipole absorption in arbitrary units, *3*—the ratio P_H/P_E. The relative error in the determination of P_H/P_E is 20%.

may be substantially higher. It changes from one material to another and does not correlate with the size of the jet. These variations in the ratio are clearly related to the internal structure of the plasma and in particular, to the presence of a finely dispersed phase. Note that fairly large droplets ($\geqslant 10\,\mu$m) removed by the vapor-gas flux from the melt make no significant contribution to the microwave absorption because of the skin effect of the field in these droplets, nor to the optical luminescence because of their comparatively low temperature.

3.2. Role of compact microclusters

We know that composites comprising a dielectric impregnated with small metal particles exhibit anomalously high absorption in the infrared.[13] In this case, the ratio P_H/P_E is equal to the ratio of the coefficients of magnetic-dipole and electric-dipole absorption[9] and may be expressed in the following form:

$$\frac{P_H}{P_E}=\frac{8\pi^2\bar{a}^2\sigma_m^2}{45c^2\varepsilon_i}, \tag{3}$$

where ε_i is the dielectric constant of the dielectric, σ_m is the active conductivity of the metal, $\bar{a}$ is the average radius of the compact particles, and c is the velocity of light. Formula (3) was obtained in the approximation of long wavelengths ($\lambda \gg a$) and low frequencies ($\omega\tau \ll 1$, where τ is the relaxation time of the conduction electrons over the pulse). For particles whose conductivity is close to metallic, the eddy current losses predominate over the conduction losses if the radius of the metal particles is $a > 50$ Å. Since the characteristic size of the compact clusters formed in a laser plasma as a result of condensation does not exceed this value,[8] the existence of merely spherical compact metallic clusters cannot explain the maximum experimental values of the ratio P_H/P_E given in Table I. In addition, the transition from metallic to poorly conducting clusters (such as Al_2O_3, Si, or C) should lead to a sharp drop in the ratio P_H/P_E.

3.3. Role of an absorbing transition layer

The electric-dipole absorption of metal particles increases if these are surrounded by a thin poorly conducting layer ($\bar{a}\sigma_c t^{-1} \ll \sigma_m$, where t is the thickness of a layer having the conductivity σ_c). In this case, in the coefficient of electric-dipole absorption the metal conductivity σ_m is replaced by $\sigma_c\bar{a}/t$ (Ref. 9). Hence, for metal particles with a poorly conducting coating we can predict that the absorption will increase by five orders of magnitude if $\sigma_c = 1$ S/cm and $t \sim 0.1\bar{a}$. Such a poorly conducting layer could be formed by oxides or nitrides on the surface of microclusters or particles. Thus, when an inert buffer gas is replaced by air, we should expect a significant increase in the electric-dipole absorption compared with the magnetic-dipole absorption, particularly since the formation of poorly conducting layers should reduce the magnetic dipole absorption. The results presented in Table I do not support this assumption. In most cases, such a replacement is accompanied by a relative increase in the eddy current losses.

3.4. Role of filamentary and annular structures

When conducting particles coagulate to form filamentary structures, the electric-dipole absorption increases. First, the conductivity of the filament (σ_f) will be substantially lower than the conductivity of the particles themselves because of the presence of poorly conducting contacts. Second, an increase in the electric-dipole absorption is promoted by the absence of any polarization shielding if the filaments are oriented in the direction of the electric field. For a filament of length b the depolarization factor L has the form[12]

$$L=\left(\frac{a}{b}\right)^2\left[\ln\left(\frac{2b}{a}\right)-1\right], \tag{4}$$

which leads to an increase in the coefficient of electric-dipole absorption by the factor $\alpha_f = L^{-2}W_f\sigma_m/\sigma_f$ (Ref. 9), where W is the probability that a conducting particle is incorporated in the filament.

The eddy losses increase when large closed conducting loops appear in the system. For randomly oriented rings of radius $R \ll \lambda$, from Ref. 9 we have (to within a constant of the order of unity) the factor $\alpha_r = (R/a)^2 W_r$ in the coefficient of magnetic-dipole absorption, where W_r is the probability that a conducting particle is incorporated in the ring.

If the probability is $W \approx 1$, i.e., all the conducting particles are involved in forming filaments or rings, then for $R/a \approx 10$ (or $b/a \approx 10$) the corresponding coefficients are increased by two orders of magnitude. However, the formation of these structures with any appreciable probability was not observed in the laser jet. Note that the filamentary structures appearing after the action of the laser pulse are the result of the directional aggregation of cooled particles.[14]

3.5. Role of fractal structures

Quite clearly, branched fractal structures, in particular percolation structures which form at fairly high microparticle concentrations in a laser jet, can substantially influence the electric- and magnetic-dipole absorption since both linear and annular fragments exist in these structures. The nature of the material distribution in a fractal cluster leads to an effective decrease in the conductivity as the radius increase and consequently increases the electric-dipole absorption whereas the magnetic-dipole absorption in a fractal cluster is increased as a result of the dominant influence of a factor which depends on its size.

In accordance with Ref. 10, for a three-dimensional fractal cluster the amplification factors for electric-dipole and magnetic-dipole absorption are respectively given by

$$\left(\frac{R_f}{a}\right)^{5(3-d_f)/2} \quad \text{and} \quad \left(\frac{R_f}{a}\right)^{(1+d_f)/2}, \tag{5}$$

where R_f is the cluster size and d_f is its fractal dimensionality. Hence, the ratio P_H/P_E will increase if the fractal dimensionality exceeds 7/3 which is satisfied in the three-dimensional case for a percolation cluster ($d_f = 2.5$).

Note that the presence of percolation structures in the laser jet is indicated by characteristic threshold dependences of the microwave absorption and the optical luminescence of the laser jet on the composition of the target.[2,8]

3.6. Role of percolation structures

Percolation models are widely used to interpret the metal–nonmetal transition in disordered condensed media.[15,16] The most comprehensive experiments in this direction have been carried out for island films and composites. We know that as the percolation threshold is approached, the real part of the permittivity diverges.[17,9] According to Bowman and Stroud,[18] close to a threshold with an even larger critical index, the diamagnetic susceptibility diverges.

For two-dimensional objects, an anomalous absorption peak near the percolation threshold was observed experimentally in the infrared.[19] Generally various percolation models[17,19–21] use the electric-dipole absorption mechanism to interpret the peak parameters while reference is only made to the magnetic-dipole mechanism as a possible source of discrepancy between theory and experiment.

No theoretical analyses or numerical experiments to study absorption by a percolation cluster have been reported for the three-dimensional case. In this context, the results of our experiments to study microwave absorption in objects whose composition can be varied so they approach the percolation threshold are of particular interest (see Fig. 4). In accordance with formula (5) for a three-dimensional percolation cluster ($d_f=2.5$ and $R_f=2$ mm) the ratio P_H/P_E increases by a factor of 45 compared with the case of isolated ~ 1 μm particles which shows good agreement with the experimental results plotted in Fig. 4. It may be predicted that by reducing the particles sizes to around 0.01 μm this ratio will increase another ten times provided that R_f remains constant. Note that in a laser plasma the typical size of the compact clusters does not exceed ~ 1 nm (Refs. 5 and 8). Consequently, the high experimental values for the ratio P_H/P_E given in Table I can be explained quite satisfactorily by assuming that percolation clusters are present in a laser plasma.

4. CONCLUSIONS

Magnetic- and electric-dipole microwave absorption has been measured experimentally in the low-temperature plasma of a laser jet and in a condensed graphite–MgO composite. It has been established that the magnetic-dipole absorption increases more rapidly than the electric-dipole absorption as the laser jet evolves and their ratio may have values of the order of unity in the advanced evaporation regime. High values of the ratio P_H/P_E were also obtained for the condensed composite above the percolation threshold. A good qualitative explanation of these results can be provided by assuming that percolation structures exist in the low-temperature laser plasma.

This work was supported by the Russian Fund for Fundamental Research (Grants Nos. 96-02-17606 and 96-15-96537).

*)E-mail: nek@srdlan.npi.msu.su

[1] V. P. Ageev, V. I. Konov, T. M. Murina *et al.*, Kratk. Soobshch. Fiz. **5**, 6 (1978).
[2] N. E. Kask, JETP Lett. **60**, 212 (1994).
[3] N. E. Kask and G. M. Fedorov, Kvantovaya Élektron. (Moscow) **23**, 1033 (1996).
[4] S. I. Anisimov, Ya. A. Imas, G. S. Romanov, and Yu. V. Khodyko, *Action of High-Power Radiation on Metals* [in Russian], Nauka, Moscow (1970).
[5] B. M. Smirnov, Usp. Fiz. Nauk **164**, 665 (1994) [Phys. Usp. **37**, 621 (1994)].
[6] A. A. Lushnikov, A. E. Negin, A. V. Pakhomov, and B. M. Smirnov, Usp. Fiz. Nauk **161**, 113 (1991) [Sov. Phys. Usp. **34**, 160 (1991)].
[7] A. A. Likalter, Phys. Rev. B **53**, 1602 (1996).
[8] N. E. Kask, S. V. Michurin, and G. M. Fedorov, Teplofiz. Vys. Temp. **37**, 9 (1999).
[9] P. N. Sen and D. B. Tanner, Phys. Rev. B **26**, 3582 (1982).
[10] P. M. Hui and D. Stroud, Phys. Rev. B **33**, 2163 (1986).
[11] Y. Watanabe, K. Maeda, S. Saito, and K. Uda, Jpn. J. Appl. Phys. **16**, 2007 (1977).
[12] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, transl. of 2nd Russ. ed. (Pergamon Press, Oxford, 1984) [Russ original, later ed., Nauka, Moscow, 1992].
[13] P. Gadenne, Y. Yagil, and G. Deutscher, J. Appl. Phys. **66**, 3019 (1989).
[14] A. A. Lushnikov, A. E. Negin, and A. V. Pakhomov, Chem. Phys. Lett. **175**, 138 (1990).
[15] S. Kirkpatrick, Rev. Mod. Phys. **45**, 574 (1973).
[16] D. Stauffer and A. Aharony, *Introduction to Percolation Theory*, (Taylor and Francis, London, 1994).
[17] A. L. Efros and B. I. Shklovskii, Phys. Status Solidi B **76**, 475 (1976).
[18] D. R. Bowman and D. Stroud, Phys. Rev. Lett. **52**, 299 (1984).
[19] F. Brouers, J. P. Clerc, and G. Girand, Phys. Rev. B **44**, 5299 (1991).
[20] T. W. Noh, P. H. Song, S.-I. Lee *et al.*, Phys. Rev. B **46**, 4212 (1992).
[21] F. Brouers, J. P. Clerc, G. Girand *et al.*, Phys. Rev. B **47**, 666 (1993).

Translated by R. M. Durham

Exact solution of the problem of the equilibrium configuration of the charged surface of a liquid metal

N. M. Zubarev*)

Electrophysics Institute, Ural Branch of the Russian Academy of Sciences, 620049 Ekaterinburg, Russia
(Submitted 1 July 1999)
Zh. Éksp. Teor. Fiz. **116**, 1990–2005 (December 1999)

A broad class of exact solutions is obtained for the problem of the equilibrium configuration of the charged surface of a conducting liquid allowing for capillary forces. An analysis of the solutions showed that when the amplitudes of the perturbations reached certain critical values, the region occupied by the liquid ceases to be singly connected, which corresponds to the formation of liquid metal droplets. It is shown that a steady-state liquid metal profile may exist for which appreciable local amplification of the electric field can be achieved. © *1999 American Institute of Physics.* [S1063-7761(99)00912-9]

1. INTRODUCTION

As we know,[1,2] the plane surface of a conducting liquid (liquid metal) becomes unstable in a fairly strong electric field. A characteristic feature of this instability which is of considerable interest for the physics of the evolution of vacuum discharges and the operation of liquid-metal charged-particle emitters (see, for example, Refs. 3–5) is that nonlinearity does not saturate linear instability but is, in contrast, a destabilizing factor.[6,7] Consequently, even if the surface is stable in the linear approximation, a fairly large-amplitude perturbation can remove the system from equilibrium. The simplest problem which can be used to assess the role of nonlinearity, and in particular to find the threshold amplitudes of the surface perturbations for the hard excitation of instability, is the problem of the steady-state profile of a liquid metal surface in an external electric field. Of considerable interest also is the problem of the equilibrium configuration of a charged liquid-metal droplet which plays a key role in understanding the conditions for its breakup.[8]

A considerable number of studies (see, for example, Refs. 9–11 and the literature cited therein) have considered a dielectric liquid in an electric field using the approximation of small surface perturbations (the case of a liquid metal of interest to us is obtained in the limit of infinite permittivity). Significantly nonlinear liquid configurations have either been analyzed by numerical methods[12] or qualitatively.[13] As regards exact analytical solutions of the problem of the steady-state profile of a charged-liquid metal surface, we are presently only aware of a single particular nontrivial solution, the so-called Taylor cone.[14] However, the force balance condition for a Taylor cone is violated at the singularity, i.e., its vertex, so that this solution is of a formal nature.

In the present study we obtain a broad class of exact analytical solutions for the problem of the equilibrium configuration of a charged liquid-metal surface, taking into account surface tension forces, for the case of planar geometry when all the quantities depend on the pair of independent variables x and y. The approach used to solve the problem is based to a considerable extent on the method used by Crapper to solve the problem of the capillary wave profile at the free surface of an ideal liquid.[15] With a correction for the known analogy between the distributions of a planar electric field in the absence of space charge and the velocity field for the two-dimensional potential flow of an ideal liquid, this method is based on changing from Cartesian coordinates $\{x,y\}$ to new variables $\{\varphi,\psi\}$, where φ satisfies the equation for the Laplace potential of the electric field, and ψ is a function harmonically conjugate to φ (in Ref. 15 these functions had the meaning of the current function and the velocity potential, respectively). Since the surface of a conducting liquid is equipotential, and all the important quantities can be expressed in terms of the harmonic function $f=\ln|\nabla\varphi|$, the initial problem with an unknown boundary reduces to an analysis of a nonlinear boundary-value problem on the half-plane for the Laplace equation for the function $f(\varphi,\psi)$.

A similar procedure was applied to the problem of the equilibrium configuration of a two-dimensional charged liquid-metal droplet. It was found that the nonlinear equations obtained can be completely integrated. The short-wavelength limit of the exact solutions constructed by us (the solutions for this limit are the same as those obtained by Crapper for a capillary wave) corresponds to another important electrostatics problem, i.e., the steady-state profile of a liquid metal in an external electric field.

The article is made up as follows. In Sec. 2 we give the equations defining the equilibrium configuration of a charged liquid-metal surface for the case of plane symmetry. By analogy with Ref. 15, we show that using a conformal mapping method can reduce the investigation to solving a nonlinear boundary-value problem on the half-plane for the Laplace equation. In Sec. 3 we put forward an additional assumption on the distribution of the electric field outside the conducting liquid and we determine the necessary conditions for its compatibility with the initial equations. In Sec. 4 we obtain exact solutions for the problem of the equilibrium configuration of a charged two-dimensional liquid-metal droplet and also for the steady-state profile of a liquid metal in an external electric field. In Sec. 5 we construct and investigate the

equilibrium surfaces corresponding to these exact solutions. In the concluding section we discuss their stability with respect to small perturbations.

2. INITIAL EQUATIONS

1. We first consider the equilibrium configuration of a charged two-dimensional conducting droplet. Let us assume that a liquid having the free surface S occupies a certain bounded, singly-connected region in the plane $\{x,y\}$ (in the direction perpendicular to this plane the conductor is unbounded, i.e., it occupies a volume bounded by a right cylindrical surface in three-dimensional space). The distribution of the electric field potential φ (the field intensity is given by $\mathbf{E}=-\nabla\varphi$) in the absence of space charge is described by the Laplace equation:

$$\varphi_{xx}+\varphi_{yy}=0, \tag{1}$$

which should be considered jointly with the condition for an equipotential conductor surface:

$$\varphi|_S=0, \tag{2}$$

and also the condition that at some appreciable distance from the surface, the field of the charged conductor will be the same as the field generated by an infinite, charged, straight filament:[16]

$$\varphi\to-2q\ln r, \quad E=|\mathbf{E}|\to 2q/r, \quad r\to\infty, \tag{3}$$

where $r=\sqrt{x^2+y^2}$, and q is the electric charge per unit length of the conductor perpendicular to the plane $\{x,y\}$.

The equilibrium relief of the liquid metal boundary is determined by the balance condition for the forces acting on the surface:[16]

$$p+\frac{E^2}{8\pi}\bigg|_S+\frac{\alpha}{R}=0, \tag{4}$$

where α is the surface tension, p is the difference between the liquid pressure and the external pressure, and R is the radius of curvature of the surface.

For convenience we now go over to the dimensionless variables:

$$\mathbf{E}\to 4\pi\alpha q^{-1}\mathbf{E}, \quad \mathbf{r}\to q^2(2\pi\alpha)^{-1}\mathbf{r},$$

$$\varphi\to 2q\varphi, \quad p\to 2\pi\alpha^2 q^{-2}p.$$

Expressions (3) and (4) are then rewritten in the form

$$\varphi\to-\ln r, \quad E\to 1/r, \quad r\to\infty, \tag{5}$$

$$p+E^2|_S+R^{-1}=0, \tag{6}$$

and Eqs. (1) and (2) remain as before.

By analogy with Ref. 15, we introduce the auxiliary function ψ for which

$$\mathbf{E}=\{\partial\psi/\partial y,-\partial\psi/\partial x\}$$

(the condition $\psi=$ const defines the electric field lines). The complex expression $w=\varphi-i\psi$ (the so-called complex potential) is an analytic function of the complex variable $z=x+iy$ (Ref. 16). Then

$$\ln(-dw/dz)=f-i\theta \tag{7}$$

will also be an analytic function, this being the analog of the Zhukovskiĭ function in the theory of the plane potential flow of an incompressible fluid, where

$$f=\ln E, \quad \theta=\arctan(E_y/E_x).$$

The function θ has the meaning of the angle of inclination of the electric field intensity to the direction of the x axis. Since the intensity vector is normal to the surface of the conductor, we can assume that $\theta|_S$ defines the angle of inclination of the vector of the external normal to the surface S relative to the abscissa. Then, in accordance with the definition, the curvature of the surface S is given by: $R^{-1}=(\partial\theta/\partial s)|_S$, where ∂s is the elementary arc of the surface. As a result of the Cauchy–Riemann conditions for the functions f and θ we obtain: $R^{-1}=(\partial f/\partial n)|_S$, where $\partial/\partial n$ denotes the normal derivative. Converting to the function E, we find on the surface S

$$R^{-1}=\frac{1}{E}\frac{\partial E}{\partial n}=-\frac{\partial E/\partial n}{\partial\varphi/\partial n}=-\frac{\partial E}{\partial\varphi}.$$

That is to say, we have shown that curvature of the liquid metal surface is given by $R^{-1}=-(\partial E/\partial\varphi)|_S$. In this case, Eq. (6) is rewritten in the form

$$\frac{\partial E}{\partial\varphi}=p+E^2, \quad \varphi=0. \tag{8}$$

For the following analysis it is convenient to convert to a system of coordinates where φ and ψ play the role of independent variables. As a result of the analytic nature of Eq. (7), the function f in terms of the new variables will satisfy the Laplace equation

$$f_{\varphi\varphi}+f_{\psi\psi}=0 \tag{9}$$

with the condition at the conductor boundary derived from formula (8):

$$\frac{\partial f}{\partial\varphi}=pe^{-f}+e^f, \quad \varphi=0, \tag{10}$$

and also the condition at infinity:

$$f\to\varphi, \quad \varphi\to-\infty, \tag{11}$$

obtained from the expressions (5) by eliminating the spatial variable r. Bearing in mind that in the limit $|z|\to\infty$ for the complex potential we have[16] $w\to-\ln z$ and consequently a closed surface corresponds to changing ψ by 2π, we add the condition for periodicity of f with respect to the variable ψ:

$$f(\varphi,\psi)=f(\varphi,\psi+2\pi). \tag{12}$$

Thus, the problem of finding the steady-state profile of a charged liquid-metal two-dimensional droplet involves analyzing the boundary-value problem (9)–(12) on the half-plane $\varphi\leqslant 0$. Note that in order for the solutions of this problem to be physically meaningful, the corresponding surfaces must not allow self-intersection. However, using the conformal mapping method cannot ensure that this condition is satisfied at this stage.

2. We shall now consider the problem of the equilibrium configuration of the free surface of a conducting liquid in an external electric field E_0. We shall assume that the vector of

the external field intensity is directed along the y axis. The distribution of the electric field potential is described by the Laplace equation (1) combined with the condition for an equipotential liquid surface (2) and also the condition at infinity:

$$\varphi \to -E_0 y, \quad y \to \infty, \tag{13}$$

which means that the perturbations introduced in the field distribution by the surface decay over distance. The force balance condition for this problem has the form:

$$-\frac{E_0^2}{8\pi} + \frac{E^2}{8\pi}\bigg|_S + \frac{\alpha}{R} = 0, \tag{14}$$

which corresponds to the pressure difference $p = -E_0^2/(8\pi)$. The conditions (13) and (14) have the simpler form:

$$\varphi \to -y, \quad E \to 1, \quad y \to \infty,$$

$$\frac{\partial E}{\partial \varphi} = E^2 - 1, \quad \varphi = 0,$$

if we convert to dimensionless variables by substituting:

$$\mathbf{E} \to E_0 \mathbf{E}, \quad \mathbf{r} \to 8\pi\alpha E_0^{-2}\mathbf{r}.$$

Introducing the function $f = \ln E$ as before in coordinates where φ and ψ are the independent variables we obtain:

$$f_{\varphi\varphi} + f_{\psi\psi} = 0, \tag{15}$$

$$\frac{\partial f}{\partial \varphi} = -e^{-f} + e^f, \quad \varphi = 0, \tag{16}$$

$$f \to 0, \quad \varphi \to -\infty. \tag{17}$$

As far as the conditions for ψ are concerned, we shall assume that the function f is periodic:

$$f(\varphi, \psi) = f(\varphi, \psi + \lambda), \tag{18}$$

where λ is the wavelength in Cartesian coordinates ($w \to iz$ for $y \to \infty$ and thus, changing ψ to λ will correspond to changing the x coordinate to $-\lambda$).

We have therefore shown that the problem of determining the steady-state profile of a liquid-metal surface in an external electric field taking into account capillary forces can be reduced to an analysis of the nonlinear boundary-value problem (15)–(18). It should be noted that apart from the transformations, these equations are the same as those derived by Crapper,[15] who analyzed the steady-state profile of a progressive capillary wave. The analogy arises because from the mathematical point of view, the equations describing the two-dimensional electric field distribution in the absence of space charge are identical to the equations for the plane potential flow of an ideal fluid.

3. ADDITIONAL ASSUMPTIONS; COMPATIBILITY CONDITIONS

In order to find solutions of the nonlinear boundary-value problems (9)–(12) and (15)–(18), it is convenient to make the assumption that, in addition to the Laplace equation, the function f for $\varphi < 0$ satisfies the equation

$$\frac{\partial f}{\partial \varphi} = P(\varphi)e^{-f} + Q(\varphi)e^f, \tag{19}$$

where P and Q are certain functions of the variable φ, which satisfy the conditions

$$P(0) = p, \quad Q(0) = 1 \tag{20}$$

to agree with condition (10) or

$$P(0) = -1, \quad Q(0) = 1 \tag{21}$$

to agree with (16). We shall subsequently show that by using this assumption, we can reduce the solution of the initial boundary-value problems to the successive integration of ordinary differential equations.

We now explain the meaning of this additional relationship. Using the function E, Eq. (19) has the form of an ordinary Riccati differential equation with the independent variable φ:

$$\frac{\partial E}{\partial \varphi} = P(\varphi) + Q(\varphi)E^2, \tag{22}$$

or, reverting to dimensional quantities

$$pP(\varphi) + \frac{E^2 Q(\varphi)}{8\pi} + \frac{\alpha}{R'} = 0,$$

where R' is the radius of curvature of the equipotential surface passing through this point. This implies that on the equipotential surface $\varphi = \varphi_1 < 0$ the following relationship must be satisfied

$$\tilde{p} + \frac{E^2}{8\pi} + \frac{\tilde{\alpha}}{R'} = 0,$$

where we have introduced the notation

$$\tilde{\alpha} = \alpha/Q(\varphi_1), \quad \tilde{p} = pP(\varphi_1)/Q(\varphi_1),$$

which has the same structure as the force balance condition (4). In this case, our assumption implies that if all the necessary equilibrium conditions are satisfied on a certain surface (corresponding to $\varphi = 0$), any other surface $\varphi = \varphi_1$ will also be a solution of this problem for new values of the surface tension $\tilde{\alpha}(\varphi_1)$ and the difference between the internal and external pressures $\tilde{p}(\varphi_1)$.

Quite clearly, in the general case the systems of equations (9)–(12) and (15)–(18) combined with condition (19) are overdetermined. We find the values of P and Q for which the Laplace equation (9) or (15) will be compatible with the additional condition. Knowing the specific form of these functions we can solve Eq. (19) or, which amounts to the same thing, the Riccati equation (22) which for arbitrary values of $P(\varphi)$ and $Q(\varphi)$ is not integrable in quadratures. For this purpose we consider the general equation

$$\frac{\partial f}{\partial \varphi} = G(f, \varphi) \tag{23}$$

and we determine the dependence of the function G on the variables f and φ required for compatibility with the Laplace

equation. We first differentiate Eq. (23) with respect to φ and substitute the expression obtained into Eq. (9), which gives:

$$f_{\psi\psi}=-G_\varphi-GG_f.$$

Multiplying both sides by f_ψ and then integrating with respect to ψ, we find:

$$\frac{1}{2}f_\psi^2=A(\varphi)-\int (G_\varphi+GG_f)df, \tag{24}$$

where A is a certain function of the variable φ. Differentiating this expression with respect to φ gives:

$$f_{\varphi\psi}f_\psi=A_\varphi-2GG_\varphi-G^2G_f-\int G_{\varphi\varphi}df. \tag{25}$$

The cross derivative $f_{\varphi\psi}$ can be obtained by differentiating expression (23) directly with respect to ψ: $f_{\varphi\psi}=G_f f_\psi$. Substituting this relationship into Eq. (25) and then eliminating f_ψ using Eq. (24), we finally obtain the following integrodifferential equation for the function G considered as a function of the independent variables f and φ:

$$2AG_f-2G_f\int G_\varphi df=A_\varphi-2GG_\varphi-\int G_{\varphi\varphi}df. \tag{26}$$

This expression is a necessary condition for the compatibility of Eqs. (23) and (9).

We now substitute into the compatibility criterion (26) the function G in the form:

$$G(f,\varphi)=P(\varphi)e^{-f}+Q(\varphi)e^{f},$$

which corresponds to Eq. (19). We find:

$$(Q_{\varphi\varphi}+2AQ)e^{f}-(P_{\varphi\varphi}+2AP)e^{-f}$$
$$=A_\varphi-4P_\varphi Q-4PQ_\varphi.$$

This condition is clearly satisfied if

$$Q_{\varphi\varphi}+2AQ=0,$$

$$P_{\varphi\varphi}+2AP=0,$$

$$A_\varphi-4P_\varphi Q-4PQ_\varphi=0.$$

Integrating this last equation gives: $A=4PQ+s$, where s is a certain constant. Eliminating the function A from our equations by using this relationship, we finally obtain:

$$Q_{\varphi\varphi}=-2sQ-8PQ^2, \tag{27}$$

$$P_{\varphi\varphi}=-2sP-8QP^2. \tag{28}$$

Thus, we have shown that in order to achieve compatibility between Eqs. (9)–(12) and (15)–(18) and the auxiliary condition (19), the functions P and Q must satisfy the linear ordinary differential equations (27) and (28).

4. CONSTRUCTION OF EXACT SOLUTIONS

1. We shall now directly solve the boundary-value problem (9)–(12). Equations (27) and (28) can be integrated completely, although in order to obtain an exact solution of the problem of the equilibrium configuration of a charged liquid-metal two-dimensional droplet it is sufficient to analyze its particular solutions obtained by substituting:

$$P(\varphi)=-(1+X(\varphi))X_0e^{\varphi},\quad Q(\varphi)=X(\varphi)X_0^{-1}e^{-\varphi}, \tag{29}$$

where we have introduced the notation $X_0=(l-1)/2$ and $l=\sqrt{1-4p}$.

Using the representation (29) for P and Q, we obtain from Eqs. (27) and (28)

$$X_{\varphi\varphi}-2X_\varphi+(1+2s)X-8X^2(1+X)=0,$$

$$X_{\varphi\varphi}+2X_\varphi+(1+2s)(1+X)-8X(1+X)^2=0.$$

It is readily seen that the condition for compatibility of these equations will be

$$X_\varphi=2X+2X^2-1/4-s/2. \tag{30}$$

In order to satisfy condition (20) this equation must be solved jointly with the condition $X(0)=X_0>0$. For convenience we rewrite condition (30) in the form

$$X_\varphi=2(X-X_1)(X-X_2),$$

using the notation $X_{1,2}=-(1\pm k)/2$ and $k=\sqrt{s+3/2}$. We assume that $X_1<X_0<X_2$ and therefore $1<l<k$ (an analysis of other possible cases does not yield a solution of the initial problem). Integrating this ordinary differential equation then gives

$$X=\frac{X_2+X_1\exp\{2k(\varphi+\varphi_0)\}}{1+\exp\{2k(\varphi+\varphi_0)\}},\quad \varphi_0=\frac{1}{2k}\ln\left(\frac{X_2-X_0}{X_0-X_1}\right). \tag{31}$$

These relationships combined with the expressions (29) for the functions P and Q ensure that the necessary compatibility conditions are satisfied for Eq. (19) and the boundary-value problem (9)–(12).

We now use Eq. (19), which may be considered to be an ordinary differential equation with the independent variable φ (the variable ψ appears in it implicitly), to determine the dependence of f on φ. Substituting the expressions (29) into Eq. (19) and introducing the new function

$$F(\varphi,\psi)=\frac{X_0-e^{f-\varphi}}{X_0+e^{f-\varphi}},$$

we obtain the following linear differential equation:

$$\frac{\partial F}{\partial\varphi}=2XF+F+1.$$

Its solution is:

$$F=\frac{Z(\varphi)-Y(\psi)}{Z_\varphi},\quad Z(\varphi)=\int\exp\left\{-\int(2X+1)d\varphi\right\}d\varphi,$$

where Y is a certain function of the variable ψ. Substituting Eq. (31) into the expression for the function Z and adopting some arbitrariness in the choice of integration constants, we find

$$Z=\sinh(k\varphi+k\varphi_0). \tag{32}$$

Returning to the initial function f, we find that the solution of Eq. (19) can be written in the form

$$f(\varphi,\psi)=\varphi+\ln X_0+\ln\left(\frac{Z_\varphi-Z+Y}{Z_\varphi+Z-Y}\right). \tag{33}$$

Bearing in mind the expression for Z (32), we observe that for $\varphi\to-\infty$ we have $f\to\varphi$, i.e., the necessary conditions at infinity (11) are naturally satisfied.

We now find the unknown function Y. Substituting expression (33) into the Laplace equation (9), we obtain

$$\frac{Z_{\varphi\varphi\varphi}-Z_{\varphi\varphi}+Y_{\psi\psi}}{Z_\varphi-Z+Y}-\frac{(Z_{\varphi\varphi}-Z_\varphi)^2+Y_\psi^2}{(Z_\varphi-Z+Y)^2}$$
$$=\frac{Z_{\varphi\varphi\varphi}+Z_{\varphi\varphi}-Y_{\psi\psi}}{Z_\varphi+Z-Y}-\frac{(Z_{\varphi\varphi}+Z_\varphi)^2+Y_\psi^2}{(Z_\varphi+Z-Y)^2}.$$

After simple transformations we arrive at:

$$F_1(\psi)+Z(\varphi)F_2(\psi)+Z^2(\varphi)F_3(\psi)=0,$$

where

$$F_1=(Y^2-k^2)(Y_{\psi\psi}-k^2Y)-2Y(Y_\psi^2+k^2),$$
$$F_2=-2(YY_{\psi\psi}-Y_\psi^2-k^2+k^4),$$
$$F_3=(1-k^2)(Y_{\psi\psi}+k^2Y).$$

This becomes an identity provided that $F_1=F_2=F_3=0$, i.e., if the function Y satisfies three equations simultaneously:

$$(Y^2-k^2)(Y_{\psi\psi}-k^2Y)=2Y(Y_\psi^2+k^2), \tag{34}$$

$$YY_{\psi\psi}=Y_\psi^2+k^2-k^4, \tag{35}$$

$$Y_{\psi\psi}=-k^2Y. \tag{36}$$

We show that the set of solutions of this overdetermined system of ordinary differential equations is not empty. The solution of this last equation will then clearly be given by

$$Y=A\cos(k\psi-k\psi_0).$$

Without any loss of generality the integration constant ψ_0 can then be set to zero. Substituting this expression into Eqs. (34) and (35), we observe that this satisfies these equations when the following condition is satisfied for the amplitude A:

$$A^2=k^2-1.$$

Hence the solution of Eqs. (34)–(36) is given by

$$Y(\psi)=\sqrt{k^2-1}\cos(k\psi). \tag{37}$$

Thus, by using the additional relationship (19), we have obtained a solution of the Laplace equation (9) satisfying the conditions (10) and (11), i.e., formula (33) combined with expressions (32) and (37).

In order to construct the equilibrium surfaces corresponding to our solution (32), (33), and (37) in the coordinates $\{x,y\}$ in Sec. 5, we require expressions for the absolute electric field intensity $E|_{\varphi=0}$ relative to the conductor boundary and also for the angle of inclination of the electric field intensity vector to the direction of the abscissa $\theta|_{\varphi=0}$. Substituting the expressions for Y and Z into formula (33) and bearing in mind that $E=\exp f$, we obtain

$$E|_{\varphi=0}=X_0\frac{1+a^2d^2+2ad\cos(k\psi)}{a^2+d^2-2ad\cos(k\psi)}, \tag{38}$$

where we have introduced the notation:

$$a=\sqrt{\frac{k-1}{k+1}},\quad d=\sqrt{\frac{k-l}{k+l}}.$$

Then, assuming that f and θ are conjugate harmonic functions and thus the Cauchy–Riemann condition $\partial\theta/\partial\psi=\partial f/\partial\varphi$ is satisfied, we obtain from Eq. (10):

$$\frac{\partial\theta}{\partial\psi}=pe^{-f}+e^f=\frac{p}{E}+E,\quad \varphi=0.$$

Then substituting expression (38), integrating with respect to ψ, and selecting the integration constant such that $\theta(0,0)=\pi/2$, we find

$$\theta|_{\varphi=0}=\frac{\pi}{2}+\psi+2\arctan\left(\frac{a+d}{a-d}\tan\left(\frac{k\psi}{2}\right)\right)$$
$$-2\arctan\left(\frac{1-ad}{1+ad}\tan\left(\frac{k\psi}{2}\right)\right). \tag{39}$$

We analyze expressions (38) and (39) for the functions E and θ. Since the function E is periodic with the period $2\pi/k$ with respect to the variable ψ, a closed curve evidently corresponds to a change in ψ by an integer number of periods n, i.e., the value of ψ should vary in the range $0\leqslant\psi<2\pi n/k$. On the other hand, a closed surface corresponds to a change in ψ by 2π, i.e., condition (12). In fact, it is readily observed [see Eq. (39)] that as ψ increases by 2π, the angle θ changes (not necessarily monotonically) by the same amount, i.e., the curve plotted in Cartesian coordinates has a single complete revolution. This means that the condition $k=n$ should be satisfied, where $n\neq1$ is a natural number. In other words, the solutions of this problem will correspond to a series of numbers $k=2,3,4\ldots$, which define the number of branches of the corresponding curve in the coordinates $\{x,y\}$.

2. We shall now construct an exact solution for the problem of an equilibrium configuration of the free surface of a conducting liquid in an external electric field. Following Ref. 15, we shall seek a solution in the form

$$f(\varphi,\psi)=\ln\left(\frac{Z(\varphi)+Y(\psi)}{Z(\varphi)-Y(\psi)}\right), \tag{40}$$

where Y and Z are unknown functions of the variables ψ and φ, respectively. This representation corresponds to the particular case when the condition $P(\varphi)=-Q(\varphi)$, which does not contradict condition (21), holds for the functions P and Q in Eq. (19). Substituting expression (40) into Eqs. (15)–(18), we observe that the following conditions should be satisfied:

$$Z(\varphi)=(k/2+1)e^{-k\varphi}+(k/2-1)e^{k\varphi},$$
$$Y(\psi)=\sqrt{k^2-4}\cos(k\psi),$$
$$k=2\pi/\lambda,\quad k\geqslant2,$$

in conjunction with expression (40), which define the exact solutions of this boundary-value problem. Note that the pa-

rameter s in the compatibility conditions (27) and (28) is related to the wave number by the simple relationship: $k=\sqrt{s}$.

We then require the dependences of the functions E and θ on the variable ψ at the conductor surface. Substituting the expressions obtained for Z and Y into Eq. (40) and assuming that $E=\exp f$, we obtain for the absolute value of the electric field intensity:

$$E|_{\varphi=0}=\frac{1+c(k)\cos(k\psi)}{1-c(k)\cos(k\psi)}, \tag{41}$$

where we introduce the notation $c(k)=\sqrt{1-4/k^2}$. For the angle $\theta|_{\varphi=0}$, using the condition $\partial\theta/\partial\psi=\partial f/\partial\varphi$, we obtain from Eq. (16) after integrating with respect to ψ

$$\theta|_{\varphi=0}=\frac{\pi}{2}+2\arctan\left(\frac{kc(k)}{2}\sin(k\psi)\right). \tag{42}$$

It should be noted that similar expressions can be obtained from Eqs. (38) and (39) if the values of k and l are made to tend to infinity so that their ratio $\tilde{k}=2k/l$ remains finite. In fact, assuming that $\tilde{\psi}=\psi l/2$, and also $\tilde{c}(\tilde{k})=\sqrt{1-4/\tilde{k}^2}$, we obtain in the fundamental order:

$$E|_{\varphi=0}\to\frac{l}{2}\,\frac{1+\tilde{c}(\tilde{k})\cos(\tilde{k}\tilde{\psi})}{1-\tilde{c}(\tilde{k})\cos(\tilde{k}\tilde{\psi})},$$

$$\theta|_{\varphi=0}\to\frac{\pi}{2}+2\arctan\left(\frac{\tilde{k}\tilde{c}(\tilde{k})}{2}\sin(\tilde{k}\tilde{\psi})\right),$$

which agrees with Eqs. (41) and (42), apart from the constant factors whose appearance is related to the various methods of dedimensionalizing the quantities. This is because in the short-wavelength limit the problem of the equilibrium configuration of a charged liquid-metal droplet is identical to the problem of the steady-state profile of a liquid metal in an external electric field.

5. STEADY-STATE SURFACE PROFILES

1. We construct equilibrium profiles of the free surface of a liquid metal in the coordinates $\{x,y\}$ using Eq. (7), where the right-hand side is the function w. Taking this as an ordinary differential equation for the unknown function z, we obtain after integration

$$z=-\int\exp(-f+i\theta)dw.$$

Bearing in mind that at the boundary $\varphi=0$ and therefore $w=-i\psi$, we obtain the expression

$$z=z_0+i\int_0^{\psi}\frac{\exp(i\theta|_{\varphi=0})}{E|_{\varphi=0}}d\psi,$$

where $z_0=x_0+iy_0$ is the integration constant. Having separated the real from the imaginary part we observe that the unknown surfaces are defined parametrically as follows:

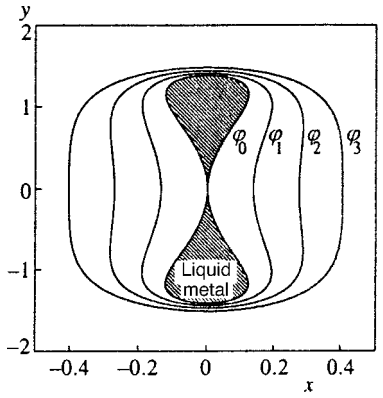

FIG. 1. Equilibrium configuration of charged two-dimensional liquid-metal droplet for $n=2$ and the critical value of the parameter $l=l_c(2)\approx 1.86$. Also plotted are typical equipotential surfaces corresponding to three different values of the parameter φ ($\varphi_1=-0.08$, $\varphi_2=-0.16$, and $\varphi_3=-0.24$). These surfaces may be considered to be a family of exact solutions of the problem, mapped in different scales and corresponding to different l.

$$y=y_0+\int_0^{\psi}\frac{\cos(\theta|_{\varphi=0})}{E|_{\varphi=0}}d\psi,\quad x=x_0-\int_0^{\psi}\frac{\sin(\theta|_{\varphi=0})}{E|_{\varphi=0}}d\psi. \tag{43}$$

When formulas (38) and (39) are used in the integrand expressions, these equations define the equilibrium configuration of a two-dimensional, liquid-metal droplet; the closed surface corresponds to a change in the parameter ψ in the range $0\leqslant\psi<2\pi$. The constants y_0 and x_0 can be conveniently taken so that the geometric center of the curve coincides with the origin.

It is easy to see that in the limit $l\to k=n$ expressions (38), (39), and (43) define circles of radius $2/(n+1)$. A reduction in the parameter l leads to an increase in the amplitude of the surface perturbations and it was found that for certain n-dependent critical values of the parameter $l=l_c(n)$, the region occupied by the liquid ceases to be singly connected and isolated liquid-metal droplets form. For $1<l<l_c$ the solutions are physically meaningless so that for fixed n the set of solutions of this problem corresponds to the interval $l_c(n)\leqslant l\leqslant n$.

Figures 1–3 show steady-state configurations of charged

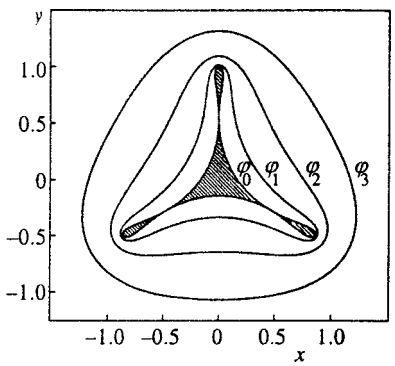

FIG. 2. Equilibrium configuration of a charged liquid-metal droplet for $n=3$ and the critical value of the parameter $l=l_c(3)\approx 2.53$. Also plotted is a family of equipotential surfaces corresponding to three different values of the parameter φ ($\varphi_1=-0.1$, $\varphi_2=-0.3$, and $\varphi_3=-0.6$).

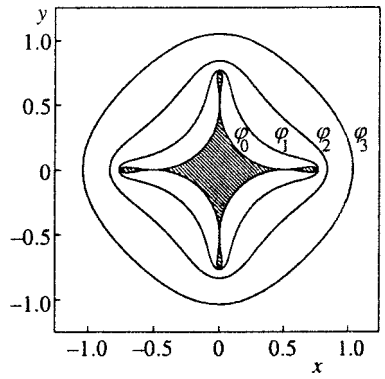

FIG. 3. Equilibrium configuration of a charged liquid-metal droplet for $n=4$ and the critical value of the parameter $l=l_c(4)\approx 3.19$. Also plotted are equipotential surfaces corresponding to three different values of the parameter φ ($\varphi_1=-0.1$, $\varphi_2=-0.3$, and $\varphi_3=-0.6$).

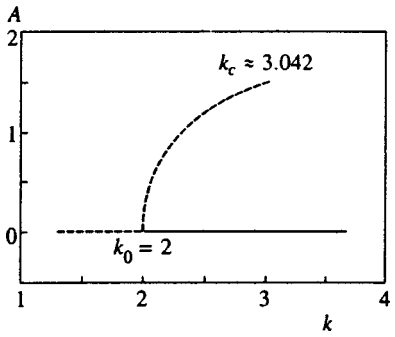

FIG. 4. Dependence of the amplitude A of the steady-state profile of the free surface of a liquid metal in an external electric field on the wave number k. The solid curve gives the stable branches of the solutions and the dashed curve gives the unstable branches.

two-dimensional liquid-metal droplets for critical l and $n=2$, $n=3$, and $n=4$. Also shown are typical equipotential surfaces corresponding to different values of $\varphi=\text{const}$ (we recall that at the liquid metal boundary $\varphi=0$). It was noted in Sec. 3 that these surfaces are also solutions of the problem of an equilibrium droplet configuration which refer only to those system parameters such as the surface tension and the difference between the internal and external pressures, which differ from the given ones. Thus, these curves not only give some idea of the electric field distribution outside a charged conducting droplet but also of its possible equilibrium configurations for various l and shown in different scales.

We introduce the characteristic of these solutions as the ratio of the maximum and minimum electric field intensities at the equilibrium surface $\gamma=(E_{\max}/E_{\min})|_S$. Quite clearly, for fixed n the highest possible amplification of the field is achieved for $l=l_c(n)$. For example, using Eq. (38) we obtain $\gamma_{\max}\approx 6.25$ for $n=2$, $\gamma_{\max}\approx 13.5$ for $n=3$, and $\gamma_{\max}\approx 18.5$ for $n=4$ ($\gamma_{\max}$ increases monotonically with any further increase in the integer-value parameter n). This demonstrates that a charged liquid-metal surface may have a steady-state profile which can ensure appreciable local amplification of the electric field, by more than an order of magnitude.

2. We now construct the steady-state profile of the free surface of a liquid metal in an external electric field for which expressions (41) and (42) should be substituted into the relationships (43) obtained above where we set: $y_0=0$ and $x_0=\pi/k$. Integration yields

$$y=1-c(k)-\frac{4k^{-2}}{1+c(k)\cos(k\psi)}, \tag{44}$$

$$x=\frac{\pi}{k}-\psi+\frac{2c(k)k^{-1}\sin(k\psi)}{1+c(k)\cos(k\psi)}, \tag{45}$$

which, apart from the substitutions $x\to -x$ and $y\to -y$, agrees with the expressions obtained by Crapper[15] as the exact solution of the problem of the capillary wave profile in coordinates moving with the wave (the liquid metal in our case corresponds to the space above the liquid surface in Ref. 15).

The parametric equations for the equilibrium surface (44) and (45) can be used to determine the dependence of the perturbation amplitude of the liquid-metal surface, defined as the difference between the maximum and minimum values of y per period ($A=y_{\max}-y_{\min}$), on the wave number k:

$$A=2\sqrt{1-4/k^2}. \tag{46}$$

Thus, in addition to the trivial solution $A=0$ of this problem for $k>k_0=2$, another nontrivial solution also appears, shown by the side branch in Fig. 4. It should be noted that a wave number of two in the initial dimensional variables corresponds to $k=k_0=E_0^2/(4\pi\alpha)$. It follows from the capillary wave dispersion law that at the plane surface of an ideally conducting fluid in an external electric field in the absence of a gravitational field[16]

$$\omega^2=\frac{\alpha}{\rho}k^3-\frac{E_0^2}{4\pi\rho}k^2,$$

where ρ is the density of the medium, this wave number is the threshold value: the trivial solution of the problem is stable with respect to small perturbations for $k>k_0$ whereas for $k<k_0$, aperiodic instability may develop (see Fig. 4).

It can be seen from Eq. (46) that as k increases, the amplitude A increases monotonically. When the wave number exceeds the critical value $k\approx k_c\approx 3.042$, the curve corresponding to expressions (44) and (45) becomes self-intersecting.[15] Consequently, the solutions of the boundary-value problem (15)–(18) given in the previous section are only physically meaningful when

$$k_0\leqslant k\leqslant k_c\simeq 1.52k_0,$$

and the largest possible amplitude of the surface perturbations will correspond to $k=k_c$. Figure 5 shows the equilibrium surface for this wave number. It can be seen that the region occupied by the liquid metal ceases to be singly connected and isolated liquid-metal droplets are formed, connected to the bulk of the metal by infinitely thin necks. Clearly, an arbitrarily small increase in the perturbation amplitude can lead to detachment of the droplets.

We give several figures characterizing the solution plotted in Fig. 5. The ratio of the amplitude of the surface perturbation to the wavelength is

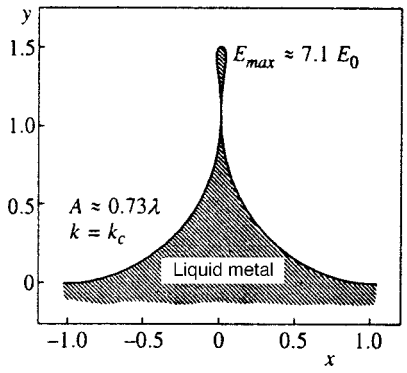

FIG. 5. Single period of the steady-state profile of the free surface of a liquid metal in an external electric field for the critical value of the wave number $k=k_c\approx 3.042$.

$$A_{\max}/\lambda = \pi^{-1}\sqrt{k_c^2-4}\approx 0.730.$$

The ratio of the maximum and minimum electric field per period, the parameter γ introduced above, has the highest possible value for $k=k_c$:

$$\gamma(k_c)=\gamma_{\max}=\left(\frac{2+A_{\max}}{2-A_{\max}}\right)^2\approx 50.64.$$

Moreover, the absolute value of the electric field at the equilibrium surface may exceed the external field E_0 by a factor of $\gamma^{1/2}$, i.e., $E_{\max}\approx 7.116\,E_0$. At the tip the curvature of the surface is more than fifty times its minimum value per period.

6. CONCLUDING REMARKS

As a result of analyzing the force balance conditions at the free surface of a conducting liquid with plane symmetry, we have obtained a broad class of exact solutions for two of the most interesting problems, i.e., the problem of the equilibrium configuration of a two-dimensional charged liquid-metal droplet and also the problem of the steady-state profile of a liquid metal in an external electric field. The question naturally arises as to whether these solutions are stable with respect to small perturbations. A comprehensive study of this aspect is outside the scope of the present paper in which we have not investigated the surface dynamics as such. Nevertheless, we have put forward several arguments which indicate that the liquid metal configurations plotted in Figs. 1–3 and 5 are unstable.

We consider the equilibrium surface of a liquid metal in an external electric field defined by expressions (44)–(45). The dependence of the amplitude of the perturbation of the free surface on the parameters of the problem is given by the relationship

$$A(k)=\frac{4}{k}\left(\frac{16\pi^2\alpha^2k^2}{E_0^4}-1\right)^{1/2},$$

obtained from Eq. (46) by converting to dimensional quantities. It can be seen that for fixed k the amplitude of the steady-state solution increases as the external electric field decreases. Conversely, if we increase the amplitude of the surface perturbation, the force balance condition can only be satisfied when the external field is slightly reduced. For the previous value of the field the capillary forces will not be able to compensate for the destabilizing influence of the electrostatic forces, which leads to a further increase in amplitude, i.e., to the development of instability. The same may also apply to liquid-metal droplets.

This reasoning also indicates that the plane surface of a liquid metal in an external electric field is globally unstable with respect to perturbations having wave numbers $k>k_0$, i.e., when the condition for its linear stability is satisfied. The instability can increase without bound if the amplitude of the surface perturbation exceeds the equilibrium value of $A(k)$ determined by us. A similar condition may well serve as the simplest criterion for the hard loss of stability by the plane surface of a conducting liquid.

The author is grateful to E. A. Kuznetsov for stimulating discussions, and also to A. M. Iskol'dskiĭ and N. B. Bolkov for their interest in this work. This work was partly financed by the Russian Fund for Fundamental Research (Project No. 97-02-16177).

*)E-mail: nick@ami.uran.ru.

[1] L. Tonks, Phys. Rev. **48**, 562 (1935).
[2] Ya. I. Frenkel', Zh. Éksp. Teor. Fiz. **6**, 347 (1936).
[3] I. Yu. Bartashyus, L. I. Pranevichyus, and G. N. Furseĭ, Zh. Tekh. Fiz. **41**, 1943 (1971) [Sov. Phys. Tech. Phys. **16**, 1535 (1971)].
[4] L. W. Swanson and G. A. Schwind, J. Appl. Phys. **49**, 5655 (1978).
[5] A. L. Pregezer and B. M. Marder, J. Appl. Phys. **60**, 3821 (1986).
[6] N. M. Zubarev, Phys. Lett. A **243**, 128 (1998).
[7] N. M. Zubarev, Zh. Éksp. Teor Fiz. **114**, 2043 (1998) [JETP **87**, 1110 (1998)].
[8] J. W. Strutt, Philos. Mag. **14**, 184 (1882).
[9] M. I. Shliomis, Usp. Fiz. Nauk **112**, 427 (1974) [Sov. Phys. Usp. **17**, 153 (1974)].
[10] E. A. Kuznetsov and M. D. Spektor, Zh. Éksp. Teor. Fiz. **71**, 262 (1976) [Sov. Phys. JETP **44**, 136 (1976)].
[11] G. V. Kolmakov and E. V. Lebedeva, Zh. Éksp. Teor. Fiz. **115**, 43 (1999) [JETP **88**, 24 (1999)].
[12] A. Ramos and A. Castellanos, J. Electrost. **33**, 61 (1994).
[13] V. V. Vladimirov and V. N. Gorshkov, Appl. Phys. A **46**, 131 (1988).
[14] G. I. Taylor, Proc. R. Soc. London, Ser. A **280**, 383 (1964).
[15] G. D. Crapper, J. Fluid Mech. **2**, 532 (1957).
[16] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, (Pergamon Press, Oxford, 1984; Nauka, Moscow, 1982).

Translated by R. M. Durham

Liquid–solid phase transition in a heterogeneous system of solid spheres

V. Ya. Rudnyak*) and A. A. Belkin

Novosibirsk State Architectural and Civil Engineering University, 630008 Novosibirsk, Russia
(Submitted 2 March 1999)
Zh. Éksp. Teor. Fiz. **116**, 2006–2011 (December 1999)

A molecular dynamics method is used to study the influence of a heavy dispersed particle on a liquid–solid phase transition in a molecular system of solid spheres. It is shown that the presence of a dispersed particle shifts the transition toward higher densities and pressures. In addition, in the liquid state the heterogeneous system has a lower pressure, whereas in the solid state the pressure is higher than that in the corresponding states of a homogeneous medium. It is established that the pressures of the heterogeneous mixture in the regions before and after the phase transition can be converted to the pressure of the single-component system by introducing various scaling factors (effective densities of the medium). © *1999 American Institute of Physics.* [S1063-7761(99)01012-4]

1. INTRODUCTION

In their pioneering work Alder and Wainwright[1] used a molecular dynamics method to establish that a liquid–solid phase transition exists in a homogeneous system of solid spheres. However, characteristic features of this phase transition in an inhomogeneous system have not yet been studied. The widespread occurrence of heterogeneous media in nature and in various technological applications makes their study a topical and important issue. This particularly applies to studies of the phase transition. The present paper is devoted to studying the influence of small dispersed particles on the liquid–solid phase transition.

We consider the evolution of a heterogeneous system of solid spheres comprising a homogeneous molecular system of solid spheres of radius r and mass m in which is immersed a dispersed particle of radius R and mass M. We studied the evolution of the system by a molecular dynamics method.[2] The mass ratio was taken to be $M/m=100$. The model was a cell in the form of a rectangular parallelepiped which contained a single particle surrounded by molecules. Initially the molecule was located at the points of a hexagonal close-packed lattice (this structure is typical, for example, of crystalline helium ^{3}He and ^{4}He). The ratio of the particle and molecular radii varied experimentally in the range $R/r=2-4$. In order to create an empty region in which the particle was located, N_m^* molecules were removed from the cell. In order to compensate for the finite number of molecules studied in this cell, we used periodic boundary conditions. All the numerical experiments were carried out using a 533 MHz DEC-Alpha computer. Between 200 and 5800 molecules were used for the calculations. The time taken to calculate a single point on the isotherm diagrams with $N_m \sim 2500$ molecules in the cell near the phase transition zone (the time taken for relaxation of the system to the liquid state is particularly long here) is around 10 h.

2. RESULTS OF CALCULATIONS

The phenomenon of a liquid–solid phase transition in a system of solid spheres is illustrated clearly in Fig. 1. Here the solid curve shows a typical isotherm of a single-phase liquid. This isotherm is the dependence of the compressibility $\kappa=(p-p_0)/p_0$ on the packing density of the molecules in the system

$$\alpha_f=\frac{V}{V_0}=\frac{V\sqrt{2}}{8r^3N_m^0}, \tag{1}$$

where p is the pressure in the system, $p_0=nkT$ is the ideal gas pressure, and $N_m^0=N_m+N_m^*$. The temperature T is determined by the mean squared molecular velocity $3kT=m\langle v^2\rangle$. The molecular dynamics method is used to calculate the coordinates and velocities of all the molecules. These calculated dynamic variables of the system are then used to determine the thermodynamic variables. In particular, the virial theorem (see, for example, Ref. 3) is used to calculate the pressure in the system.

In a pure single-phase liquid, as the density of the system increases the compressibility has a discontinuity at a certain density $\alpha_f^*\approx 1.472$ (see Fig. 1) and at higher densities it goes over to a new branch corresponding to the solid state. This phenomenon is interpreted as a first-order phase transition.

In physical terms, the phase transition in the solid sphere model can be explained as follows. Initially the molecules are situated at points in a hexagonal close-packed lattice. This packing corresponds to the maximum average distance between neighboring molecules for a given density. In this case, the correction to the pressure caused by the molecular interaction is minimal. If the density of the system is fairly high, the molecule remains near the lattice point for the entire time. This medium is an analog of a solid. As the molecular concentration decreases, the ordered structure is destroyed and the system is converted to the ''liquid'' state, whereupon the pressure increases abruptly.

If one of the molecules in this particular molecular system is replaced by its ''heavy isotope'' (for example, the mass of the substituted molecule may be $100m$, but the radii of the heavy and light molecules are the same), the character

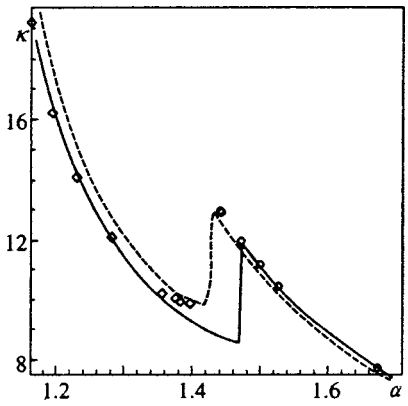

FIG. 1. Dependences of the compressibility factors of a ''pure liquid'' on α_f (solid curve) and of a binary mixture ($R=4r$, $N_m=2637$) on α_{fp} (dashed curve), α_{fp}^s (squares), and α_{fp}^l (circles).

of the observed phase transition changes little. The compressibility is again described by the solid curve in Fig. 1.

However, the nature of the phase transition changes substantially if a large dispersed particle is added to the homogeneous liquid. For the calculations the ratio of the particle and molecular masses was taken as $M/m=100$ and the ratio of their radii was taken to be $R/r=2, 3, 4$. The parameters of this heterogeneous system are the ratios of the volume n_p/n_f and mass ρ_p/ρ_f concentrations. Here and subsequently the subscripts ''p'' and ''f'' refer to the dispersed particles and molecules of the carrier medium, respectively. Thus, for instance, we have $n_f=N_m/V$, where V is the total volume of the cell. The density of this mixture is determined by the parameter [cf. Eq. (1)]

$$\alpha_{fp}=\frac{V\sqrt{2}}{8r^3N_m+8R^3}, \quad N_m+N_m^*=N_m^0. \tag{2}$$

In the calculations plotted in Fig. 1 the dashed curve gives the isotherm for this heterogeneous system (liquid + dispersed particle). Here the radius of the dispersed particle is $R=4r$ and the cell contains $N_m=2637$ molecules. When a dispersed particle is present in the system, the phase transition takes place at higher densities and pressures. The density at which the phase transition takes place increases which corresponds to a decrease in the parameter α_{fp} to 1.43. In this case, the compressibility κ increases from 11.9 to 12.95.

In addition, in the liquid state the pressure in the mixture is lower than that in a single-phase liquid of the same density. This is because in the liquid state the entire region around the particle is accessible to molecules. The addition of a dispersed particle to the homogeneous molecular liquid reduces the accessible volume of the cell for the molecular carrier medium by the particle volume $V_p=4\pi R^3/3$ and therefore reduces the number of molecules. The effective density parameter of the mixture in the liquid state can thus be written as

$$\alpha_{fp}^l=\frac{(V-V_p)\sqrt{2}}{8r^3N_m}. \tag{3}$$

It is easy to establish that $\alpha_{fp}<\alpha_{fp}^l$, and so a dispersed system in the liquid state has a lower pressure than a homogeneous system, as shown in Fig. 1 (right-hand branch of the isotherm for $\alpha_{fp}>1.43$).

Conversely, if the mixture remains in the solid state shielded empty regions form around the particle. The effective particle volume increases in the solid state. Quantitatively this is equal to the volume which would be occupied by N_m^* removed molecules. The corresponding density parameter is given by

$$\alpha_{fp}^s=\frac{V\sqrt{2}}{8r^3(N_m+N_m^*)}. \tag{4}$$

Since $\alpha_{fp}^s<\alpha_{fp}$, a dispersed system in the solid state is effectively denser than a homogeneous system (left-hand branch of the isotherm for $\alpha_{fp}<1.43$ in Fig. 1).

Note that if the branches of the isotherms in the liquid and solid states of a heterogeneous liquid are plotted as functions of the parameters α_{fp}^l and α_{fp}^s, respectively, we arrive at the branches of the isotherms for a homogeneous liquid. These branches are indicated by the circles and squares, respectively, in Fig. 1. Thus, we can say that the parameters α_{fp}^l and α_{fp}^s are scaling factors which can be used to construct the isotherms of a homogeneous liquid if we know the isotherms of the heterogeneous system.

The increase in the density α_{fp}, at which a phase transition is observed is also related to shielding effects. The critical density α_{fp}^* for a heterogeneous system, i.e., the density at which the isotherm may have a discontinuity, may be obtained from the critical density α_f^* for a homogeneous liquid:

$$\alpha_{fp}^*=\alpha_f^*\frac{\alpha_f}{\alpha_{fp}^l}. \tag{5}$$

An analysis of the data plotted in Fig. 1 indicates that the presence of a dispersed particle in a molecular system substantially alters the nature of the phase transition in the system. However, it is clear from physical reasoning that as the volume density n_p of the dispersed particles decreases, their influence on the phase transition in the system should decrease. The experiments confirm this viewpoint. Figure 2 gives dependences of the critical density α_{fp}^* on the reduced particle concentration $\nu_p=4\pi r^3n_p/3$ for various particle radii. The scale on the abscissa is logarithmic. The horizontal line corresponds to the critical density α_f^*, and the theoretical curves are obtained using formula (5). It can be seen that for $R=2$ and $\nu_p\lesssim 3\times10^{-4}$ the influence of the particles on the isotherms can be neglected. As the radius of the dispersed particles increases, the boundary of the ''concentrations of influence'' becomes increasingly lower.

Using the virial theorem is not the only method of making a numerical analysis of the equation of state for a system of solid spheres. It is easily shown that for a system of solid spheres, the compressibility κ is proportional to the volume density and the pair distribution function of the molecules at the interaction point n_2 (see, for example, Ref. 4). The function $n_2(r')$ is the concentration of molecules at the distance

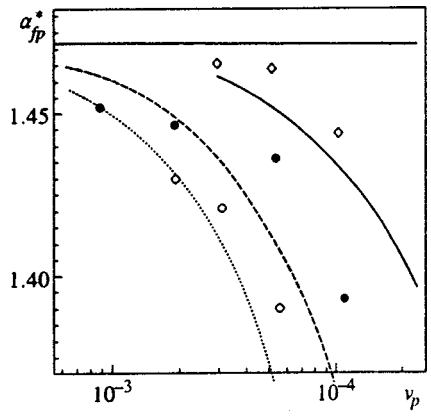

FIG. 2. Dependences of the critical density of a mixture on the particle concentration ν_p. Results of numerical experiments: $R=2$ ($\diamond$), $R=3$ ($\bullet$) and , $R=4$ ($\circ$). Theoretical curves: $R=2$ (solid curve), $R=3$ (dashed curve), and $R=4$ (dotted curve).

r' from the center of one of these, normalized to the average molecular concentration n_f. Figure 3 gives calculated pair distribution functions of a homogeneous molecular liquid (solid curve) and a heterogeneous system (dashed curve, $R=4$, $N_m=2637$) for the same density $\alpha=1.446$, i.e., in the liquid state for a homogeneous system and in the solid state for a heterogeneous one (see Fig. 1). This is confirmed by Fig. 3. The pair function for the heterogeneous system is periodic even for large r', which is typical of the ordered structure of a solid. The first maximum of the function n_2 for a mixture is substantially higher than that for a homogeneous system and thus the pressure is the liquid state is higher. By measuring the value of n_2 at the point of the first maximum, we can calculate the pressure in the system. The isotherms calculated by both methods show good agreement over almost the entire density range (the relative deviation is $<2\%$ for $\alpha>1.3$). Appreciable changes are only observed in the solid region far from the phase transition point. In this region, however, the adequacy of this equation of state is somewhat doubtful.

3. CONCLUSIONS

The solid-sphere potential is presently the main model for studying dense gases and liquids using direct numerical modeling methods. The distances between the molecules in these media are short so that the main factor in the intermolecular interactions is the repulsive component of the potential which is accurately modeled by the solid-sphere potential. In this sense, these results not only accurately describe the qualitative pattern of a liquid–solid phase transition in a heterogeneous system but can also provide important quantitative information.

The modeling of a phase transition in a heterogeneous system performed in the present study was limited to low concentrations of dispersed particles and small sizes. The low concentrations of the dispersed particles (10^{-3}–10^{-5}) imply that the influence of interparticle interaction could be neglected. We essentially studied the influence on the phase transition of an isolated dispersed particle immersed in a molecular medium. The dispersed particles were either large molecules or nanoparticles, in particular clusters. The widespread development of nanotechnologies over the last decade makes it relevant to study such small particles.

The size of the dispersed particle plays a decisive role in liquid–solid phase transition processes. This is because the transition itself is of a geometrical nature to a considerable extent. A variation in the mass of the dispersed particle between $50m$ and 10^3m negligibly influences the nature of the isotherms although of course, heterogeneous media containing dispersed particles of different mass will have different relaxation times.

Of particular note is the fact that these results can be used to calculate the isotherms of heterogeneous systems from the corresponding data for a homogeneous molecular system (at least at fairly low concentrations of dispersed particles).

It is predicted that the observed behavior of the phase transition process will be conserved as the particle sizes increase. In order to check this, we made trial calculations of several isotherm points for a system with $N_m=13341$ and $R=8r$. A shift of the transition toward higher densities is also observed in this system. The error in estimates of the shift using these formulas does not exceed at most 30%.

The authors thank the referee for comments which have resulted in improvements to the content of the article.

This work was supported by the Russian Fund for Fundamental Research (Grants Nos. 98-01-00719 and 96-15-96310).

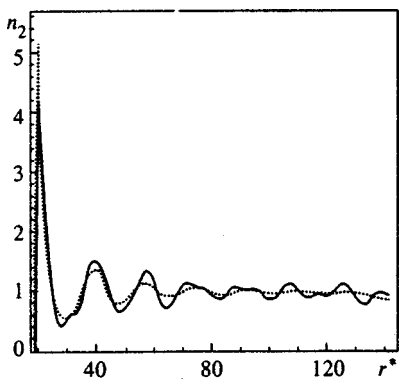

FIG. 3. Pair distribution functions of a pure gas (solid curve) and a mixture (dashed curve) of the same density ($\alpha=1.446$, $r^*=r'/r$).

*)E-mail: rudyak@ngasu.nsk.su

[1] B. J. Alder and T. E. Wainwright, J. Chem. Phys. **27**, 1208 (1957).
[2] V. Ya. Rudyak, G. V. Kharlamov, A. A. Belkin, *Direct Numerical Simulation of Transport Processes in Heterogeneous Media.* I. *Diffusion Coefficient of a Brownian Particle*, Preprint No. 2(12)–98 [in Russian], Novosibirsk State Architectural and Civil Engineering University, Novosibirsk (1998).
[3] D. N. Zubarev, *Nonequilibrium Statistical Thermodynamics* (Consultants Bureau, New York, 1974) [Russ. original, Nauka, Moscow, 1971].
[4] C. A. Croxton, *Liquid State Physics. A Statistical Mechanical Introduction* (Cambridge University Press, Cambridge, 1974) [Russ. transl., Mir, Moscow, 1978].

Translated by R. M. Durham

Structure, melting, and potential barriers in mesoscopic clusters of repulsive particles

Yu. E. Lozovik*) and E. A. Rakoch

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia
(Submitted 12 January 1999)
Zh. Éksp. Teor. Fiz. **116**, 2012–2037 (December 1999)

This paper discusses two-dimensional mesoscopic clusters of particles that repel according to dipole, Coulomb, and logarithmic laws and are confined by an external parabolic potential. These models describe a number of physical systems, in particular, electrons in semiconductor structures or on a liquid-helium surface allowing for image forces, indirect excitons in coupled semi-conductor dots, and a small number of vortices in an island of a second-order superconductor or in superfluid helium. Two competing forms of ordering are detected in the particles in the mesoscopic clusters—the formation of a triangular lattice or of a shell structure. The temperature dependences of the potential energy, the mean-square radial and angular deviations, the radial and angular distributions of the particles, and the distribution of the particles over the local minima are studied. Melting in mesoscopic clusters occurs in two stages: at lower temperatures, there is orientation melting, from the frozen phase into a phase with rotational reorientation of ''crystalline'' shells with respect to each other; subsequently, a transition occurs in which the radial order disappears. Melting in dipole macroclusters occurs in a single stage. However, in Coulomb and logarithmic macroclusters, orientation melting occurs only for the outer pairs of shells. Orientation melting is also detected in three-dimensional Coulomb clusters. A connection is established between the character of the melting and the ratio of the energy barriers that describe the breakdown of the orientational and radial structure of a cluster. [S1063-7761(99)01112-9]

1. INTRODUCTION

Much attention has been paid in recent years to the study of the properties of systems made up of a finite number of particles with a repulsive interaction potential. Coupled systems made up of a small number of particles whose physical properties do not coincide (because of their small size) with the properties of crystals are called clusters.

We shall consider here clusters with logarithmic, Coulomb, and dipole laws for the interaction between the clusters. Each of these systems has interesting physical implementations (see below).

Clusters having different interaction laws and a small number of particles possess many common properties, in particular, a shell structure competing with the appearance inside the cluster of a nucleation center with the ''bulk-phase'' structure (i.e., a triangular lattice for two-dimensional systems). We shall consider mesoscopic clusters that possess a shell structure. They are an intermediate case between microscopic clusters consisting of one shell and macroscopic clusters in which a large part of the particles form the bulk phase. For example, in the two-dimensional case, a large part of the particles inside the cluster form a fragment of a slightly distorted two-dimensional triangular lattice. In mesoscopic clusters, the number N of particles varies between 6 and 50–100, depending on the interaction law between the particles in the cluster (the longer-range is the interaction law, the more numerous the particles that correspond to macroscopic clusters). The shell structure of a mesoscopic cluster can sharply change when only one ''particle'' is added (structural sensitivity) all the way to some number N of particles, at which a region with the structure of the bulk phase appears within this cluster. What is most interesting, the melting of a mesoscopic cluster can possess interesting specific features by comparison with the melting of the bulk phase.[1–3] It turns out that these mesoscopic clusters melt in two stages—the mutual-orientation melting of the shells first occurs, and the shell structure disappears at a higher temperature (magic clusters are an exception). As will be shown, these features are common to mesoscopic clusters of different types, and this makes it reasonable to consider them jointly. At the same time, the criterion for a cluster to be mesoscopic depends on the range of the interaction between the particles. Namely, the transition from mesoscopic to macroscopic clusters (at which the features of mesoscopic clusters mentioned above disappear) occurs at a lower number of particles in dipole clusters than, for example, in Coulomb and logarithmic clusters.

The orientation melting indicated above is also possible in an extended system made up of repulsive particles existing in an external (random) field created by impurities, defects, boundary roughness, etc. A structure recalling a cluster also forms close to the minimum of a random potential (or close to individual defects, if their concentration is small), and orientation melting can occur here as the temperature is increased (it was observed in Ref. 4 for a vortex lattice in an impurity system).

We shall now briefly consider physical implementations of the clusters under consideration.

We shall first concentrate on the most important physical

implementation for logarithmic clusters. A magnetic field penetrates into a second-order superconductor in the form of Abrikosov vortices, which form an ideal triangular lattice in the low-temperature region.[5] As the temperature increases, this lattice can melt and form a liquid phase made up of vortices, as has been observed for high-temperature superconductors (see the review in Ref. 6). There is great interest in the question of the structure of a vortex system with a small number of vortices included in islands of the superconducting phase.

Moreover, the vortices in a rotating vessel with superfluid helium also interact according to a logarithmic law when the number of vortices is small. The electrons in a semiconductor nanostructure surrounded by a medium with low permittivity also interact according to a logarithmic law in a certain interval.[7]

The system under consideration is equivalent to the problem of the two-dimensional analog (see Ref. 1) of the classical Thomson atom[8] with a finite number of ''charges'' obeying the laws of two-dimensional electrostatics and with a compensating incompressible background.

Physical implementations of two-dimensional Coulomb clusters are, for example, electrons in craters on a liquid-helium surface[9] and electrons in a quantum dot.[10] A system of electrons in a three-dimensional quantum dot is an analog of a Thomson atom.

When the image forces close to the semiconductor-metal boundary are taken into account, the Coulomb interaction law is replaced at large distances by a dipole law, and this is reflected in the phase diagram of the system, leading to reversible quantum melting of a Wigner crystal, with a change in the density.[11] Repulsion at large distances also follows a dipole law for excitons with spatially separated electrons and holes,[12–16] particles in a layer of magnetic liquid, a layer of dielectric clusters on the surface of an electrolyte, etc. (see Ref. 17 and the literature cited therein).

For this paper, we considered the temperature dependence of the cluster structure. We show that a two-dimensional mesoscopic cluster melts in two stages: inter-shell (orientation) melting occurs first, and the shells smear out at a significantly higher temperature. (Clusters with a small number of particles but with a structure close to a true triangular lattice—magic clusters-can be an exception.) However, a macrocluster (in the interval of the numbers of particles studied here) melts in one or two stages, depending on the interaction law between the clusters. It is shown that this is because the energy barrier relative to reorientation of the shells is substantially less than the barrier with respect to radial jumping of the particles in the case of microclusters, whereas these barriers are of the same order of magnitude in the case of macroclusters. We also have analyzed how the indicated potential barriers disappear as the temperature increases.

For this paper, we consider the classical regime for electrons, in which their characteristic de Broglie wavelength is much less than the mean distance between them [the quantum regime, in particular, the quantum melting of clusters (see, for example, Ref. 18) is not considered here].

Section 2 briefly describes the physical model. Section 3 presents the numerical methods used in this paper. Section 4 describes the configurations of the clusters at the global and local minima of the potential energy. Section 5 presents the results of calculations of the melting in the clusters. Section 6 analyzes the potential barriers when the shells in the clusters rotate with respect to each other and when the particles jump from one shell to another. Section 7 discusses how the anisotropy of the confining potential affects the structure and the melting of the clusters. Section 8 discusses the melting of three-dimensional Coulomb clusters. Conclusions are presented in Section 9.

2. THE PHYSICAL MODEL

We shall consider two-dimensional clusters with dipole, Coulomb, and logarithmic laws for the interaction between particles confined by external potential $U_{\text{ext}}(r)$. For electrons in a semiconductor nanostructure, the role of the confining potential is played by the boundary of the semiconductor nanostructure. For electrons above a helium film, the role of lateral confining potential can be played by the potential of a (small) metallic electrode immersed in the helium. For vortices in a superconducting island, the role of confining potential is played by the compensating incompressible background of charges of opposite sign (effectively taking into account the correct mean vortex density in the given magnetic field, which determines the minimum of the Ginzburg-Landau functional in the field). In all these cases, the confining potential can be calculated for small clusters from the quadratic expression $U_{\text{ext}}(r_i)=\alpha r_i^2$, where α is a positive constant.

For two-dimensional dipole clusters (the dipoles are perpendicular to the plane of the cluster) with pairwise interaction $U_{ij}=D^2/r_{ij}^3$, we carry out the scale transformations

$$r\to\frac{\alpha^{1/5}\epsilon^{1/5}}{D^{2/5}}r,\quad T\to\frac{k\epsilon^{2/5}}{\alpha^{3/5}D^{4/5}}T,\quad U\to\frac{\epsilon^{2/5}}{\alpha^{3/5}D^{4/5}}U. \quad (1)$$

Likewise, for the Coulomb interaction law $U_{ij}=q^2/r_{ij}$, we obtain dimensionless quantities by means of the transformations

$$r\to\frac{\alpha^{1/3}\epsilon^{1/3}}{q^{2/3}}r,\quad T\to\frac{k\epsilon^{2/3}}{\alpha^{1/3}q^{2/3}}T,\quad U\to\frac{\epsilon^{2/3}}{\alpha^{1/3}q^{2/3}}U. \quad (2)$$

However, for the logarithmic interaction $U_{ij}=-q^2\ln r_{ij}$, we use the transformation

$$r\to\frac{\alpha^{1/2}}{q}r,\quad T\to\frac{k_B}{q^2}T,\quad U\to\frac{1}{q^2}U. \quad (3)$$

As a result, the potential energy of a cluster acquires the form

$$U=\sum_{i>j}U_{ij}+\sum_i r_i^2, \quad (4)$$

where $U_{ij}=1/r_{ij}^3$, $U_{ij}=1/r_{ij}$, and $U_{ij}=-\ln r_{ij}$, respectively, for the cases considered above.

3. NUMERICAL METHODS USED IN THIS PAPER

To search for the equilibrium configurations of the particles, we used a random search for the potential-energy minimum of the system, with random motion of the individual shells as a whole and random motion of the particles. It is convenient to take regular polygons inscribed in circles as the initial configurations in the case of an isotropic confining potential, since the shells have shapes close to circles, and it is convenient to take regular polygons inscribed in ellipses as the initial configurations in the case of anisotropic confinement. In the case of very strong anisotropy, the particles were placed on a straight line as the initial configurations. The maximum step was decreased from 1×10^{-2} to 1×10^{-6} in dimensionless units. Each 10^3 steps, the step was decreased by a factor of 0.8–0.98.

The potential barrier against rotation (allowing for the extremely substantial relaxation, i.e., alignment of the particles to rotation of the shell) can be found by the following procedure: We fix all the particles of the cluster, other than the particles of a specific shell, in the configuration corresponding to the global minimum of the potential energy. We rotate all the particles of the indicated shell by angle $\delta\phi$ with respect to the others. We then fix the angle of one particle from the rotated shell and the angle of one particle from the stationary shell, and we find the minimum of the potential energy in terms of $2N-2$ variables by the method of random search (N variable radii r and $N-2$ variable angles ϕ). We note the minimum potential energy of this system. We then repeat this procedure, varying ϕ, until one shell has been rotated with respect to the other by an angle equal to the mean angular distance between the particles in the rotating shell. It is possible in this way to find the dependence of the potential energy of the system on the angle of rotation ϕ of one shell and the potential barrier against rotation.

The potential barrier against the jumping of a particle from one shell to another characterizes the radial (total) melting of the shells. We use the following procedure to find it: We fix all the particles of the cluster in the configuration corresponding to the global minimum of the potential energy, except for one particle at the site, and we move this particle away from the center of the system (or toward the center) toward its position at the local minimum of the potential energy by distance δr. We then fix the distance from this particle to the center of the system, and we find the minimum of the potential energy in terms of $2N-1$ variables by the method of random search ($N-1$ variables r and N variables ϕ). We note the minimum potential energy of this system. We then repeat this procedure until the system is at a local minimum of the potential energy. It is possible in this way to find the dependence of the potential energy on the coordinate r of a particle ''by changing shell'' and the potential barrier against jumping of a particle. The proposed method of taking into account the ''relaxation'' of the positions of the particles as the shells rotate or as the particles jump between shells substantially reduces the barriers with respect to reorientation of the shells and with respect to jumping. Otherwise, as shown by calculation, unrealistic, overestimated values of these barriers are obtained (this effect is especially large in the former case).

We can use the techniques described above to find the potential barriers in clusters at zero temperature. However, the variation of the barriers with temperature is of interest. The technique for finding the ''temperature'' potential barriers differs from that for the ''nontemperature'' ones as follows: Instead of finding the minimum in terms of $2N-2$ variables in the case of the barrier against rotation or in terms of $2N-1$ variables in the case of the barrier against jumping, in the case of the temperature barrier, we find the mean potential energy for fixed T and $\Delta\phi$ in the case of the angular barrier and fixed T and r in the case of the radial barrier. The averaging is carried out over approximately 1×10^5 Monte Carlo steps.

The Monte Carlo method with the Metropolis algorithm was used in this paper to study the dependence of the physical quantities on the temperature and the melting of the system. After the equilibrium configurations were found, the temperature of the system was increased by ΔT ($\Delta T=5\times10^{-7}-5\times10^{-3}$), and the system was then confined until it reached equilibrium at the new temperature in $(2-4)\times10^4$ Monte Carlo steps. The statistical characteristics were then calculated by averaging over 1×10^6 Monte Carlo steps. This was followed by further heating, using the procedure described here. The following quantities were computed:

(1) The total potential energy U_{pot}.

(2) The radial mean-square displacements: the total value,

$$\langle \delta R^2\rangle = \frac{1}{N}\sum_{i=1}^{N}\frac{\langle r_i^2\rangle-\langle r_i\rangle^2}{a^2}, \tag{5}$$

and the individual value for each shell,

$$\langle \delta r^2\rangle = \frac{1}{N_R}\sum_{i=1}^{N_R}\frac{\langle r_i^2\rangle-\langle r_i\rangle^2}{a^2}, \tag{6}$$

where N_R is the number of particles in a shell, and the averaging, indicated by angle brackets, is carried out over different Monte Carlo configurations.

(3) The angular mean-square displacements with respect to the nearest particles of its own shell,

$$\langle \delta\phi_1^2\rangle = \frac{1}{N_R}\sum_{i=1}^{N_R}\frac{\langle(\phi_i-\phi_{i_1})^2\rangle-\langle(\phi_i-\phi_{i_1})\rangle^2}{\phi_0^2}, \tag{7}$$

and with respect to the nearest particles of the neighboring shell,

$$\langle \delta\phi_2^2\rangle = \frac{1}{N_R}\sum_{i=1}^{N_R}\frac{\langle(\phi_i-\phi_{i_2})^2\rangle-\langle(\phi_i-\phi_{i_2})\rangle^2}{\phi_0^2}, \tag{8}$$

where i_1 and i_2 relate to the closest particle from the same and from the neighboring shell, and $2\phi_0=2\pi/N_R$ is the mean angular distance between the neighboring particles for the given shell.

Only the relative angular mean-square deviations are calculated, since the rotation of the system as a whole is of no interest.

TABLE I. Shell structure and potential energy of clusters made up of two-dimensional vortices.

Number of particles	Occupation numbers of shells	Potential energy	Number of particles	Occupation numbers of shells	Potential energy
1	1	0.000000×10^{0}	26	3,9,14	-1.940569×10^{2}
2	2	5.000000×10^{-1}	27	3,9,15 Cr_3	-2.156137×10^{2}
3	3	8.918023×10^{-1}	28	4,9,15	-2.384294×10^{2}
4	4	1.090457×10^{0}	29	4,10,15	-2.625912×10^{2}
5	5	9.764052×10^{-1}	30	4,10,16	-2.881028×10^{2}
6	1,5	4.354169×10^{-1}	31	4,10,17	-3.149268×10^{2}
7	1,6 Cr_1	-7.512442×10^{-1}	32	4,11,17	-3.431329×10^{2}
8	1,7	-2.514746×10^{0}	33	5,11,17	-3.727473×10^{2}
9	1,8	-4.914510×10^{0}	34	1,5,11,17	-4.037308×10^{2}
10	2,8	-8.100414×10^{0}	35	1,6,11,17	-4.361606×10^{2}
11	3,8	-1.209333×10^{1}	36	1,6,12,17	-4.700331×10^{2}
12	3,9	-1.697858×10^{1}	37	1,6,12,18 Cr_1	-5.053534×10^{2}
13	4,9	-2.271610×10^{1}	38	1,6,12,19 Cr_1	-5.420929×10^{2}
14	4,10	-2.942793×10^{1}	39	1,7,13,18	-5.803155×10^{2}
15	4,11	-3.706118×10^{1}	40	1,7,13,19	-6.200430×10^{2}
16	5,11	-4.573707×10^{1}	41	1,7,13,20	-6.612310×10^{2}
17	1,5,11	-5.541308×10^{1}	42	1,7,14,20	-7.039416×10^{2}
18	1,6,11	-6.620692×10^{1}	43	2,8,14,19 Cr_2	-7.481666×10^{2}
19	1,6,12	-7.811655×10^{1}	44	2,8,14,20 Cr_2	-7.939606×10^{2}
20	1,6,13	-9.110199×10^{1}	45	2,8,14,21 Cr_2	-8.412619×10^{2}
21	1,7,13	-1.052696×10^{2}	46	3,9,14,20 Cr_3	-8.901514×10^{2}
22	1,7,14	-1.205683×10^{2}	47	3,9,15,20 Cr_3	-9.406122×10^{2}
23	1,8,14	-1.370647×10^{2}	48	3,9,15,21 Cr_3	-9.926554×10^{2}
24	2,8,14	-1.548203×10^{2}	49	3,9,15,22 Cr_3	-1.046250×10^{3}
25	3,8,14	-1.737968×10^{2}	50	4,10,15,21	-1.101460×10^{3}
107	3,9,15,21,27,32 Cr_3	-7.155730×10^{3}			
108	3,9,15,21,27,33 Cr_3	-7.316694×10^{3}	192	3,9,15,21,27,33,39,45 Cr_3	-2.834568×10^{4}

The following quantities were also calculated for two-dimensional Coulomb clusters with $N=19$ and $N=20$.

(4) The radial distribution function

$$g(r)=\left\langle \sum_{i=1}^{N} \delta(r_i-r) \right\rangle .$$

(5) The correlation function of the angles between the particles of two shells:

$$g(\phi)=\left\langle \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} \delta(\phi_i)\,\delta(\phi_j-\phi_i+\phi) \right\rangle .$$

(6) The distribution of particles over local minima, i.e., the probability of detecting the system close to different minima of the potential energy $W_{\mathrm{glob,loc}}$. To calculate this quantity, the system was periodically (once per 10^4-10^5 iterations) sharply cooled to a temperature several orders of magnitude lower than all the melting temperatures (1×10^{-7}) for 2×10^4 iterations. The minimum energy was then compared with the energy at the different local minima, and, if it was close to the value at some minimum, it was assumed that the system fell into this minimum (the local and global minima were calculated ahead of time).

4. EQUILIBRIUM STRUCTURE OF TWO-DIMENSIONAL CLUSTERS

We shall investigate the equilibrium structure of clusters with logarithmic and dipole interaction laws. The Coulomb interaction is intermediate in range. The equilibrium structure of Coulomb clusters was investigated earlier (see, for example, Refs. 1 and 3).

Mesoscopic clusters have a shell structure at low temperatures, and this is explained by the influence of the centrally symmetric confining potential. As is well known, an unbounded two-dimensional crystal has a triangular lattice. However, the triangular structure originates inside a cluster with a rather large number of particles (see also Ref. 3). In the case of dipole clusters, because of the short-range character of the interaction, the triangular lattice originates when $N>11$, but it is difficult to distinguish the shells when $N>40$; therefore we have studied dipole clusters with $N\leqslant40$. In logarithmic clusters, a triangular lattice is formed at much larger N; therefore we investigated logarithmic clusters with $N\leqslant50$ and selectively for larger N. The following were found for two-dimensional clusters: (a) the local, and (b) the global (the deepest of the local) minima of the potential energy.

4.1. The structure of logarithmic clusters

The dependences of the internal, external, and total potential energy per particle on the number of particles are close to linear. In this case, the dependences of the internal and total potential energy per particle on the number of particles N are decreasing and smooth for N from 3 to 50, while the dependence of the external potential energy per particle as the number of particles N goes from 1 to 50 is increasing and smooth.

Table I shows the occupation numbers of the shells and the corresponding potential energies [measured from $C_N^2 \ln(q/\alpha^{1/2}a)$—see above] for the global minima of two-dimensional logarithmic clusters. The successive filling of the shells recalls to some extent the periodic table of the elements (compare the filling of the shells in a two-dimensional classical system of particles with three-dimensional Coulomb interaction[3]). Each shell can contain no more than a definite number of particles. Thus, there can be no more than five particles in the first shell (counting from the center of the system) and no more than 11 particles in the second. When all the shells are filled, i.e., when they contain the maximum number of particles, a new shell begins to appear: first one additional particle appears at the center of the system, then (as the number of particles increases), two particles form an inner shell, etc. In this case, one particle appears at the center of the system after a particle is added to a system with configuration (5,...), two appear after configuration (1,7,...) or (1,8,...), three after configuration (2,8,...), four after configuration (3,9,...), and five after configuration (4,11,...).

We shall use the following concept of shell (which differs from the definitions used earlier, which are not universal for different numbers of particles). We define a shell as a convex polygon made up of the maximum possible number of particles (inside which the preceding shell is found), satisfying the following rule: The maximum distance from a particle of the given shell to the center of the system must be less than the minimum distance to the center of the system from a particle of the neighboring shell, external with respect to the given shell. As a result of this definition, the law for filling the shells in clusters, obtained from our calculations using the definition indicated above, is similar to the law for filling the shells in the Thomson atom.[3] In particular, in our work, unlike Ref. 19, the number of shells cannot decrease as the number of particles increases.

Since the confining quadratic potential is centrosymmetric, it should be expected that the shells in a cluster must have the shape of regular polygons inscribed in circles. However, this is valid only for clusters consisting of one shell or of two shells the inner of which has only one particle. As the number of particles in clusters with small N increases further, the symmetry spontaneously breaks down. This shows up most strongly in a cluster with two particles at the center (in the clusters with $N=10$ and $N=24$). In these cases, since the first shell (counting from the center) consists of two particles, it is convenient to assign the shape of an ellipse to the second shell.

The basis of certain configurations at the global minima of the potential energy consists of parts of a crystal lattice with hexagonal symmetry. When describing and analyzing the properties of such configurations, it is suitable to introduce into the discussion ''crystal shells'' (Cr_c)—concentric groups of nodes of an ideal two-dimensional crystal at the center of which are found c particles (see Table I).

An analytical calculation of the radius and potential energy of clusters consisting of one shell, as well as of clusters consisting of two shells the second of which contains only one particle, confirms that the computer calculation is sufficiently accurate (nine places).

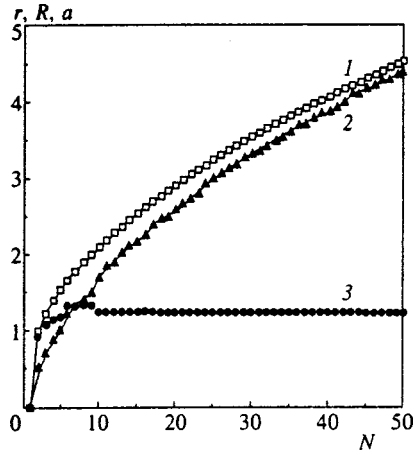

FIG. 1. Mean distance r between particles (*1*), size R of the system (*2*), and lattice period a (*3*) for two-dimensional logarithmic clusters vs the number of particles.

It can be seen from Fig. 1 that the mean distance between the particles monotonically increases as the number of particles increases, while the rate at which it grows decreases somewhat with N. As far as the size of the system is concerned, although the overall tendency is for it to increase, it cannot be said to monotonically depend on N for all N. Comparing Fig. 1 and Table I, it can be concluded that the size R of the system experiences discontinuities when a new shell appears and when a particle is added to the first shell (from the center of the system). However, if the cluster becomes more symmetrical when the number N is increased by one, the size of the system may just barely increase or even slightly decrease, as can be seen from Fig. 1. Moreover, the lattice period (the mean distance between the nearest particles) becomes approximately constant for $N>9$ (see Fig. 1). The last two circumstances show that the density of particles is constant.

We should point out that it is a rather difficult task to search for the configuration corresponding to the global minimum of the potential energy for large N, because there are numerous local minima with a potential energy that differs very little from the energy of the system at the global minimum. For example, in the case of $N=49$, the difference between the potential energy of the system at the global minimum (3,9,15,22) and the local minimum closest to it (3,9,16,21) is only $4\times10^{-4}\%$. For this reason, the search for the minimum has to be done very precisely (no less than seven places).

4.2. The structure of dipole clusters

For dipole clusters, the inner, outer, and total potential energy per particle E/N increases almost linearly as the number of particles increases (see Fig. 2). This indicates that, for a small number of particles $N\leqslant40$, the system has the properties of a cluster and has not yet acquired the properties of a crystal (for which $E/N=\text{const}$).

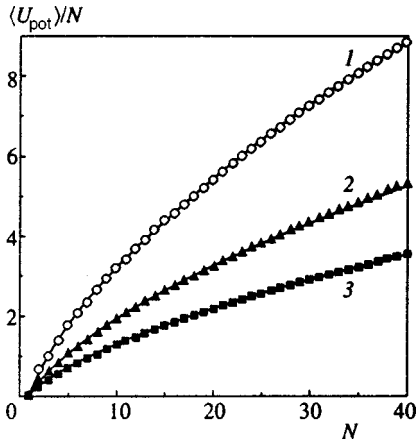

FIG. 2. Potential energy per particle $\langle U_{\rm pot}\rangle/N$ vs the number N of particles for two-dimensional dipole clusters: *1*—total potential energy $\langle U_{\rm pot}\rangle/N$, *2*—mean potential energy of all the interactions between the particles $\langle U_{\rm int}\rangle N=(1/2N)\Sigma\langle U_{ij}\rangle=(1/2N)\Sigma(1/r_{ij}^3)$, *3*—external potential energy $\langle U_{\rm ext}\rangle/N=(1/N)\alpha\Sigma r_i^2$.

Table II shows the occupation number of the shells and the corresponding potential energies for the global minima of two-dimensional dipole clusters. The rules of the filling of the shells are similar to the rules of their filling in logarithmic clusters.

As N increases, it is found that a triangular lattice begins to grow inside the cluster. Fragments of a triangular lattice first appear for a cluster of 12 particles [the (3,9) configuration—see Table II]. Beginning with 32 particles, a triangular lattice continually predominates in the cluster: it is impossible to completely assign certain particles to definite shells—they appear between the shells, forming a fragment of a triangular lattice (the latter is centered close to the boundary of the cluster, and not at the center of symmetry of the confining potential). For example, in a cluster made up of 37 particles, it is no longer possible to unambiguously distinguish the shells in the configuration corresponding to the global minimum of the potential energy (see Table II). A defect exists in the shell structure, but the particles mainly form a triangular lattice. The ordering thus has two forms—either a triangular lattice or a shell structure—which compete with each other. The triangular structure appears at a smaller N for dipole clusters than for Coulomb and logarithmic clusters,[2] and this can be explained by the fact that the dipole interaction potential has a shorter range. Since a regular triangular lattice possesses hexagonal symmetry, deviations of the shape of the shells from circles are observed for $N>30$. This phenomenon is analogous to the ''faceting'' of a crystal (see also Ref. 20). Thus, certain particles in a dipole cluster are found not at the boundary of the cluster but at a distance from it of the order of a lattice period. The introduction of the concept of ''crystal shells'' (Cr_c) is more natural for dipole clusters than for logarithmic, since the majority of the configurations at the global minima can be classified in this way (see Table II).

5. THE MELTING AND PHASE TRANSITIONS OF TWO-DIMENSIONAL MESOSCOPIC CLUSTERS

We have studied the melting of logarithmic, Coulomb, and dipole micro- and macroclusters with increasing temperature.

5.1. The melting of logarithmic clusters

The temperature dependence of the radial mean-square displacement for a two-dimensional logarithmic cluster with $N=37$ is shown in Fig. 3. The radial mean-square displacements for all the shells experience an infection point at the same temperature $T_c=8\times10^{-3}$.

It can be seen from Fig. 4 that the angular mean-square displacements with respect to the nearest particles of its own

TABLE II. Shell structure and potential energy of dipole clusters confined by a harmonic potential.

Number of particles	Occupation number of shells	Potential energy	Number of particles	Occupation number of shells	Potential energy
1	1	0.0000000×10^0	21	2,7,12 Cr_2	1.1740007×10^2
2	2	1.2932046×10^0	22	2,8,12 Cr_2	1.2715322×10^2
3	3	3.0418217×10^0	23	3,8,12 Cr_3	1.3727919×10^2
4	4	5.5208363×10^0	24	3,8,13 Cr_3	1.4753113×10^2
5	5	8.7856477×10^0	25	3,9,13 Cr_3	1.5814029×10^2
6	1,5	1.2289769×10^1	26	4,9,13	1.6921679×10^2
7	1,6 Cr_1	1.6281382×10^1	27	4,9,14 Cr_4	1.8047079×10^2
8	1,7	2.1083395×10^1	28	4,10,14 Cr_4	1.9198318×10^2
9	2,7	2.6313547×10^1	29	5,10,14 Cr_5	2.0404328×10^2
10	3,7 Cr_3	3.1901163×10^1	30	5,10,15 Cr_5	2.1616304×10^2
11	3,8 Cr_3	3.7616955×10^1	31	1,5,10,15 Cr_1	2.2839087×10^2
12	3,9 Cr_3	4.3999784×10^1	32	1,6,12,13 Cr_1	2.4093329×10^2
13	4,9	5.0634105×10^1	33	1,6,12,14 Cr_1	2.5368468×10^2
14	4,10 Cr_4	5.7895957×10^1	34	1,6,12,15 Cr_1	2.6669867×10^2
15	5,10	6.5399893×10^1	35	1,6,12,16 Cr_1	2.8012640×10^2
16	1,5,10	7.3049228×10^1	36	1,6,12,17 Cr_1	2.9407878×10^2
17	1,6,10 Cr_1	8.1136231×10^1	37	1,7,13,16 Cr_1	3.0825097×10^2
18	1,6,11 Cr_1	8.9506331×10^1	38	2,8,13,15 Cr_2	3.2244908×10^2
19	1,6,12 Cr_1	9.8421773×10^1	39	3,8,13,15 Cr_3	3.3690883×10^2
20	1,7,12	1.0776650×10^2	40	3,9,14,14 Cr_3	3.5144690×10^2

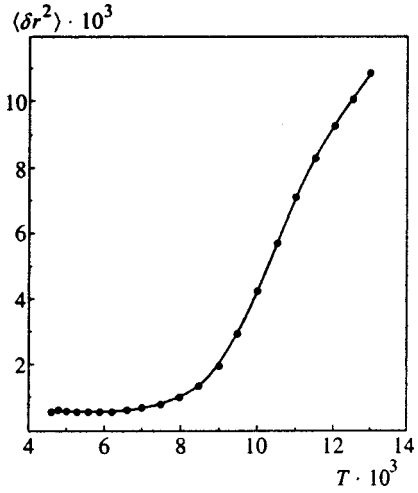

FIG. 3. Total radial mean-square displacement $\langle \delta r^2 \rangle$ vs temperature for a two-dimensional logarithmic cluster with $N=37$ particles.

shell experience an inflection point at the same temperature for both shells. Consequently, a phase transition occurs at a temperature of $T_c = 8\times10^{-3}$ (see Table III) in a cluster made up of $N=37$ particles: The system loses its ordered structure. For $T>T_c$, the number of particles in the shells begins to change, and the shells exchange particles and smear out. For $T \gg T_c$, it is impossible to distinguish any shells. The particles move chaotically. The characteristic value of the dimensionless parameter $\Gamma = q^2/k_B T_c$ at which the system melts is 125 when $N=37$, and this is not much different from the value of Γ at which a system with a relatively large number of vortices melts ($\Gamma=\Gamma_c \simeq 130$).

It turns out, however, that the angular mean-square deviations with respect to the nearest particles of the adjacent shell experience an inflection point at much lower temperatures. For a cluster made up of 37 particles, the angular mean-square displacements of the outer shell with respect to the nearest particles of the middle shell begin to quickly increase at a temperature of $T_{c1} = 8\times10^{-4}$ ($\Gamma = 1/T_{c1} = 1250$) (see Fig. 5a), while the angular displacements of the middle shell relative to the closest particles of the inner shell do this at a temperature of $T_{c2} = 3\times10^{-3}$ ($\Gamma = 1/T_{c2} = 333$) (see Fig. 5b). This means that, at these temperatures, specific orientation melting characteristic of clusters with a shell structure occurs for corresponding pairs of shells; i.e., the shells in the two-dimensional clusters, while maintaining their crystallinity, begin to rotate with respect to each other at a certain instant.

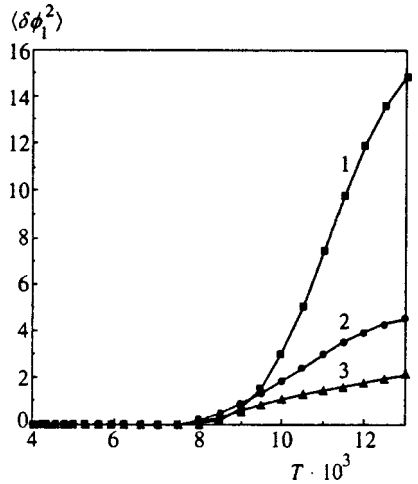

FIG. 4. Angular mean-square deviations $\langle \delta \phi_1^2 \rangle$ of the outer (*1*), middle (*2*), and inner (*3*) shells with respect to the closest particles from their own shell vs temperature for a two-dimensional logarithmic cluster with $N=37$ particles.

TABLE III. Melting temperatures T_c and potential barriers $U_{j,\mathrm{rot}}$ of two-dimensional clusters.

	T_c	$U_{j,\mathrm{rot}}$
Logarithmic cluster, $N=11$		
Orientation melting of outer shell with respect to inner shell	4.0×10^{-7}	2.32×10^{-6}
Complete melting	4.5×10^{-3}	3.71×10^{-2}
Logarithmic cluster, $N=37$		
Orientation melting of outer shell with respect to the middle shell	8.0×10^{-4}	2.30×10^{-3}
Orientation melting of the middle shell with respect to the inner shell	2.4×10^{-3}	1.61×10^{-2}
Complete melting	8.0×10^{-3}	6.61×10^{-2}
Logarithmic cluster, $N=107$		
Orientation melting of the outer shell with respect to the neighboring shell	7.5×10^{-3}	3.0×10^{-2}
Complete melting	8.5×10^{-3}	5.8×10^{-2}
Coulomb cluster, $N=19$		
Orientation melting	–	4.5×10^{-2}
Complete melting	4.0×10^{-3}	5.0×10^{-2}
Coulomb cluster, $N=20$		
Orientation melting of the outner shell with respect to the middle shell	1.8×10^{-6}	1.2×10^{-5}
Complete melting	1.4×10^{-2}	5.4×10^{-2}
Coulomb cluster, $N=37$		
Orientation melting of the outer shell with respect to the middle shell	4.0×10^{-5}	6.1×10^{-3}
Orientation melting of the middle shell with respect to the inner shell	4.0×10^{-5}	6.1×10^{-3}
Complete melting	6.5×10^{-3}	5.17×10^{-2}
Coulomb cluster, $N=54$		
Orientation melting of the outer shell with respect to the neighboring shell	1.5×10^{-3}	4.59×10^{-2}
Orientation melting of the middle shells	1.5×10^{-3}	5.72×10^{-2}
Orientation melting of the inner shell with respect to the adjoining shell	4.0×10^{-3}	1.01×10^{-1}
Complete melting	4.0×10^{-3}	5.31×10^{-2}
Dipole cluster, $N=10$		
Orientation melting of the outer shell with respect to inner shell	1.2×10^{-5}	3.5×10^{-5}
Complete melting	7.0×10^{-3}	5.6×10^{-2}
Dipole cluster, $N=38$		
Orientation melting absent		
Complete melting	9.0×10^{-3}	

We should point out that the phase transitions in the mesoscopic system under consideration, as shown by calculation, are still fairly well-defined (i.e., the broadening $\Delta T/T$ of the transition region is not yet large).

The potential energy of a logarithmic cluster increases

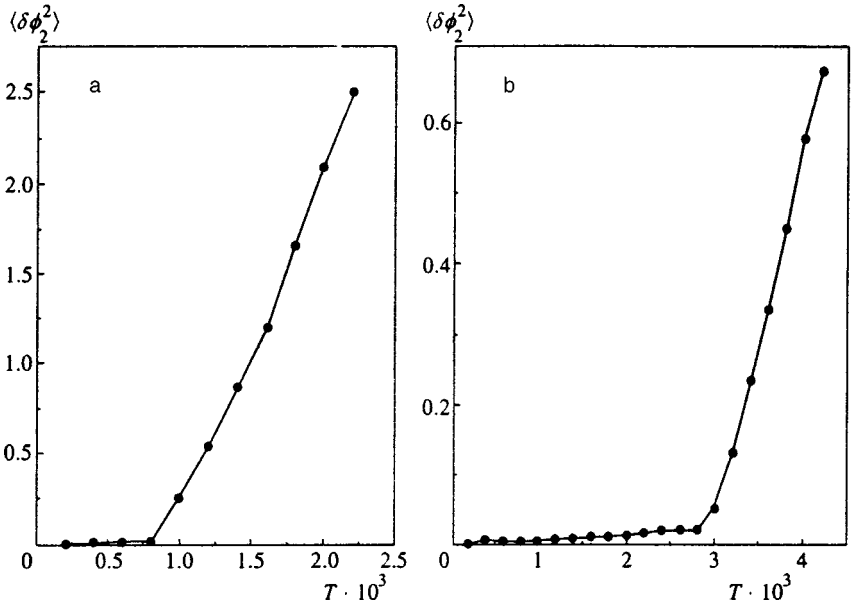

FIG. 5. Temperature dependences of the angular mean-square deviations $\langle\delta\phi_2^2\rangle$ of the middle shell with respect to the nearest particles from the outer shell (a) and of the inner shell with respect to the nearest particles from the middle shell (b) for a two-dimensional logarithmic cluster with $N=37$ particles.

almost linearly with temperature and has no singularities. Therefore, the temperature dependence of the potential energy of the system cannot be used to determine the melting temperature.

For logarithmic macroclusters with $N\simeq 100$, orientation melting occurs only for the outer pairs of shells, since a region of a triangular lattice forms inside.

5.2. The melting of Coulomb clusters

The temperature dependences of the total radial displacement and of the radial mean-square displacements for all the shells individually are shown in Fig. 6 for a two-dimensional Coulomb cluster with $N=54$. These dependences experience an inflection point at the same temperature $T_{c1}=4\times 10^{-3}$ for all the shells.

It can be seen from Fig. 7 that the angular displacements with respect to the nearest particles of its own shell experience an inflection point at the same temperature for all the shells. Consequently, a phase transition occurs at a temperature of $T_{c1}=4\times 10^{-3}$ in a cluster made up of $N=54$ particles: The system loses its ordered structure.

However, it turns out that the angular mean-square displacements with respect to the nearest particles of the adjacent shell experience an inflection point at a lower temperature $T_{c2}=1.5\times 10^{-3}$ for the two outer pairs of shells, and at the same temperature T_{c1} for only the inner pair of shells (see Fig. 7 and Table III). This means that orientation melting occurs at a temperature of $T=T_{c2}$ for the corresponding pairs of shells.

The absence of orientation melting of the inner pair of shells is explained by the fact that there is a fragment of a

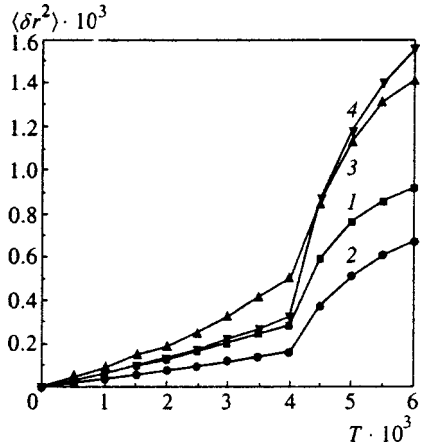

FIG. 6. Temperature dependences of the radial mean-square displacements $\langle\delta r^2\rangle$ for a two-dimensional Coulomb cluster, $N=54$: *1*—total, *2*—of the first shell, *3*—of the second shell, *4*—of the third shell.

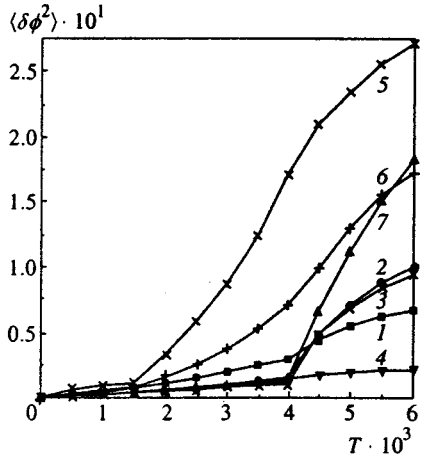

FIG. 7. Temperature dependences of the angular mean-square displacements $\langle\delta\phi^2\rangle$ for a two-dimensional Coulomb cluster with respect to the nearest particles of its own and the neighboring shells, $N=54$: *1*—of the first shell with respect to the particles of its own shell, *2*—of the second shell with respect to the particles of its own shell, *3*—of the third shell with respect to the particles of its own shell, *4*—of the fourth shell with respect to the particles of its own shell, *5*—of the second shell with respect to the particles of the first shell, *6*—of the third shell with respect to the particles of the second shell, *7*—of the fourth shell with respect to the particles of the third shell.

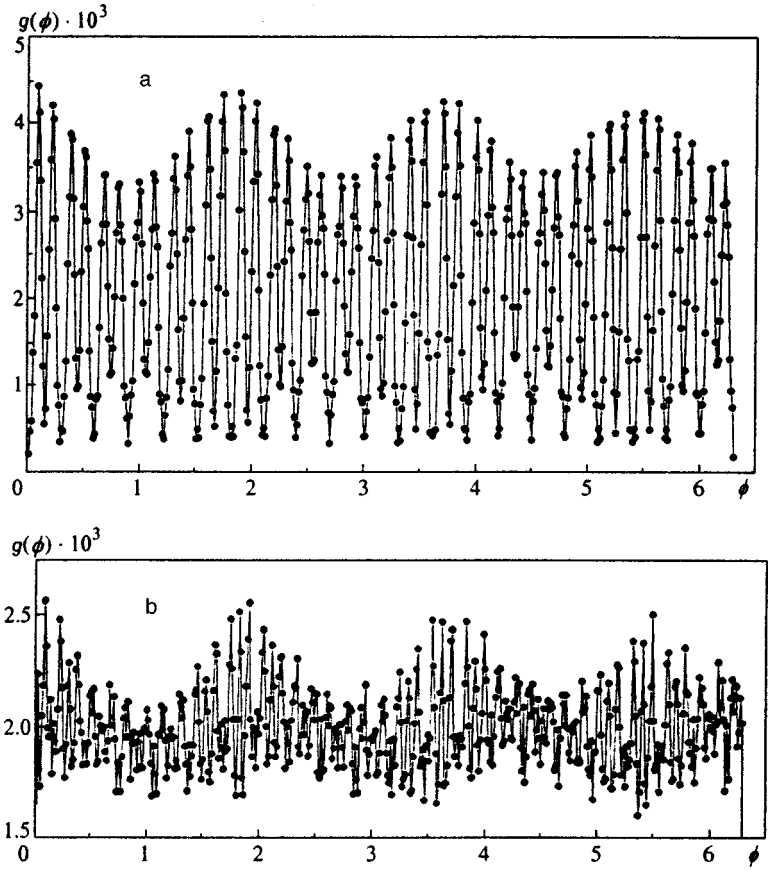

FIG. 8. Distribution functions $g(\phi)$ of the angles between the particles of two shells for $N=20$ at $T=1\times10^{-6}$ (a) and $T=4\times10^{-6}$ (b).

triangular lattice inside a cluster made up of 54 particles. The entire cluster cannot have a triangular lattice, since, because of the centrally symmetric confining potential, clearly expressed circular shells are observed on the outside.

The potential energy of a Coulomb cluster, like that of a logarithmic cluster, increases almost linearly with temperature and has no singularities.

The melting of a Coulomb cluster made up of 37 particles also occurs in two stages: orientation melting and complete melting (see Table III). Unlike a cluster made up of 54 particles, all the pairs of shells in the $N=37$ case (including the inner one) experience orientation melting, which is because there is no region of triangular lattice inside a cluster with $N=37$.

We shall consider the melting of clusters made up of 19 and 20 particles in more detail.

Two-stage melting was observed in a cluster with $N=20$, as expected for Coulomb mesoscopic clusters, with the orientation melting temperature and the complete melting temperature differing by several orders of magnitude. Thus, at $T=T_{c2}=1.8\times10^{-6}$, the angular mean-square displacement with respect to the nearest particles of the adjacent shells experiences an inflection point. Moreover, the fluctuation amplitude of the $g(\phi)$ function (see Section 3) is large at $T<T_{c2}$ (see Fig. 8a), whereas the amplitude of $g(\phi)$ sharply decreases at $T>T_{c2}$ (see Fig. 8b). This all indicates that orientation melting occurs at $T=T_{c2}$. The radial mean-square displacements experience an inflection point at $T=T_{c1}=1.4\times10^{-2}$ (see Table III), which indicates the presence of complete melting of a Coulomb cluster with $N=20$ at the point $T=T_{c1}$.

In a cluster with $N=19$, the radial and angular mean-square displacements with respect to the nearest particles of the adjacent shell experience a discontinuity at the same temperature of $T=T_{c1}=4\times10^{-3}$ (see Table III). For $T>T_{c1}$, two sharp maxima of the outer shell that exist in the $g(r)$ function when $T<T_{c1}$ (see Section 3) run together (Fig. 9); meanwhile, the zeroes that exist in the $g(\phi)$ function disappear at $T<T_{c1}$ (Fig. 10). All this indicates that the orientational and radial ordering in a Coulomb cluster with $N=19$ simultaneously disappear at $T=T_{c1}$. Consequently, a Coulomb microcluster with $N=19$ melts in one stage—there is no orientation melting. This behavior is unique for mesoscopic clusters and results from the fact that the $N=19$ cluster has the configuration (1,6,12) at the global minimum and is almost an exact fragment of a triangular lattice.

We also investigated the distribution of the system over the local minima of the potential energy as the temperature increases. It turned out for both $N=19$ and $N=20$ that, at a temperature below the complete melting temperature, the cluster is continually close to the configuration corresponding to the global minimum of the potential energy. At a temperature above that of complete melting, the probability that a cluster is close to the configuration corresponding to the global minimum (1,7,12) decreases and goes to a constant value. The cluster shuffles between configurations corresponding to different local minima (1,7,12), (1,6,13), and (1,8,11).

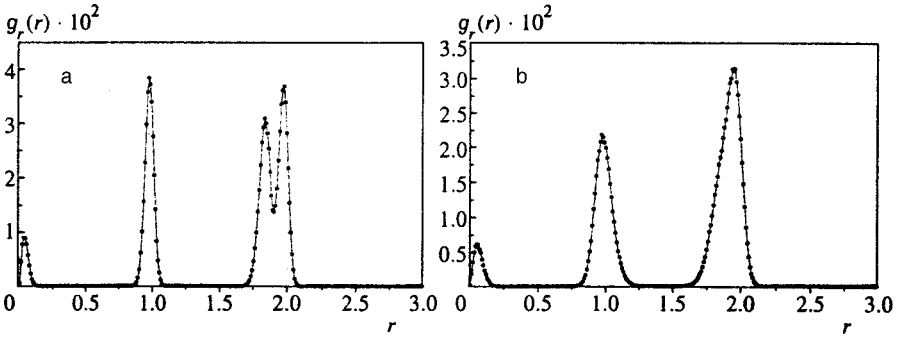

FIG. 9. Radial distribution functions $g(r)$ for $N=19$ at $T=3\times10^{-3}$ (a) and $T=7\times10^{-3}$ (b).

5.3. The melting of dipole clusters

The temperature dependences of the radial mean-square displacements, total and for each shell, experience an inflection point at a temperature of T_{c1} for a two-dimensional dipole cluster with $N=38$ (see Table III). At the same temperature, the angular mean-square displacements with respect to the closest particles of their own and the neighboring shells experience an inflection point for all shells. Consequently, a phase transition occurs in a cluster made up of $N=38$ particles at a temperature of T_{c1}: The system loses its ordered structure. For $T>T_{c1}$, the number of particles in the shells begins to change, and the shells exchange particles and smear out. For $T\gg T_{c1}$, it is no longer possible to distinguish any shells. The particles move chaotically.

The dimensionless parameter Γ_d at which the system melts, in units of $D=k_B=1$, equals

$$\Gamma_d=\frac{1}{2a^3T_c}.$$

For $N=38$, the value $\Gamma_d=69$ does not strongly differ from the value $\Gamma_d=62$ at which a dipole crystal melts.[21]

The scenario of the melting of a two-dimensional dipole cluster made up of $N=38$ particles is analogous to that of the melting of a two-dimensional dipole crystal with a triangular lattice, except that there is no jump in the potential energy at the melting temperature.[21] However, this is not valid for mesoscopic clusters consisting of only two shells. Here, despite the short-range character of the dipole interaction, the quadratic confining potential plays a major role: The shell structure competes with the triangular lattice. The melting of a two-dimensional dipole cluster made up of 10 particles occurs in two stages, as does the melting of logarithmic and Coulomb clusters.

It can be seen from Table III and the preceding discussion that the melting of two-dimensional logarithmic, Coulomb, and dipole mesoscopic clusters (except for magic clusters with a triangular structure) occurs in two stages. The melting of logarithmic and Coulomb macroclusters also occurs in two stages, but orientation melting occurs only in the outer pairs of shells, and, finally, the melting of dipole macroclusters occurs in one stage.

6. POTENTIAL BARRIERS WITH RESPECT TO THE ROTATION OF SHELLS AND THE JUMPING OF PARTICLES FROM ONE SHELL TO ANOTHER

We can use the techniques described above to find the potential barrier $U_{\rm rot}$ against rotation of the shells with respect to each other and the potential barrier U_j against jumping of a particle from one shell to another in the clusters in which melting was studied (see Table III). Rotation of the shells is the lowest excitation in the case of mesoscopic clusters. It turned out that, for mesoscopic clusters, the orientation barrier is substantially lower than the radial one. The fact that the barriers are incommensurate, along with the discontinuity of the angular mean-square displacement, is still more powerful evidence that orientation melting can occur in two-dimensional clusters. However, the barriers against rotation of the inner shells (for logarithmic and Coulomb clusters) or of all the shells (for dipole clusters) are of the same order of magnitude as the barriers against jumping for macroclusters.

It can be seen from Table III that the ratio between the potential barriers against rotation and jumping for a given number N and a given interaction law equals in order of magnitude the ratio between the temperatures of orientation melting and complete melting,

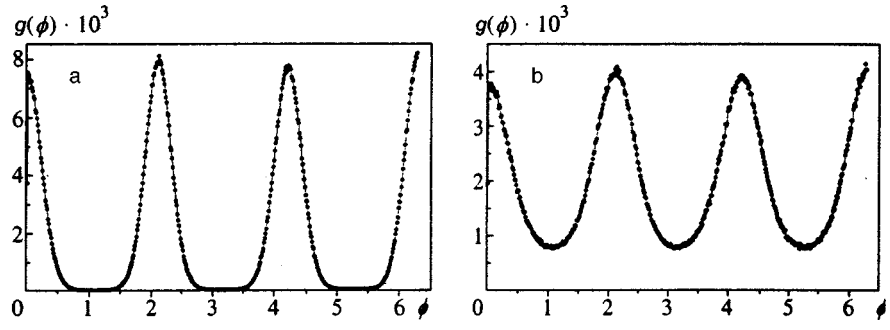

FIG. 10. Distribution functions $g(\phi)$ of angles between the particles of two shells for $N=19$ at $T=3\times10^{-3}$ (a) and $T=7\times10^{-3}$ (b).

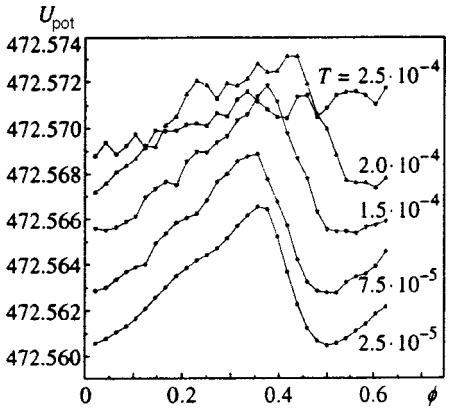

FIG. 11. Temperature variation of the dependence of the potential energy on the angle of rotation of the outer shell with respect to the middle shell for a two-dimensional Coulomb cluster with $N=37$.

$$\frac{U_{\rm rot}}{U_j} \simeq \frac{T_{c2}}{T_{c1}}.$$

This makes it possible to predict the possibility of orientation melting in clusters with different N and different interaction laws.

We also investigated how the potential barriers against rotation and jumping vary as temperature increases. It is easy to see that all the barriers must disappear for a very high temperature $T \gg T_c$, for which it is no longer possible to distinguish the shells in a cluster. However, it is interesting to compare the temperatures at which the barriers with respect to rotation of the shells and with respect to jumping of particles between the shells disappear with the temperatures of orientation melting and complete melting of the clusters.

We have studied how the dependences of the potential energy on the angle of rotation of the shells with respect to each other for two pairs of shells vary with temperature and how the dependence of the potential energy on the radius of the particle, changing shell, varies with temperature for a Coulomb cluster with $N=37$. It turned out that the barrier against rotation of the outer shell with respect to the middle shell disappears at $T=2.5\times10^{-4}$ (Fig. 11), while that of the middle shell with respect to that of the inner shell disappears at $T=6.0\times10^{-4}$. The barrier against the jumping of particles between shells disappears at $T=3.0\times10^{-2}$. When these temperatures are compared with the melting temperatures, it is seen that the melting temperature is less than the temperature at which the corresponding barrier disappears by a factor of about 5–15. This is apparently because the particles begin to partly overcome the barrier at the melting temperature, whereas, when the temperature is increased further and the barrier disappears, the particles cease to notice it and move chaotically in the direction of the corresponding barrier.

7. CLUSTERS WITH ANISOTROPIC CONFINEMENT

Let us consider how the equilibrium structure of two-dimensional clusters varies because of the anisotropy of the confining potential. An anisotropic confining potential has the form

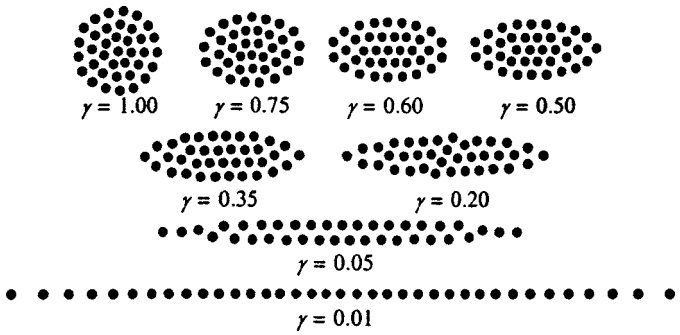

FIG. 12. Configurations of a two-dimensional Coulomb cluster with $N=37$ and various degrees of anisotropy γ.

$$U_{\rm ext} = \gamma \sum_i x_i^2 + (2-\gamma) \sum_i y_i^2, \tag{9}$$

where $0 \leqslant \gamma \leqslant 2$ ($\gamma=1$ corresponds to the isotropic case). We shall calculate the equilibrium configurations by the method described above. The configurations corresponding to the local and global minima of the potential energy were considered for logarithmic and Coulomb clusters with $N=37$ and for logarithmic clusters with $N=7$ and 11 with different degrees of anisotropy γ (the configurations of a Coulomb cluster with $N=37$ are shown in Fig. 12), and the following effects were detected:

1. The greater the degree of anisotropy, the larger the tendency to find the global minimum in the configuration with the smallest number of shells and with the smallest number of particles in the inner shells (see Table IV). The shells become more and more elongated and closer in shape to polygons inscribed in ellipses. We must alter the definition of shell given above. Now the greatest value of $r_i \sqrt{(2-\gamma)\sin^2\phi_i + \gamma\cos^2\phi_i}$ for the particles of each shell must be greater than the smallest value of this quantity for the particles of the outer shell with respect to the given shell. (Because of the anisotropy of the confining potential, the maximum value of r_i for the particles of each shell can be much greater than the minimum value of r_i for the particles of the next outer shell.)

2. If the confinement is strongly anisotropic, the inner shell can have two tails, directed along the y axis (if $\gamma<1$, i.e., $\gamma<2-\gamma$), or can simply turn into a straight line (see Table IV). If a cluster consists of one shell, its structure does not undergo any more changes after it turns into a straight line (other than an increase of the distance between the particles as the degree of anisotropy increases). The cluster becomes one-dimensional. It should be pointed out that the density of a one-dimensional cluster increases toward its center.

3. The greater the degree of anisotropy of the confining potential, the lower the local minima possessed by a cluster with a given number N. In particular, a one-dimensional cluster (a two-dimensional cluster with very strong anisotropy) has only one minimum—the particles lie on a straight line.

The effects described above can be traced in Table IV.

TABLE IV. Structure of two-dimensional anisotropic clusters.

Number of particles	Degree of anisotropy γ	Occupation number of shells $N_1, N_2, \ldots$	Potential energy U_{pot}
		Coulomb clusters	
37	1.00	1,7,12,17	4.725601×10^{2}
37	0.90	1,7,13,16	4.714139×10^{2}
37	0.75	conc.7,13,17	4.651811×10^{2}
37	0.60	L5,14,18	4.528417×10^{2}
37	0.50	L4,14,19	4.403236×10^{2}
37	0.35	15, 22	4.124505×10^{2}
37	0.20	$4+X3+X4$, conc.26	3.652442×10^{2}
37	0.05	$35+2*X1$	2.546257×10^{2}
37	0.01	L37	1.652735×10^{2}
		Logarithmic clusters	
7	1.00	1, 6	-7.512442×10^{-1}
7	0.89	1, 6	-8.891301×10^{-1}
7	0.67	1, 6	-2.136851×10^{0}
7	0.50	1, 6	-4.172614×10^{0}
7	0.40	1, 6	-6.029061×10^{0}
7	0.33	l7	-7.802550×10^{0}
7	0.18	l7	-1.416698×10^{1}
11	1.00	3, 8	-1.709333×10^{1}
11	0.89	2, 9	-1.251753×10^{1}
11	0.67	2, 9	-1.566826×10^{1}
11	0.50	11	-2.082752×10^{1}
11	0.40	11	-2.560793×10^{1}
11	0.33	11	-2.983309×10^{1}
11	0.18	L11	-4.548732×10^{1}
37	1.00	1, 6,12,18	-5.053534×10^{2}
37	0.89	6,12,19	-5.095770×10^{2}
37	0.67	L4,13,20	-5.456363×10^{2}
37	0.50	15, 22	-6.037877×10^{2}
37	0.40	$9+2*X2$, 24	-6.580007×10^{2}
37	0.33	L9, 28	-7.064780×10^{2}
37	0.18	37	-8.840317×10^{2}
37	0.10	$31+2*X3$	-1.073160×10^{3}
37	0.06	$17+2*X10$	-1.237329×10^{3}
37	0.03	L37	-1.468088×10^{3}

Note: The following symbols are used here: Lz is a line consisting of z particles; $x*Xy$ are x tails, each of which consists of y particles; and ''conc.'' means that this shell is concave.

We next consider how the degree of anisotropy of the confining potential affects the ratio of the potential barriers with respect to rotation of the shells and jumping of the particles between shells. It can be assumed that, in accordance with the distortion of the circular symmetry of the confining potential, the orientation melting disappears, and the melting becomes single-stage. To prove this hypothesis, we investigated the barriers against rotation U_2 and against jumping U_1, using logarithmic and Coulomb clusters consisting of 37 particles as examples (see Table V). As the degree of anisotropy γ increases, the ratio U_{rot}/U_j increases [an exception is the case of a logarithmic cluster with $\gamma=0.89$, for which this ratio decreases; this is because the global minimum is displaced from configuration (1,6,12,18), similar to a fragment of triangular lattice, into configuration (6,12,19), which possesses a clearly expressed shell structure and does not resemble a fragment of triangular lattice] and for certain γ ($\gamma=0.50$ for a logarithmic cluster and $\gamma=0.60$ for a Coulomb cluster) becomes larger than unity for all pairs of shells (see Table V). This agrees with our assumptions.

TABLE V. Potential barriers of clusters with $N=37$ with various degrees of anisotropy of the confining potential.

γ	$U_{\text{rot},a}$	$U_{\text{rot},b}$	U_j
		Logarithmic cluster	
1.00	2.31×10^{-3}	1.60×10^{-2}	6.60×10^{-2}
0.89	1.14×10^{-4}	5.84×10^{-2}	5.01×10^{-2}
0.67	1.17×10^{-2}	6.67×10^{-2}	3.04×10^{-2}
0.50	9.47×10^{-2}	—	4.64×10^{-2}
		Coulomb cluster	
1.00	6.1×10^{-4}	6.1×10^{-3}	5.17×10^{-2}
0.90	1.96×10^{-2}	5.73×10^{-2}	5.12×10^{-2}
0.75	1.13×10^{-1}	1.77×10^{-2}	6.93×10^{-2}
0.60	8.95×10^{-2}	1.21×10^{-1}	6.65×10^{-2}

Note: Here γ is the anisotropy parameter, $U_{\text{rot},a}$ is the barrier against rotation of the outer shell with respect to the middle shell, $U_{\text{rot},b}$ is the barrier against rotation of the middle shell with respect to the inner shell, and U_j is the barrier against jumping of a particle between shells.

On the other hand, in order to confirm that the melting of a two-dimensional cluster has a single stage for strong anisotropy of the confining potential, we calculated the temperature dependences of the radial and relative angular displacements for a logarithmic cluster with $N=37$ and $\gamma=0.50$ and for a Coulomb cluster with $N=37$ and $\gamma=0.60$. We found that the dependences of the radial and relative angular displacements with respect to the nearest particles from their own and from neighboring shells experience a discontinuity at the same temperatures ($T_c=0.0045$ for a logarithmic cluster and $T_c=0.0065$ for a Coulomb cluster). Consequently, a two-dimensional cluster with a sufficiently strong degree of anisotropy melts in one stage.

It should also be pointed out that, for a very strong degree of the confinement anisotropy, the melting of a cluster in general does not occur, since the cluster becomes a one-dimensional system.

8. THE MELTING OF THREE-DIMENSIONAL COULOMB CLUSTERS

In this section, we study three-dimensional Coulomb clusters in a quadratic confining potential. Since the confining potential is centrosymmetric, the clusters have a shell structure (as do two-dimensional clusters) at low temperatures (see, for example, Refs. 3 and 22). However, in the three-dimensional case, at low values of N ($N<100$), the shells in the clusters have the shape of polyhedra close to those inscribed in a sphere. Moreover, the larger the dimensionality, the larger the critical number of particles for which a new shell is formed (see Table VI).

There are a number of papers devoted to the structure of three-dimensional Coulomb clusters (see, for example, Refs. 3, 22, and 23). However, the melting of three-dimensional

TABLE VI. Formation of new shells in Coulomb clusters.

Formation of shells	$D=2$	$D=3$
Formation of a second shell	$N_{cr}=6$	$N_{cr}=13$
Formation of a third shell	$N_{cr}=17$	$N_{cr}=61$

TABLE VII. Potential barriers in three-dimensional Coulomb clusters.

Barrier	$N=33$	$N=48$	$N=60$	$N=100$
Barrier with respect to rotation of the shells	1.44×10^{-2}	3.00×10^{-2}	7.5×10^{-3}	3.13×10^{-2}
Barrier against jumping of particles between shells	6.42×10^{-2}	6.74×10^{-2}	1.18×10^{-1}	5.71×10^{-2}

Coulomb clusters has not been studied in detail (preliminary data were obtained in a paper by one of the authors[3]). Therefore, the interesting question arises of whether orientation melting (which is usual for a two-dimensional cluster) is possible in three-dimensional clusters.

We first found the configurations of the clusters at the local and global minima of the potential energy for $N=33,48,60,100$. We next found the potential barriers with respect to rotation of the shells and jumping of the particles between shells for the given N. It turned out that the barrier against rotation is not much different from the barrier against jumping only for $N=100$, whereas, for clusters consisting of two shells with $N=33,48,60$, the barrier against rotation is substantially less than the barrier against jumping (see Table VII). Therefore, orientation melting is possible in three-dimensional mesoscopic clusters.

To confirm this fact, we calculated the melting of three-dimensional Coulomb clusters with $N=33,100$. We studied the temperature dependences of the following quantities, which are analogs of the quantities presented above for two-dimensional systems:

(1) The total potential energy.

(2) The total radial mean-square displacement, and the radial mean-square displacement for each shell.

(3) The angular mean-square displacement relative to the closest particles from a given shell:

$$\langle \delta\Omega^2_{\rm int}\rangle = \frac{1}{N_R}\sum_{i=1}^{N_R} \frac{\langle(\phi_i-\phi_{i_1})^2+(\theta_i-\theta_{i_1})^2\rangle-\langle(\phi_i-\phi_{i_1}+\theta_i-\theta_{i_1})\rangle^2}{\Omega_0^2}. \tag{10}$$

(4) The angular mean-square displacement with respect to the nearest particles of the neighboring shell:

$$\langle \delta\Omega^2_{\rm ext}\rangle = \frac{1}{N_R}\sum_{i=1}^{N_R} \frac{\langle(\phi_i-\phi_{i_2})^2\rangle+\langle(\theta_i-\theta_{i_2})^2\rangle-\langle(\phi_i-\phi_{i_2}+\theta_i-\theta_{i_2})\rangle^2}{\Omega_0^2}. \tag{11}$$

Here $2\Omega_0=4\pi/N_R$ is the mean angular distance between the nearest particles of a given shell.

For $N=33$, the temperature dependences of the radial mean-square displacement for both shells simultaneously experience a jump at $T_{c1}=7\times10^{-3}$. As the temperature increases further, the shells in a cluster flow together, and the cluster loses its ordered structure. Consequently, $T=T_{c1}$ is the temperature of complete melting for a three-dimensional Coulomb mesoscopic cluster with $N=33$.

However, the temperature dependence of the angular mean-square displacement with respect to the nearest particles of the neighboring shell experiences an inflection point at the lower temperature of $T_{c2}=4\times10^{-4}$. The dependences of the radial and angular mean-square displacements with respect to the nearest particles of a given shell also experience inflection points at this temperature for both shells. Thus, orientation melting occurs at $T=T_{c2}$ in a three-dimensional cluster with $N=33$; i.e., the ''bulk'' shells, maintaining their crystallinity, begin to rotate with respect to each other at $T=T_{c2}$.

However, the melting of a three-dimensional Coulomb macrocluster with $N=100$ occurs in one stage. The dependences of the radial mean-square displacements experience an inflection point at $T_c=9.0\times10^{-3}$ for all the shells, while the angular mean-square displacements with respect to the closest particles of their own and the neighboring shells have no singularities as a function of temperature. These results confirm the results of the calculation of the potential barriers.

9. CONCLUSIONS

1. It has been shown that two-dimensional mesoscopic clusters of particles that repel according to dipole, Coulomb, and logarithmic laws and that are confined by an external quadratic potential have a shell structure at low temperatures. The configurations of the system at the local and global minima of the potential energy have been found. As the number of particles in the clusters increases, a fragment of a triangular lattice appears. The longer-range is the interaction potential between the particles, the greater the number of particles needed to form the triangular lattice. The following physical implementations have been considered: electrons on the surface of liquid helium, electrons in a quantum dot, particles in a layer of magnetic liquid, and vortices in a small superconducting island or in a rotating vessel with superfluid helium, as well as electrons in a semiconductor nanostructure surrounded by a medium with low permittivity, etc.

2. The temperature dependences of the potential energy, the mean-square radial and angular displacements, the radial and angular distribution functions of the particles, and the

distribution of the particles over the local minima of the potential energy have been studied in detail. As a result, the melting of the system has been studied. It has been shown that the melting occurs in one stage in dipole macroclusters. In macroclusters with a longer-range interaction potential and in all mesoscopic clusters (except for magic microclusters with a triangular lattice structure), the melting comes in two stages: The first stage, at lower temperatures, is orientation melting—from the frozen phase into a state with rotational reorientation of ''crystalline'' shells with respect to each other. The second stage is a transition in which the radial ordering disappears. In Coulomb and logarithmic macroclusters, orientation melting occurs only for the outer pairs of shells. We should point out that orientation melting is possible in suitable crystals with impurities. There the impurity plays the role of the confining potential, and the entire crystal in a certain sense is broken up into clusters, so that, as the temperature increases, orientation melting occurs around the impurities.

3. The cause of the orientation melting is that the potential barrier with respect to rotation of the shells is incommensurate (small) by comparison with the barrier with respect to jumping of particles from one shell to another (for all the pairs of shells in mesoscopic clusters and for the outer pairs of shells in macroclusters with Coulomb and logarithmic interaction laws). When the temperatures at which the potential barriers disappear are compared with the corresponding melting temperatures, it can be seen that the melting temperature is lower than the temperature at which the corresponding barrier disappears by about a factor of 5–15. A method is proposed for predicting the character of the melting in shell clusters by comparing the potential barriers against rotation of the shells and against jumping of particles between the shells. The melting temperatures can be estimated from the temperatures at which the potential barriers disappear.

4. We have studied how the structure of the clusters varies for various degrees of anisotropy of the confining potential. For a small degree of anisotropy, the global minima shift to configurations with a smaller number of particles in the inner shells and into configurations with a smaller number of shells. For a strong degree of anisotropy of the confining potential, the inner shell can have two tails directed along the axis or can even degenerate into a straight line. If the cluster consists of one shell, the cluster becomes one-dimensional when it degenerates into a straight line. As the degree of anisotropy increases, the barriers against rotation increase with respect to the barrier against jumping; at a certain degree of anisotropy, the orientation melting disappears and the melting becomes single-stage.

5. The higher the dimensionality of the cluster, the larger the critical number of particles for which a new shell forms. Orientation melting also takes place in three-dimensional Coulomb mesoscopic clusters. However, for large N values, orientation melting disappears and the melting of three-dimensional macroclusters is single-stage.

This work was supported by INTAS, the Russian Fund for Fundamental Research, and the Atomic Surface Structures Program.

*)E-mail: lozovik@isan.msk.su

[1] Yu. E. Lozovik, Usp. Fiz. Nauk **153**, 356 (1987) [Sov. Phys. Usp. **30**, 912 (1987)].
[2] Yu. E. Lozovik and E. A. Rakoch, JETP Lett. **65**, 282 (1997); Yu. E. Lozovik and E. A. Rakoch, Phys. Rev. B **57**, 1214 (1998).
[3] Yu. E. Lozovik and V. A. Mandelshtam, Phys. Lett. A **145**, 269 (1990); F. M. Peeters, V. A. Schweigert, and V. M. Bedanov, Physica B **212**, 237 (1995); Yu. E. Lozovik and E. A. Rakoch, Phys. Lett. A **235**, 55 (1997).
[4] M. E. Gracheva, V. A. Kashurnikov, and I. A. Rudnev, JETP Lett. **66**, 291 (1997).
[5] A. A. Abrikosov, Zh. Eksp. Teor. Fiz. **32**, 1442 (1957) [Sov. Phys. JETP **5**, 1174 (1957)].
[6] G. Blatter, M. V. Feigel'man, V. B. Geshkenbein *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
[7] L. V. Keldysh, JETP Lett. **29**, 658 (1979).
[8] J. J. Thomson, Phil. Mag. S **6, 7**, 238 (1904).
[9] P. Leiderer, W. Ebner, and V. B. Shikin, Surf. Sci. **113**, 405 (1987).
[10] M. A. Read and W. P. Kirk, *Nanostructure Physics and Fabrication* (Academic Press, Boston, 1989).
[11] Yu. E. Lozovik and V. I. Yudson, JETP Lett. **22**, 274 (1975); S. M. Apenko, A. V. Kluchnik, and Yu. E. Lozovik, Solid State Commun. **36**, 485 (1980); B. A. Abdulaev and Yu. E. Lozovik, Fiz. Tverd. Tela **24**, 2663 (1982) [Sov. Phys. Solid State **24**, 1510 (1982)].
[12] Yu. E. Lozovik and V. I. Yudson, Zh. Éksp. Teor. Fiz. **71**, 738 (1976) [Sov. Phys. JETP **44**, 389 (1976)]; Yu. E. Lozovik and O. L. Berman, JETP Lett. **64**, 573 (1996); Yu. E. Lozovik and O. L. Berman, Phys. Scr. **58**, 86 (1998).
[13] D. Yoshioka and H. Fukuyama, J. Phys. Soc. Jpn. **45**, 137 (1978).
[14] L. Brey, Phys. Rev. Lett. **65**, 903 (1990).
[15] D. Yoshioka and A. H. MacDonald, J. Phys. Soc. Jpn. **59**, 4211 (1990).
[16] X. M. Chen and J. J. Quinn, Phys. Rev. Lett. **67**, 895 (1991).
[17] A. T. Skjeltorp, Phys. Rev. Lett. **51**, 2306 (1983).
[18] Yu. E. Lozovik and V. A. Mandelshtam, Phys. Lett. A **165**, 469 (1992).
[19] L. J. Campbell and R. M. Ziff, Phys. Rev. B **20**, 1886 (1979).
[20] G. E. Volovik and U. Parts, JETP Lett. **58**, 774 (1993).
[21] V. M. Bedanov, G. V. Gadiyak, and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **88**, 1622 (1985) [Sov. Phys. JETP **61**, 967 (1985)].
[22] R. M. Hasse and V. V. Avilov, Phys. Rev. A **44**, 4506 (1991).
[23] K. Tsuruta and S. Ichimaru, Phys. Rev. A **48**, 1339 (1993).

Translated by W. J. Manthey

Vacancy mechanism of the anomalous behavior of surface atoms at elevated temperatures

Yu. N. Devyatko,*) S. V. Rogozhkin, and V. I. Troyan

Moscow Engineering-Physics Institute, 115409 Moscow, Russia

E. P. Gusev

IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598 U.S.A.

T. Gustafsson

Department of Physics and Chemistry, Rutgers University, Piscataway, New Jersey 08855 U.S.A.
(Submitted 10 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 2038–2044 (December 1999)

A model that takes into account the significant contribution of vacancies at moderate and high temperatures is proposed to explain the anomalous increase in the mean-square displacements of atoms in transition-metal surface monolayers. The dependence of the effects on the orientation of the surface is explained. The vacancy formation energies and the values of the Debye temperature are obtained within the proposed model for differently oriented surfaces.
 [S1063-7761(99)01212-3]

1. Numerous experiments performed by various methods [medium-energy ion scattering (MEIS), low-energy electron diffraction (LEED), elastic scattering of He^+ ions, electron-energy-loss spectroscopy (EELS), angle-resolved photoemission spectroscopy (ARPES), and Rutherford backscattering (RBS)] have revealed significant anomalies in the behavior of atoms in surface monolayers of transition-metal single crystals at moderate and high temperatures:

the vibrational amplitudes of the surface atoms significantly exceed the bulk values;[1–7]

the amplitudes of the in-plane vibrations of atoms are of the same order as the vibrational amplitudes in the perpendicular direction and, in some cases, appreciably exceed them;[3,8,9]

considerable anharmonicity has been discovered in the vibrations of surface atoms;[10]

an anomalous increase in the thermal-expansion coefficient has been noted;[11,12]

softening of the phonon spectrum has been observed.[13,14]

All these effects are manifested differently for different orientations of the single-crystal cut. For fcc crystals the sharpest dependences are observed for faces having the (110) orientation, and the weakest dependences are observed for (111) faces.[1,4,10,12] In particular, anomalous thermal expansion of a Ni(100) surface was observed using LEED at 900–1300 K.[10] However, an MEIS investigation of a Ni(111) surface did not reveal any increase in the expansion of the surface layers up to 1100 K.[4] Unlike nickel surfaces, Ag(111) (Ref. 2) and Cu(111) (Ref. 6) surfaces expand considerably [Ag(111) by up to 10% and Cu(111) by up to 4.3%] at high temperatures.

The atomic potential on a surface differs from the bulk value because of 3D symmetry breaking, which leads to the appearance of anharmonicity of the vibrations of atoms in the direction perpendicular to the atomic plane. However, attempts to explain the observed phenomena only on the basis of this anharmonicity were unsuccessful (see, for example, Ref. 15). Moreover, this anharmonicity offers no explanation for the anomalous in-plane vibrations.

The theoretical approaches proposed to account for the temperature effects just enumerated[16–22] utilize molecular dynamics. The interaction between atoms is described by potentials obtained within the method of pairwise interatomic interactions,[23] the pseudopotential method,[24] and the embedded-atom method (EAM).[25] Yang *et al.*[16] were able to attribute the anomalous amplitudes of atomic vibrations on Ag(110) and Cu(100) surfaces to surface phonon modes. For Ni(100) the anharmonicity of the vibrations was explained in a quasiharmonic approximation using an EAM potential in Ref. 17.

In the theoretical models described in Refs. 16–20 the surface is treated as a perfect (defect-free) structure determined by the volume. However, studies of the Ni(110) (Ref. 10) and Pb(100) (Ref. 26) surfaces (using high-resolution LEED) and the Pt(111) surface[27] (using reflection electron microscopy) revealed an anomalous decrease in the diffraction intensity when the temperature was raised. An increase in the intensity of the fundamental background was simultaneously observed and was attributed to the formation of surface vacancies. Adatom-vacancy pairs were introduced to account for the mean-square displacements on Ag(110) and Cu(110) surfaces in the models in Refs. 21 and 22. The formation of diadatom-divacancy pairs was discussed in Ref. 28 for Ni(110). Anomalously large (5–30%) enrichment of the surface with vacancies was obtained in high-resolution LEED experiments for Ni(110) (Ref. 10), Pb(100) (Ref. 26), and Pt(111) (Ref. 27) surfaces. The authors of these reports note that vacancy formation plays a dominant role in the

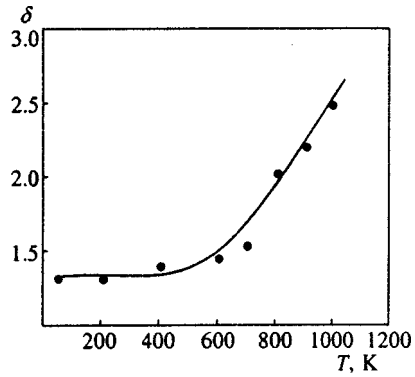

FIG. 1. Temperature dependence of $\delta=\sqrt{\langle\Delta u^2\rangle_s/\langle\Delta u^2\rangle_b}$. The solid line was obtained by the least-squares method from Eq. (4) using the experimental data (points) for Cu(110) (Ref. 1).

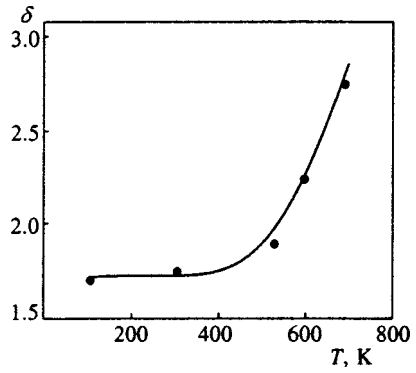

FIG. 2. Same as in Fig. 1 for Cu(100) (Ref. 3).

roughening and disordering of surfaces at moderate temperatures.

The EAM calculations of the formation energy ε_f of surface vacancies in Ref. 29 showed that this parameter is comparatively small and falls in the range $\varepsilon_f=0.2-0.8$ eV for the group of metals considered (Ag, Cu, and Ni). Estimates show that at elevated temperatures the concentration of surface vacancies corresponding to the formation energies indicated reach a value of $\sim 1.0\%$. Since the melting of metals begins already when the bulk concentration of vacancies is $\sim 0.01\%$, the high concentration of surface vacancies indicated must have a significant effect on the surface phenomena at elevated temperatures.

The purpose of the present work was to investigate the role and contribution of vacancies to the mean-square displacements of atoms in transition-metal surface monolayers at elevated temperatures.

2. The microscopic calculation of the mean-square displacements is carried out within an expansion of the potential energy Φ in the deviations of the atoms $\mathbf{u}(\mathbf{R}_k)$ from their equilibrium positions $\mathbf{R}_k$:

$$\Phi=\Phi_0+\frac{1}{2}\sum_{\mathbf{R}_k,\mathbf{R}_i} u_\alpha(\mathbf{R}_k)D_{\alpha\beta}(\mathbf{R}_k-\mathbf{R}_i)u_\beta(\mathbf{R}_i)+\ \ldots\ , \tag{1}$$

where $D_{\alpha\beta}$ is the dynamic matrix (see, for example, Ref. 23).

For surface atoms the $\mathbf{R}_k$ are the positions of the atoms under equilibrium conditions (in a defect-free crystal) at low temperatures (with consideration of the low-temperature relaxation of the surface).

The presence of vacancies leads to a change in the potential energy of the system $\tilde{\Phi}\to\Phi+\delta V$ due to both the disappearance of atoms and the associated reorganization of the electronic structure. The parameters of the dynamic matrix $D_{\alpha\beta}$ vary as a consequence. A decrease in density clearly leads to a decrease in the elastic constants. We note that the perturbation δV is not small for crystal-lattice sites next to vacancies. Vacancies make a fundamentally new contribution to the potential energy Φ of the system[30] by breaking the translational symmetry of the crystal. In the absence of defects, the forces acting on the atom in the site at $\mathbf{R}_k$ are equal to zero. A different situation is created in the presence of defects. The force exerted by other atoms on the atom in the site at $\mathbf{R}_k$ is

$$F_\alpha=-\frac{\varepsilon\tilde{\Phi}}{\varepsilon(\mathbf{R}_k)_\alpha}\to\begin{cases}=0, & c_v=0,\\ \neq 0, & c_v\neq 0.\end{cases} \tag{2}$$

Here c_v is the concentration of vacancies. As a result, a linear term appears in the expansion (1) of the potential energy in a series in small displacements, and the effect of the force **F** (2) leads to displacement of the equilibrium positions of the atoms.

Thus, the appearance of vacancies leads to a change in the mean-square displacements $\langle\Delta u^2\rangle_s$ of the atoms in the outer monolayer: static displacement of the atoms and alteration of their vibrational amplitudes occur.

The experiments were devised to study the temperature dependence of the magnitude of the relative displacement $\delta=\sqrt{\langle\Delta u^2\rangle_s/\langle\Delta u^2\rangle_b}$, where $\langle\Delta u^2\rangle_b$ is the mean-square displacement of the atoms in the bulk of the sample. In order to describe the experimental dependences of the behavior of the mean-square displacements of atoms, let us consider a simple model. We assume that all the changes occur only in the atoms in the local environment of a vacancy (z is the number of such atoms) and that the appearance of vacancies does not have any effect on the behavior of other atoms. The vacancies do not interact with one another, and the contribution of the vacancies is therefore proportional to their concentration.

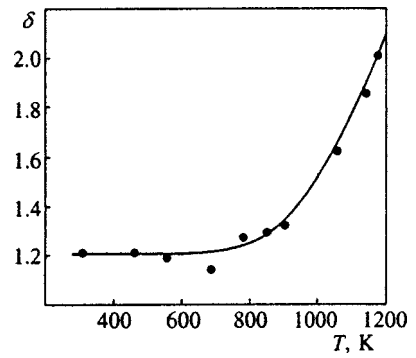

FIG. 3. Same as in Fig. 1 for Cu(111) (Ref. 6).

TABLE I. Vacancy formation energy for differently oriented surfaces of Cu, Ag, and Ni single crystals: values obtained from experimental data[1–7] by the least-squares method using Eq. (4) *, results of EAM calculations,[29] and results of charge-density calculations.[34] The bulk values of the vacancy formation energy ε_b (Refs. 35 and 36) are given for comparison.

	ε_v			
Metal	(110)	(100)	(111)	ε_b
Ag	0.16*		0.43*	1.11
	0.22 (Ref. 29)	0.45 (Ref. 29)	0.58 (Ref. 29), 0.67 (Ref. 34)	
Cu	0.35*	0.35–0.5*	0.7*	1.28
	0.29 (Ref. 29)	0.58 (Ref. 29)	0.72 (Ref. 29), 0.92 (Ref. 34)	
Ni			0.5–0.8*	1.8
	0.34 (Ref. 29)	0.66 (Ref. 29)	0.82 (Ref. 29)	

Averaging over an ensemble of atoms on a surface, for the mean-square displacement we obtain

$$\langle \Delta u^2 \rangle_s = \langle \Delta u^2 \rangle_0 + (\langle \Delta u^2 \rangle_v - \langle \Delta u^2 \rangle_0) z c_v , \tag{3}$$

where $\langle \Delta u^2 \rangle_v$ and $\langle \Delta u^2 \rangle_0$ are the mean-square displacements of the atoms near and far away from a vacancy.

Using the Arrhenius dependence of the vacancy concentration on the temperature T, we obtain the following expression for $\delta(T)$ (the Boltzmann constant $k_B=1$):

$$\delta = \sqrt{\xi + \frac{\psi}{T} \exp\left(- \frac{\varepsilon_f}{T} \right)}, \tag{4}$$

where ε_f is the effective vacancy formation energy, which decreases as the surface is approached,[31] ξ is the ratio of the mean-square displacements of the atoms on the surface to the corresponding bulk value at low temperatures ($T \ll \varepsilon_f$), and the parameter ψ is $\psi = z \langle u^2 \rangle_v K_b e^{\Delta S}$ (ΔS is the entropy change caused by the alteration of the properties of the atoms as a consequence of the appearance of a defect, particularly by the change in the vibrational frequency, and K_b is the bulk value of the elastic modulus).

The experimental data for differently oriented Ag, Cu, and Ni surfaces were used to determine the parameters of the temperature dependence of the relative displacement δ (4) by the least-squares method. As examples, Figs. 1–3 present the theoretical plots of (4) obtained for Cu(111), Cu(100), and Cu(110). The experimental data[1,3,6] are shown as points in Figs. 1–3. The vacancy formation energies on different surfaces are listed in Table I. It is seen that for the same material the vacancy formation energy increases successively from the (110) surface to the (100) surface and then to the (111) surface. Consequently, at an assigned temperature the concentration of vacancies on a surface decreases from surface to surface in the following order: (110)→(100)→(111). This sequence coincides with the experimentally observed dependence of the magnitude of the mean-square displacements of the atoms on the orientation of the surface. The empirical values of the vacancy formation energies are close to the values calculated by the embedded-atom method[29] (see Table I).

Apart from the formation energy ε_f, the dependence (4) contains two more parameters, viz., ξ and ψ. Let us relate them to physical quantities. Without allowing for vacancies, the magnitude of the relative deviation δ equals $\delta=\sqrt{\xi}$. In the Debye model (for temperatures above the Debye temperature) δ is the ratio between the Debye temperatures within the sample and on the surface: $\delta=\Theta_b/\Theta_s$. In the harmonic approximation δ is determined by the ratio of the force constants in the volume to the corresponding constants on the surface $\delta=\sqrt{K_b/K_s}$. The values of the Debye temperatures Θ_s that we obtained for differently oriented surfaces from the experimental data are presented in Table II. The value of Θ_s has previously been discussed in the literature (see, for example, Ref. 1). The corresponding data[1,32] on Θ_s are also presented in Table II. A more detailed comparison of the Debye temperatures Θ_s requires allowance for the anisotropy of surfaces with low Miller indices and the scattering geometry in the experiments considered.

The values of ψ were also obtained from experimental data by the least-squares method. The values obtained from ψ for the mean-squares displacements of atoms near vacancies $\langle \Delta u^2 \rangle_v$ include both local strains and vibrations. At elevated temperatures the values of $\sqrt{\langle \Delta u^2 \rangle_v}$ are comparable to the lattice constant. Such a large value for $\langle \Delta u^2 \rangle_v$ calls for

TABLE II. Debye temperatures for differently oriented surfaces of Cu, Ag, and Ni single crystals: values obtained from experimental data[1–7] by the least-squares method using Eq. (4)* and results of EAM calculations.[1] The bulk values of the Debye temperature Θ_b (Ref. 23) are given for comparison.

	Θ_s			
Metal	(110)	(100)	(111)	Θ_b
Ag	205*		151*	215
Cu	220*, 250 (Ref. 32)	185*	260*	315
	150–240 (Ref. 1),	125–190 (Ref. 1)	150–190 (Ref. 1)	
Theoret.	224 (Ref. 1)	246 (Ref. 1)	213 (Ref. 1)	340 (Ref. 1)
Ni			282*	375

a more detailed theoretical treatment extending beyond the scope of the model proposed here, which would take into account the change in the behavior only of atoms next to a vacancy. We note that such a situation corresponds to melting according to the Lindemann criterion. In our case we may be dealing with local melting.

It was postulated above that the parameters ξ and ψ in (4) do not depend on temperature. A possible complication of the temperature dependence (4) is associated with allowance for the temperature dependence of the mean-square displacements of atoms near vacancies $\langle \Delta u^2 \rangle_v$ and the force constants $K_s(T)$. Calculations showed that the temperature dependence of $\langle \Delta u^2 \rangle_v$ only slightly reduces the vacancy formation energy.

3. Thus, consideration of the imperfection of the surface, i.e., the presence of vacancies, enables us to explain the increase in the mean-square displacements of atoms obtained in the behavior of the last layers of transition-metal single crystals at moderate and high temperatures. A vacancy creates a field for the lattice atoms, which displaces the atoms in its local environment and simultaneously alters their vibrations. In particular, the amplitudes (of both in-plane vibrations and vibrations in the perpendicular direction) and the anharmonicity of the atomic vibrations increase sharply. The vibrational frequencies of the atoms in the local environment also vary.

We note that, apart from the factors taken into account, the displacement caused by the anharmonicity appearing as a consequence of the formation of vacancies, whose diffusion leads to surface melting at high temperatures, makes a contribution to the mean-square displacement of the atoms. At these temperatures the values of the concentration exceed 1%, and the vacancies cannot be treated as isolated: the interaction between the vacancies becomes appreciable, its influence increases with increasing concentration. Consideration of the interaction between the vacancies leads to lowering of the vacancy formation energy ε_f and alters their diffusion coefficient.[33] At elevated temperatures the formation of vacancies that are not isolated and vacancy–native-adatom pairs is quite likely. Since native adatoms make their own contribution to the processes occurring on the crystal surface, their contribution to the observed values should be considered separately.

This research was carried out with partial financial support from NATO (Grant No. HTECH.LG 960939) and the NSF (Grant No. DMR-9705367).

*)E-mail: devyatko@veshn.mephi.msk.su

[1] R. Schneider, H. Durr, T. Fauster, and V. Dose, Phys. Rev. B **42**, 1638 (1990).
[2] P. Statiris, H. C. Lu, and T. Gustafsson, Phys. Rev. Lett. **72**, 3574 (1994).
[3] D. E. Fowler and J. V. Barth, Phys. Rev. B **52**, 2117 (1995).
[4] H. C. Lu, E. P. Gusev, E. Garfunkel, and T. Gustafsson, Surf. Sci. **352–354**, 21 (1996).
[5] E. P. Gusev, H. C. Lu, E. Garfunkel, and T. Gustafsson., Surf. Rev. Lett. **3**, 1349 (1996).
[6] K. H. Chae, H. C. Lu, and T. Gustafsson, Phys. Rev. B **54**, 14 082 (1996).
[7] B. W. Busch and T. Gustafsson, Surf. Sci. **407**, 7 (1998).
[8] Q. T. Jiang, P. Fenter, and T. Gustafsson, Phys. Rev. B **42**, 9291 (1990).
[9] Q. T. Jiang, P. Fenter, and T. Gustafsson, Phys. Rev. B **44**, 5773 (1991).
[10] Y. Cao and E. Conrad, Phys. Rev. Lett. **64**, 447 (1990).
[11] J. W. M. Frenken, F. Huussen, and J. F. Van der Veen, Phys. Rev. Lett. **58**, 401 (1987).
[12] Y. Cao and E. Conrad, Phys. Rev. Lett. **65**, 2808 (1990).
[13] L. L. Kesmodel, M. L. Xu, and S. Y. Tong, Phys. Rev. B **34**, 2010 (1986).
[14] M. Wuttig, R. Franchy, and H. Ibach, Solid State Commun. **57**, 445 (1986).
[15] J. W. Flocken, Phys. Rev. B **9**, 5153 (1974).
[16] L. Yang, T. Rahman, and M. S. Daw, Phys. Rev. B **44**, 13 725 (1991).
[17] Y. Beaudet, L. J. Lewis, and M. Persson, Phys. Rev. B **47**, 4127 (1993).
[18] L. J. Lewis, Phys. Rev. B **50**, 17 693 (1994).
[19] S. Narshimhan and M. Scheffler, Z. Phys. Chem. **202**, 253 (1997).
[20] A. Kara, A. Al-Rawi, and T. S. Rahman, Phys. Rev. B **50**, 13 440 (1997).
[21] T. S. Rahman, Z. J. Tian, and J. E. Black, Surf. Sci. **374**, 9 (1997).
[22] L. Yang and T. Rahman, Phys. Rev. Lett. **67**, 2327 (1991).
[23] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).
[24] V. Heine and D. Weaire, ''Pseudopotential theory of cohesion and structure,'' in *Solid State Physics, Vol. 24*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York–London, 1970), pp. 250–463.
[25] S. M. Folies, M. I. Baskes, and M. S. Daw, Phys. Rev. B **33**, 7983 (1986).
[26] H..-N. Yang, K. Fang, G.-C. Wang, and T.-M. Lu, Phys. Rev. B **44**, 1306 (1991).
[27] Y. Uchida and G. Lehmpfuhl, Surf. Sci. **243**, 193 (1991).
[28] E. T. Chen, R. N. Barnett, and u. Landman, Phys. Rev. B **41**, 439 (1990).
[29] S. V. Eremeev, A. G. Lipnitskii, A. I. Potekaev, and E. V. Chulkov, Phys. Low-Dimens. Semicond. Struct. **3–4**, 127 (1997).
[30] A. M. Stoneham, *Theory of Defects in Solids* (Clarendon Press, Oxford, 1975).
[31] Yu. N. Devyatko and O. V. Tapinskaya, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 1414 (1990).
[32] I. Stensgaard, R. Feidenhans'l, and J. E. Sorensen, Surf. Sci. **128**, 281 (1983).
[33] Y. N. Devyatko and V. N. Tronin, Phys. Scr. **41**, 355 (1990).
[34] H. M. Palatoglou, M. Methfessel, and M. Scheffler, Phys. Rev. B **45**, 1877 (1993).
[35] H. E. Schaefer, Phys. Status Solidi A **102**, 47 (1987).
[36] P. Erhart, in *Physics of Radiation Effects in Crystals*, edited by R. A. Jonson and A. N. Orlov (North-Holland, 1986).

Translated by P. Shelnitz

Berezinskiĭ–Kosterlitz–Thouless phase transitions in two-dimensional systems with internal symmetries

S. A. Bulgadaev*)

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia
(Submitted 24 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 2045–2057 (December 1999)

The Berezinskiĭ–Kosterlitz–Thouless (BKT) phase transitions in two-dimensional systems with internal continuous Abelian symmetries are investigated. In order for phase transitions to occur, the kinetic part of the action of the system must have conformal invariance, and the vacuum manifold must be degenerate and have a discrete Abelian homotopy group π_1. In this case topological excitations have a logarithmically divergent energy and can be described by effective theories that generalize the two-dimensional Euclidean sine-Gordon theory, which is an effective theory of the original XY model. In particular, the effective actions are found for chiral models on the maximal Abelian tori T_G of the simple compact Lie groups G. The critical properties of the possible effective theories are found, and it is shown that they are characterized by the Coxeter numbers h_G of lattices of the $\mathbb{A}$, $\mathbb{D}$, $\mathbb{E}$, and $\mathbb{Z}$ series and can be interpreted as properties of conformal theories with an integer central charge $C=n$, where n is the rank of the groups π_1 and G. The possibility of reconstructing the complete symmetry of G in the massive phase is also discussed. © *1999 American Institute of Physics.*
[S1063-7761(99)01312-8]

1. INTRODUCTION

The discovery of the possibility of a phase transition in the two-dimensional XY model[1] at once attracted a great deal of interest on the part of theoreticians due to its unusual properties. First and foremost, it seemed that the existence of such a phase transition contradicts the Peierls–Landau[2,3] and Bogolyubov–Goldstone[4,5] theorems, which state that spontaneous symmetry breaking and spontaneous magnetization cannot occur in low-dimensional systems ($d\leqslant 2$) with continuous symmetry.[6,7] Second, because of the absence of spontaneous magnetization the correlation functions in the low-temperature phase should decay according to a power law.[8,9] This means that the entire low-temperature phase must be massless.

All these contradictions were brilliantly resolved in a series of papers by Berezinskiĭ,[10] Popov,[11] Kosterlitz, and Thouless.[12,13] They were the first to demonstrate the important role of topological excitations, i.e., vortices, with a logarithmically divergent energy in these phase transitions. Vortices exist because the value manifold of the XY model $\mathcal{M}=S^1$ has a nontrivial topology, which can be described by a discrete Abelian homotopy group $\pi_1(\mathcal{M})=\mathbb{Z}$, while the logarithmic divergence of the energy is associated with the conformal symmetry of the model. Allowance for vortices transforms the continuous compact symmetry $U(1)$ into the dual discrete symmetry $Z_2\times\mathbb{Z}$, where Z_2 is the automorphism group of a circle S^1, which coincides with the reflection $U(1)=S^1$. Similar phase transitions take place in other systems with the same symmetry: two-dimensional SOS and 6-vertex lattice models, quantum spin XXZ chains,[14] and the Euclidian sine-Gordon model with a noncompact field.[15–18] All these systems belong to a single critical universality class. The sine-Gordon model can be regarded as an effective theory of Berezinskiĭ–Kosterlitz–Thouless (BKT) phase transitions, just as the Ginzburg–Landau–Wilson theories are effective theories of second-order phase transitions (see, for example, Ref. 19).

BKT phase transitions can also be related to a conformal theory, but here there is an additional subtlety. Unlike an ordinary second-order phase transition in a two-dimensional system, where an infinite-dimensional conformal symmetry with a rational central charge C exists only at the phase-transition point,[20] in systems with a BKT phase transition an infinite-dimensional conformal symmetry with an integer central charge $C=1$ exists not only at the transition point (with logarithmic corrections), but also in the entire low-temperature phase.

Hence it follows that a BKT phase transition is intimately related to two fundamental properties of two-dimensional systems: 1) the nontrivial topology described by the discrete Abelian homotopy group π_1 and 2) the conformal symmetry. It would be interesting to examine the properties of BKT phase transitions in systems having internal symmetries in addition to these two properties.

Such systems are associated, for example, with tori, which are a natural generalization of the circle S^1 with the necessary properties. It is easy to see that the same critical properties exist in two-dimensional chiral models on tori T^n with $\pi_1(T^n)=\mathbb{Z}^n$. This case effectively reduces to the preceding case, since only excitations with the minimal topological charges $e_i=\pm 1$, $i=1,\ldots,n$, are important and charges of this kind that correspond to different circles do not interact with one another. The same properties are exhibited by σ models on common tori associated with arbitrary

nondegenerate lattices L (Ref. 21). However, as was shown in Ref. 21, apart from such tori there are maximal Abelian tori T_G of simple compact Lie groups G, which have (in the case of simply connected G) $\pi_1(T_G)=L_v\neq\mathbb{Z}^n$ (here L_v is the lattice of dual roots of the corresponding Lie algebra $\mathcal{G}$) and in which excitations with different vectorial topological charges interact with one another.

The following question thus arises: how do the critical properties of the topological phase transition mentioned above depend on G? This question is important, for example, for string theory, where various aspects of compactification on T_G are considered (more precisely, it has heretofore been considered only on the nondegenerate tori $T^n=T_{U(n)}$ or $T_L=\mathbb{R}^n/L$, where L is a nondegenerate lattice of rank n),[22,23] or for chiral models on G with a reduced (or partially broken) symmetry $G\searrow T_G$.[24]

In this paper it is shown that:

1) all the critical properties of nonlinear σ models on compact T_G can be described in terms of effective field theories with discrete symmetries that generalize the sine-Gordon theory;

2) these properties depend only in the Coxeter number h_L of the corresponding lattice of topological charges L^t;

3) different universality classes of a BKT phase transition are defined by the $\mathbb{A}$, $\mathbb{D}$, $\mathbb{E}$, and $\mathbb{Z}$ series of integer-valued lattices;

4) all the critical and low-temperature properties of these σ models (apart from the logarithmic corrections at the phase-transition point) can be described by corresponding conformal theories with an integer central charge $C=n$, where n is the rank of $\pi_1(T_G)$ and G.

The possibility of reconstructing the complete symmetry group G in the massive (high-temperature) phase is also discussed.

2. NONLINEAR σ MODEL ON T_G AND VORTICES

Let us consider the two-dimensional Euclidean chiral field theories on T_G that generalize the nonlinear σ model on a circle S^1 or the continuous XY model. Their action has the following form:

$$\begin{aligned}\mathcal{S}&=\frac{1}{2\alpha}\int d^2x\,\mathrm{Tr}_\tau(\mathbf{t}_\nu^{-1}\mathbf{t}_\nu)\\&=\frac{(2\pi)^2}{2\alpha}\int d^2x\,\mathrm{Tr}_\tau(\mathbf{H}\cdot\boldsymbol{\phi}_\nu)^2\\&=\frac{(2\pi)^2}{2\alpha}N_\tau\int d^2x\,(\boldsymbol{\phi}_\nu)^2,\end{aligned}\tag{1}$$

where $\mathbf{t}=\exp\{2\pi i(\mathbf{H}\cdot\boldsymbol{\phi})\}\in T_G$, $\mathbf{H}=(H_1,\ldots,H_n)$ belongs to $\mathcal{C}$, i.e., the maximal Cartan subalgebra of the corresponding Lie algebra $\mathcal{G}$, $[H_i,H_j]=0$, n is the rank of the group G, $\boldsymbol{\phi}_\nu=\partial_\nu\boldsymbol{\phi}$, and $\nu=1,2$. Here we have utilized the isotropy of the system of weights $\{\mathbf{w}\}_\tau$ of any $\tau(G)$ representation, which is a consequence of the invariance of systems of weights toward the discrete Weyl group $W_G\in O(n)$,

$$\sum_a w_i^a w_k^a=N_\tau\delta_{ik},\quad a=1,\ldots,\dim\tau(G).\tag{1a}$$

Below it will be convenient to include the constant N_τ as a normalizing factor in the definition of the trace Tr_τ. This gives the canonical Euclidean metric in the space of topological charges.

The theories (1), like other two-dimensional chiral models, are invariant with respect to the action of the direct product of the right-hand (R) and left-hand (L) groups $N_G^{R(L)}$, which are the semidirect product of T_G and W_G:

$$N_G=T_G\times W_G.\tag{2}$$

The group N_G, which is called the normalizer of the torus T_G, is the symmetry group of the torus T_G.

These theories comprise a multicomponent generalization of the XY model, which has properties similar to the properties of the XY model:

1) a zero beta function $\beta(\alpha)$ (due to the flatness of T_G) and unbroken symmetry when the topology is disregarded;

2) a nontrivial homotopy group π_1 and corresponding vortex solutions.

The classical equations of the theories (1)

$$(\partial_\nu)^2(\mathbf{H}\cdot\boldsymbol{\phi})=0\tag{3}$$

have the following vortex solutions in the region $R>r>a$, where R is the radius of the system and a is the short-wavelength cutoff parameter (for example, the vortex core radius):

$$\mathbf{t}(\vartheta)=\exp\{2\pi i(\mathbf{H}\cdot\boldsymbol{\phi}(\vartheta))\},\quad \boldsymbol{\phi}=\frac{1}{2\pi}\mathbf{q}(\vartheta).\tag{4}$$

Here ϑ and r are the angular and radial coordinates in the $\mathbb{R}^2$ plane, $\mathbf{q}$ is the vectorial topological charge of the vortex, $\mathbf{q}\in L_\tau^t=L_\tau^{-1}$, L_τ^t is the lattice of all possible topological charges of the τ representation, and L_τ^{-1} is the lattice of vectors which are reciprocals of all the weights of the τ representation:

$$\mathbf{q}\in L_\tau^t,\quad \mathbf{w}_a\in\{\mathbf{w}_\tau\},\quad (\mathbf{q}\cdot\mathbf{w}_a)\in\mathbb{Z}.\tag{5}$$

For the minimal fundamental representations of the simply connected groups $\tau(G)=\min$ we have $L^t_{\min}=L_v$, and for the adjoint representations $\tau=\mathrm{ad}$ we have $L^t_{\mathrm{ad}}=L_r^{-1}=L_{w*}$, where L_r is the lattice of roots of the group G, and L_{w*} is the lattice of dual weights or the lattice of weights of the dual group G^*. Just these solutions for all the groups G which are such that $L_\tau^t\supseteq L_w$ can give a topological interpretation of all their quantum numbers.[21] The energy of these vortices diverges logarithmically:

$$E=\frac{(2\pi)^2}{2\alpha}\int(\partial_\mu\boldsymbol{\phi})^2\,d^2x=\frac{2\pi}{2\alpha}\mathbf{q}^2\ln\left(\frac{R}{a}\right).\tag{6}$$

By virtue of formula (2), which defines the effective metric in the space of topological charges,[21] there is a logarithmic interaction between vortices with different vectorial topological charges:

$$E=(\mathbf{q}_1\cdot\mathbf{q}_2)\frac{2\pi}{2\alpha}\ln\frac{|\mathbf{x}_1-\mathbf{x}_2|}{a}.\tag{7}$$

The general N-vortex solution has the following form:[20]

$$\boldsymbol{\phi}(\mathbf{x})=\sum_{i=1}^{N}\ \mathbf{q}_i\frac{1}{\pi}\arctan\left(\frac{y-y_i}{x-x_i}\right), \tag{8}$$

$$\mathbf{q}_i\in L_\tau^t,\quad (\mathbf{q}_i\cdot\mathbf{w}_a)\in\mathbb{Z},\quad (x,y)\in\mathbb{R}^2.$$

The energy of the Nth vortex solution with a zero total topological charge $\Sigma_{i=1}^{N}\mathbf{q}_i=0$ equals

$$E_N=\sum_i\ E_{q_i}^0+E_{N\,\mathrm{int}},\quad E_{q_i}^0=\frac{1}{2\alpha}C(a)(\mathbf{q}_i\cdot\mathbf{q}_i),$$

$$E_{N\,\mathrm{int}}=\frac{2\pi}{2\alpha}\sum_{i\neq k}^{N}(\mathbf{q}_i\cdot\mathbf{q}_k)\ln\frac{|x_i-x_k|}{a}, \tag{9}$$

where $E_{q_i}^0$ is the ''self-energy'' (or the core energy) of a vortex with the topological charge $\mathbf{q}_i$, and $C(a)$ is a nonuniversal constant, which depends on the vortex core regularization method. Only such solutions make a finite contribution to the partition function $\mathcal{Z}$ of the theory. Since $E_q\sim q^2$ and $\mathbf{q}\in L^t$, the maximal contribution to each N-vortex sector of solutions is made by the vortices with the minimal $|q|_i$. Therefore, in the quasiclassical approximation (or in the low-temperature expansion) we can represent the partition function of the theory (5)

$$\mathcal{Z}=\int D\boldsymbol{\phi}\exp(-\mathcal{S}[\boldsymbol{\phi}]) \tag{10}$$

in the form of the large partition function of a classical, neutral as a whole, Coulomb gas of vortex solutions with the minimal vectorial topological charges $\mathbf{q}_i\in\{\mathbf{q}\}_\tau$, where $\{\mathbf{q}\}_\tau$ is the set of minimal vectors of the lattice L_τ^t:

$$\mathcal{Z}=\mathcal{Z}_0\mathcal{Z}_{CG},\quad \mathcal{Z}_{CG}=\sum_{N=0}^{\infty}\frac{\mu^{2N}}{N!}{\sum_{\{\mathbf{q}\}}}'\ \mathcal{Z}_N(\{\mathbf{q}\}|\beta). \tag{11}$$

Here the summation Σ' is carried out over all the neutral configurations of the minimal charges $\mathbf{q}_i\in\{\mathbf{q}\}_\tau$ under the condition $\Sigma_1^N\mathbf{q}_i=0$, and $\mathcal{Z}_0$ is the partition function of a free massless isovectorial boson field, which corresponds to the ''spin waves'' of the XY model,

$$\mathcal{Z}_0=\int D\boldsymbol{\phi}\exp(-\mathcal{S}_0[\boldsymbol{\phi}]), \tag{12}$$

$$\mathcal{Z}_N(\{\mathbf{q}\}|\beta)=\prod_{i=1}^{N}\int d^2x_i\exp(-\beta H_N(\{\mathbf{q}\})), \tag{13}$$

$$H_N(\{\mathbf{q}\})=\sum_{i<j}^{N}(\mathbf{q}_i\cdot\mathbf{q}_j)D(x_i-x_j), \tag{14}$$

$$D(x)=\int\frac{d^2k}{(2\pi)^2}[\exp(i(\mathbf{k}\cdot\mathbf{x}))-1]\times\frac{f(ka)}{k^2}\underset{|x|\gg a}{\sim}\frac{1}{2\pi}\ln\left|\frac{x}{a}\right|, \tag{15}$$

where

$$\mu^2=a^{-2}y^2(\det)^{-1/2},\quad y^2=\exp(-E_q^0) \tag{15a}$$

is the chemical activity of the Coulomb gas, det denotes the determinant of the quadratic fluctuations on a background of the vortex solution (we shall henceforth set det$=1$),

$$\beta=4\pi^2/\alpha, \tag{15b}$$

and $f(ka)$ is a regularizing function, which is such that

$$\lim_{k\to 0}f(ka)=1,\quad \lim_{k\to\infty}f(ka)=0.$$

The next section clarifies how allowance for vortices alters the original symmetry group N_G of a σ model.

3. DUALITY OF COMPACT AND NONCOMPACT THEORIES

In the case of the XY model, in the long-wavelength quasiclassical approximation there is a very important relation between the partition function of the compact chiral theory (1) on S^1 and the partition function of the noncompact sine-Gordon theory[15–17] with the action (modulo of $\mathcal{Z}_0$)

$$\mathcal{S}_{SG}=\int d^2x\left(\frac{1}{2\beta}(\partial_\mu\phi)^2-2\mu^2\cos\phi\right). \tag{16}$$

This action is invariant relative to the dual discrete group $Z_2\times\mathbb{Z}$. There is a similar relation between the compact chiral models on T_G and the noncompact generalized sine-Gordon theories.

In order to see this, we note that the large partition function $\mathcal{Z}_{CG}$ from (11) is, in turn, equivalent to the partition function of a noncompact scalar isovectorial field theory:

$$\mathcal{Z}_{CG}=\int D\boldsymbol{\phi}\exp(-\mathcal{S}_{\mathrm{eff}}),\quad \mathcal{S}_{\mathrm{eff}}=\int\frac{1}{2\beta}(\partial\boldsymbol{\phi})^2-\mu^2V(\boldsymbol{\phi}), \tag{17}$$

$$V(\boldsymbol{\phi})=\sum_{\{\mathbf{q}\}}\ \exp\{i(\mathbf{q}\cdot\boldsymbol{\phi})\}, \tag{18}$$

where the summation is carried out over the set of minimal topological charges $\{\mathbf{q}\}$, and $\boldsymbol{\phi}\in\mathbb{R}^n$ (Ref. 25). Strictly speaking, the theories (17) with the arbitrary parameters μ and β are more general than the original σ models (1). The latter have only one parameter, viz., the interaction constant α. Representing the σ models in the form of (11) and (17) provides a means to fit them into the general theories (17), since the constraints (15a) and (15b), which relate the parameters μ and β, exist. This fact will be important below in a discussion of the possibility of raising the symmetry of σ models (Sec. 6).

Since the set of minimal charges $\{\mathbf{q}\}_\tau$ is invariant toward the dual Weyl group W_{G*}, it can be seen that allowance for vortices reduces the original symmetry group N_G to the discrete dual group $W_{G*}\times L_q^{-1}$. Here L_q^{-1} is the periodicity lattice of the potential V and is consequently the reciprocal to all $\mathbf{q}\in\{\mathbf{q}\}$. It follows from the definition of these lattices that $L_q^{-1}=L_\tau$. This dual group generalizes the dual group $Z_2\times\mathbb{Z}$ of the XY model.

Thus, in this quasiclassical and long-wavelength approximation the compact theory on the torus T_G with the continuous symmetry N_G is equivalent (modulo of $\mathcal{Z}_0$) to the noncompact theory with a periodic potential and an infinite discrete symmetry. These potentials contain the sum over all

the minimal vectors $\{\mathbf{q}\}$ and can coincide with the characters of some representations of G. For example, in the case of $L_\tau^{-1}=L_v$ the summation in (18) is carried out over all the dual minimal roots. Therefore, the corresponding potentials V for simply laced groups from the A, D, and E series coincide with the characters of the adjoint representations of these groups (modulo a constant corresponding to zero weight). In this case the general theories (17) can describe systems with the symmetry of G broken to N_G (Ref. 25).

The noncompact theories (17) can also be regarded as corresponding linear σ models. As a result, the compact nonlinear σ models on the tori T_G turn out to be equivalent (in the approximation under consideration) to the noncompact linear σ models on Cartan tori of the dual group T_{G*}.

The ensuing treatment calls for classification of all the possible effective theories of this type. It follows from (17) and (18) that they are determined by the sets of minimal vectors $\{\mathbf{q}\}$ of the lattice L_τ^t which satisfies the following constraint:

$$L_{w*}\supseteq L_\tau^t\supseteq L_v. \tag{19}$$

For $\tau=\min$ we have $L_\tau^t=L_v$, and for $\tau=\mathrm{ad}$ we have $L_\tau^t=L_{w*}$. The lattices L_v and L_{w*} differ by a factor which is isomorphic to the center Z_G of the group G:

$$L_{w*}/L_v=Z_G.$$

Therefore, the set $\{\mathbf{q}\}$ can vary from the set of minimal vectors (it defines the so-called Voronoi polyhedron or Wigner–Seitz cell of the corresponding lattice) of the weight lattice to the set of minimal vectors of the root lattice. All the possible cases are specified by subgroups of the group of the center Z_G. For groups G with $Z_G=1$ the lattices L_v and L_{w*} coincide.

4. PHASE TRANSITION IN CHIRAL MODELS ON T_G

In this section we consider topological phase transitions in chiral models on T_G using the approximate equivalence obtained above between these theories and the noncompact theories (17), which generalize the sine-Gordon field theory. These theories can be regarded as effective theories for topological phase transitions, just as the Ginzburg–Landau–Wilson theories are effective theories for second-order phase transitions.[19]

An investigation of the BKT phase transitions for all the effective field theories of the type (17) by renormalization group analysis was carried out in Ref. 25. It was shown there that only theories associated with even integer-valued $\mathbb{A}$, $\mathbb{D}$, and $\mathbb{E}$ lattices can have new critical properties. They have the structure of the root lattices of the corresponding simple groups G from the so-called simply laced groups of the A, D, and E series. All theories associated with other lattices have the same critical properties as the sine-Gordon theory or superpositions thereof that are associated with the lattice $\mathbb{Z}^n$. Therefore, here we give only a brief review of the results obtained, focusing mainly on the symmetry and universality properties.

Upon renormalization, both μ and β are renormalized. It is convenient to introduce two dimensionless parameters into the treatment:

$$(\mu a)^2=g,\quad \delta=\frac{\beta q^2-8\pi}{8\pi}, \tag{20}$$

where q^2 is the square of the norm of the minimal vectorial topological charges from $\{\mathbf{q}\}$. The theories (17) are renormalizable only if vectors from $\{\mathbf{q}\}$ belong to some lattice (in our case L_τ^t). New critical properties can appear only if the geometry of the set $\{\mathbf{q}\}$ is such that each vector $\mathbf{q}\in\{\mathbf{q}\}$ can be presented in the form of a sum of two other vectors from $\{\mathbf{q}\}$ (Ref. 25). The latter property is very restrictive and coincides with the definition of the root systems $\{\mathbf{r}\}$ of simple groups from the A, D, and E series (Ref. 25) or with the definition of the root system of even integer-valued (on some scale) lattices of types $\mathbb{A}$, $\mathbb{D}$, and $\mathbb{E}$ (Ref. 26). The sets of minimal roots (or minimal dual roots) of all simple groups belong to the four series of integer-valued lattices $\mathbb{A}$, $\mathbb{D}$, $\mathbb{E}$, and $\mathbb{Z}$. For the theories (17), for which the sets $\{\mathbf{q}\}\notin\mathbb{A},\mathbb{D},\mathbb{E}$, all the critical properties will be the same as for the theories with $\{\mathbf{q}\}\in\mathbb{Z}^n$ (Ref. 25).

The renormalization group equations for the theories (17) with $\{\mathbf{q}\}$ belonging to lattices of the series $G=A,D,E$ have the following form:[25]

$$\frac{dg}{dl}=-2\delta g+B_G g^2,\quad \frac{d\delta}{dl}=-C_G g^2. \tag{21}$$

Here $B_G=\pi\theta_G$, θ_G is the reproduction multiplicity of the potential $V(\boldsymbol{\phi})$ upon the renormalization of (17) or the number of different ways of representing each root in the form of a sum of two other roots, and $C_G=2\pi K_G$, where K_G is the value of the second-order Casimir operator in an adjoint representation (where $\mathbf{w}_a=\mathbf{r}_a$)

$$\sum_a r_i^a r_j^a=K_G\delta_{ij}. \tag{22}$$

Renormalization group equations of the form (21) with coefficients corresponding to the case $G=A_2$ were first obtained in Ref. 27 in a study of the melting of two-dimensional triangular lattices. The renormalization group equations for lattices not belonging to the $\mathbb{A}$, $\mathbb{D}$, and $\mathbb{E}$ series have the form (21) with the coefficient $B_G=0$.

The value of the second-order Casimir operator K_G for groups of the A, D, and E series can be expressed in terms of the corresponding Coxeter number h_G:

$$K_G=2h_G,\quad h_G=\frac{(\text{number of roots})}{(\text{rank of group})}. \tag{23}$$

This definition of the Coxeter number coincides with the one for the Coxeter number of the corresponding lattices from the $\mathbb{A}$, $\mathbb{D}$, and $\mathbb{E}$ series. The coefficient B_G can be calculated by different methods and can also be expressed in terms of the Coxeter number:

$$\theta_G=K_G-4=2(h_G-2). \tag{23a}$$

Thus, we see that all the coefficients in the renormalization group equation are expressed in terms of the Coxeter number

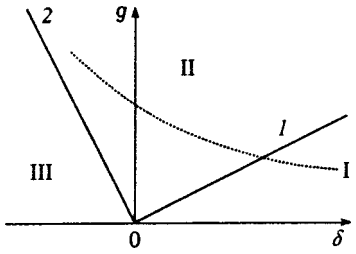

FIG. 1. Schematic phase diagram.

h_G or in terms of the value of the second-order Casimir operator K_G. The renormalization group equations (21) have two separatrices:[25,27]

$$u_{1,2}\equiv\left(\frac{g}{\delta}\right)_{1,2}=\frac{1}{2C_G}[\pm(B_G^2+8C_G)^{1/2}-B_G], \tag{24}$$

where u_1 corresponds to the phase separation line. The critical exponent ν_G, which determines the divergence of the correlation length ξ as the transition point is approached from above with respect to the temperature T,

$$\xi\sim a\exp(A\tau^{-\nu_G}),\quad \tau=\frac{T-T_c}{T_c},$$

is given by the following expression:

$$\nu_G=1/\kappa_G=u_1[(B_G/C_G)^2+8/C_G)]^{-1/2}, \tag{25}$$

where $1/\kappa_G$ is the Lyapunov exponent on separatrix *1*.[25] Substituting the corresponding values for the coefficients into (24), we obtain the following values for the separatrix slopes:

$$u_{1,2}=\begin{cases}1/\pi h_G,\\ -1/2\pi.\end{cases} \tag{26}$$

We note that the separatrix $u_2=-1/2\pi$ does not depend on G and is equal to a universal constant. This fact is very important for the possibility of reconstructing the complete symmetry of G in the massive (high-temperature) phase (see Sec. 6 below). A schematic phase diagram is shown in Fig. 1.

The dotted line of initial values corresponds to the original σ model. This line is specified by the dependences of the parameters β and μ on the interaction constant α [Eqs. (15a) and (15b)]. Region I corresponds to the low-temperature (decompactified, massless) phase, and the other regions correspond to the high-temperature (compact, massive) phase. In region I the correlation length $\xi=\infty$, and in region II near separatrix *1*

$$\xi\sim a\exp(A\tau^{-\nu_G}),\quad \nu_G=2/(2+h_G)=4/(4+K_G). \tag{27}$$

Using known values for the Coxeter number h_G and data on the geometry of sets of minimal dual roots, we obtain the following expressions for the critical exponent ν_G:[25]

G:	A_n	B_n	C_n	D_n	G_2	F_4	E_6	E_7	E_8
ν_G:	$\frac{2}{n+3}$	$\frac{1}{n}$	$\frac{1}{2}$	$\frac{1}{n}$	$\frac{2}{5}$	$\frac{1}{4}$	$\frac{1}{7}$	$\frac{1}{10}$	$\frac{1}{16}$

The dependence of the critical properties of the theories (1) and (17) only on such a fairly rough characteristic of groups as the Coxeter number h_G or K_G leads to their convergence for fairly different groups. In particular, it is interesting to note that the $D_{16}=O(32)$ and E_8 groups, which are used to construct anomaly-free string theories,[22] have identical values of ν_G (together with the A_{29} group). The largest number of possible values of ν_G is given by the A_n series: $1/k$ and $2/(2k+1)$, where k is an integer. For theories with the potentials V containing sets of minimal roots, all the exponents, except ν_{B_n} and ν_{C_n}, remain unchanged. The latter exponents transform into one another due to the mutual duality of these groups.

5. LOW-TEMPERATURE PHASE AND CONFORMAL SYMMETRY

The equality of the correlation length $\xi\to\infty$ everywhere in the low-temperature phase corresponds to the existence of conformal symmetry at large distances. This can also be seen from the renormalized effective action $\mathcal{S}_{\rm eff}$ of the theory, which takes the following asymptotic form in the infrared (IR) limit:

$$\mathcal{S}_{\rm eff}=\int d^2x\,\frac{1}{2\overline{\beta}}(\partial\boldsymbol{\phi})^2, \tag{28}$$

where $\overline{\beta}$ is the value of the renormalized parameter $\beta(l)$ in the IR limit

$$\overline{\beta}=\lim_{l\to\infty}\beta(l). \tag{29}$$

At the phase-transition point $\overline{\beta}=\beta^*=8\pi/q_{\min}^2$. At other points of the low-temperature phase $\overline{\beta}$ depends on the initial values of the system parameters. It is generally known that the action (28) describes a free conformal theory with a central charge $C=n$, where n is the rank of the group G, which specifies both the rank of the torus T_G and the rank of the group $\pi_1(T_G)$. Hence it follows that the long-wavelength low-temperature properties of σ models defined on different tori T_G, will be identical for all groups with the same rank n. Only the logarithmic corrections at the phase-transition point will depend on G through the Coxeter number h_G. It thus becomes clear why the critical exponent depend only on h_G and K_G. This is consistent with the fact that in compact groups all the quantum anomalies also depend only on h_G (or the dual Coxeter number $\tilde{h}_G$).[28] Here it should be noted that ν_G coincides with the ''screening'' factor in the formula for the central charge C_k of the affine algebra $\hat{G}$ (Ref. 28),

$$C_k=\frac{k}{k+h_G}\dim G, \tag{30}$$

at the $k=2$ level or in the formula for C_k in the ''coset'' realization of the minimal unitary conformal models $\hat{G}_k\otimes\hat{G}_1/\hat{G}_{k+1}$ (Ref. 29),

$$C_k=n\left(1-\frac{h_G(h_G+1)}{(h_G+k)(h_G+k+1)}\right), \tag{31}$$

at the $k=1$ level. It also follows from Eq. (31) and Ref. 29 that the phase-transition point in a σ model on T_G ($G=A,D,E$) is the limiting case ($k\to\infty$) for the sequence of minimal unitary conformal models corresponding to the respective expanded conformal groups W_G.

The fact that the theory effectively becomes free in the low-temperature phase permits the calculation of correlation functions. For example, the following expressions are obtained for the correlation functions of the exponential functions of a field:

$$\left\langle \prod_{s=1}^{t} \exp(i(\mathbf{r}_s\cdot\boldsymbol{\phi}(x_s))) \right\rangle = \prod_{i\neq j}^{t} \left|\frac{x_i-x_j}{a}\right|^{\bar{\beta}(\mathbf{r}_i\cdot\mathbf{r}_j)/2\pi},$$

$$\sum_{i=1}^{t}\mathbf{r}=0. \tag{32}$$

At the phase-transition point (where $\bar{\beta}=\beta^*=8\pi/q^2=4\pi$) an additional logarithmic factor, which is associated with the ''zero-charge'' behavior of g and δ on the critical separatrix, appears in them:

$$\prod_{i\neq j}^{t}\left(\ln\left|\frac{x_i-x_j}{a}\right|\right)^{\beta^*(\mathbf{r}_i\cdot\mathbf{r}_j)/2\pi A_G} = \prod_{i\neq j}^{t}\left(\ln\left|\frac{x_i-x_j}{a}\right|\right)^{h_G\cos(\mathbf{r}_i\cdot\mathbf{r}_j)}. \tag{33}$$

Here $A_G=4/h_G$ is the coefficient in the renormalization group equation for δ on the critical separatrix.

6. MASSIVE PHASE, ASYMPTOTIC FREEDOM, AND GLOBAL SYMMETRY

Regions II and III correspond to the IR limit of the high-temperature (in the language of statistical physics) or massive (in the language of field theory) phase. In the ultraviolet (UV) limit region III will be asymptotically free. Separatrix *2* with the slope $u_2=-1/2\pi$ also plays an important role. In the UV limit it marks the boundary of asymptotically free region III. There is also another possibility for increasing the symmetry group of the original nonlinear σ model on this separatrix. On the classical level the σ model (1) has two symmetries: 1) gauge (or conformal) symmetry and 2) isotopic global symmetry, $N_G=T_G\times W_G$. On the quantum level the former symmetry is spontaneously broken in the IR region by vortices in the general case [see (11)]. For this reason the σ model has a finite correlation length $\xi\sim m^{-1}$, where m is the characteristic mass scale of the theory, in the massive phase. It should depend on the interaction constant α or β. The behavior of m near the phase-transition point is described by formula (27), where

$$\tau\sim\frac{\alpha-\alpha_c}{\alpha_c}.$$

In the massive phase there is another region, separatrix *2*, where the function $m(\alpha)$ can also be found. Since this separatrix attracts all the renormalization group trajectories in the massive (or high-temperature) phase, it is very important to know the effective mass scale on it. In the leading logarithmic approximation with respect to g it is given by the pole in the solution of the renormalization equation on this separatrix[30] or by the formula

$$m\sim\Lambda\exp\left(-\int^{g}\frac{dx}{\beta(x)}\right),$$

where $\Lambda\sim a^{-1}$ is the UV cutoff parameter in momentum space, $\beta(x)$ is the β function on separatrix *2*, and

$$\beta(g)=2\pi g^2K_G/2=2\pi g^2h_G.$$

Hence we obtain

$$m\sim\Lambda\exp(-1/2\pi g h_G)=\Lambda\exp(-1/\pi g K_G). \tag{34}$$

The numerical factor in the β function can vary as a function of the normalization of the interaction constant, but the fact that $\beta\sim K_G\sim h_G$ on separatrix *2* is a consequence of the aforementioned lack of a dependence of the slope of the separatrix u_2 on G.

The expression (34) for the mass scale on separatrix *2*, which depends only on K_G, coincides with the expression for chiral models on the groups G (Ref. 30) and the expression obtained from the exact solution of the corresponding chiral and fermionic theories (in the leading approximation with respect to g)[31]

$$m\sim\Lambda\exp(-2\pi/(gK_G/2)).$$

Thus, it has been shown that on separatrix *2* the mass scale coincides (at least for groups $G=A,D,E$) with the one in G-invariant theories (chiral and fermionic theories) and can be expressed only in terms of the value of the Casimir operator K_G or the Coxeter number h_G by the universal formula [here $g\to g/(2\pi)^2$]:

$$m\sim\Lambda\exp(-4\pi/gK_G)=\Lambda\exp(-2\pi/gh_G). \tag{35}$$

This means that the theories (1) and (17) can be G-invariant on separatrix *2*. This is also indicated by the equivalence of the general theories (17) with $\{\mathbf{q}\}=\{\mathbf{r}\}$ in the case of $G=A_{n-1},D_n,E_{6,7}$ to fermionic theories with the same global symmetry G (Ref. 25a).

It follows from the results obtained that for the massive phase of chiral theories on T_G ($G=A,D,E$) in the minimal representation (where $L^t_{\min}=L_v$) there is a strong dependence of the mass scale on the interaction constant, which interpolates between formula (27) near the phase-transition point and formula (34) near separatrix *2*. In the former region the theory has T_G symmetry, which is described by its normalizer $N_G=T_G\times W_G$, while in the latter region the theory can be more symmetric and G-invariant. Similar crossover in $m(\alpha)$ also takes place in σ models on other groups and in other representations, but the relationship between the symmetry properties in the two limiting regions in these cases remains not so clear.

This work was carried out with the support of the Russian Fund for Fundamental Research (Grants 96-02-17331 and 96-15-96861).

*)E-mail: bulgad@itp.ac.ru

[1] H. E. Stanley and T. A. Kaplan, Phys. Rev. Lett. **17**, 913 (1966).
[2] L. D. Landau, Zh. Éksp. Teor. Fiz. **7**, 627 (1937).
[3] R. E. Peierls, Ann. Inst. Henri Poincare **5**, 177 (1935).
[4] N. N. Bogolyubov, *Selected Works* [in Russian], Naukova Dumka, Kiev (1971), Vol. 3.
[5] J. Goldstone, Nuovo Cimento **19**, 154 (1961).
[6] N. Mermin and H. Wagner, Phys. Rev. Lett. **17**, 1133 (1966).
[7] P. C. Hohenberg, Phys. Rev. **158**, 383 (1967).
[8] T. M. Rice, Phys. Rev. **140**, 1889 (1965).
[9] B. Jancovici, Phys. Rev. Lett. **19**, 20 (1967).
[10] V. L. Berezinskiĭ, Zh. Éksp. Teor. Fiz. **59**, 907 (1970) [Sov. Phys. JETP **32**, 493 (1970)]; Zh. Éksp. Teor. Fiz. **61**, 1144 (1971) [Sov. Phys. JETP **34**, 610 (1972)].
[11] V. N. Popov, *Functional Integrals in Quantum Field Theory and Statistical Physics*, Reidel, Dordrecht (1983) [Russ. original, Atomizdat, Moscow (1976)].
[12] J. M. Kosterlitz and J. P. Thouless, J. Phys. C **6**, 118 (1973).
[13] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[14] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics*, Academic Press, New York (1982) [Russ. transl., Mir, Moscow (1985)].
[15] J. Jose, L. Kadanoff, S. Kirkpatrick, and D. Nelson, Phys. Rev. B **16**, 1217 (1977).
[16] P. B. Wiegmann, J. Phys. C **11**, 1583 (1978).
[17] T. Ohta, Prog. Theor. Phys. **60**, 968 (1978).
[18] D. J. Amit, Y. Y. Goldschmidt, and G. Grinstein, J. Phys. A **13**, 585 (1980).
[19] A. Z. Patashinskii and V. L. Pokrovskii, *Fluctuation Theory of Phase Transitions*, Pergamon Press, Oxford (1979) [newer ed. of Russ. original, Nauka, Moscow (1982)].
[20] A. A. Belavin, A. M. Polyakov, and A. B. Zamolodchikov, Nucl. Phys. B **241**, 333 (1984); Vl. S. Dotsenko and V. A. Fateev, Nucl. Phys. B **240**, 312 (1984); , **251**, 691 (1985); D. Friedan, Z. Qiu, and S. H. Shenker, Phys. Rev. Lett. **53**, 1575 (1984); G. E. Andrews, R. J. Baxter, and P. J. Forrester, J. Stat. Phys. **35**, 193 (1984); D. A. Huse, Phys. Rev. B **30**, 3908 (1984).
[21] S. A. Bulgadaev, JETP Lett. **63**, 780 (1996); S. A. Bulgadaev, Landau Institute Preprint 29/05/97 (1997), e-print Archive hep-th/9901036; Zh. Éksp. Teor. Fiz. **116**, 1131 (1999) [JETP **89**, 603 (1999)].
[22] M. Green, J. H. Schwarz, and E. Witten, *Superstring Theory*, Cambridge Univ. Press, Cambridge–New York (1987), Vols. 1 and 2; Ya. I. Kogan, JETP Lett. **45**, 709 (1987); D. Gross and I. Klebanov, in *Proceedings of the Trieste Spring School ''String Theory and Quantum Gravity'91''*, World Scientific, Singapore (1991).
[23] T. Banks, W. Fischler, S. H. Shenker, and L. Susskind, Phys. Rev. D **55**, 5112 (1997); T. Banks and N. Seiberg, Nucl. Phys. B **497**, 41 (1997).
[24] S. A. Bulgadaev, talk given at the International Conference ''Conformal Field Theories and Integrable Models,'' Chernogolovka, Russia (1996); Landau Institute Preprint 02/06/97 (1997).
[25] S. A. Bulgadaev, Phys. Lett. A **86**, 213 (1981); Teor. Mat. Fiz. **49**, 7 (1981); Nucl. Phys. B **224**, 349 (1983); JETP Lett. **63**, 780 (1996); S. A. Bulgadaev, Nucl. Phys. B **224**, 349 (1983); Landau Institute Preprint 27/05/97 (1997).
[26] G. H. Conway and N. J. A. Sloane, *Sphere Packing, Lattices and Groups*, Springer-Verlag (1988), Vols. 1 and 2.
[27] D. R. Nelson, Phys. Rev. B **18**, 2318 (1978); D. R. Nelson and B. I. Halperin, Phys. Rev. B **19**, 2457 (1979).
[28] V. N. Kac, *Infinite Dimensional Lie Algebras*, Cambridge University Press (1990).
[29] P. Goddard, A. Kent, and D. Olive, Phys. Lett. B **152**, 88 (1985); Commun. Math. Phys. **103**, 105 (1986); V. A. Fateev and S. L. Lukyanov, Int. J. Mod. Phys. A **13**, 507 (1988); T. Eguchi and S. K. Yang, Phys. Lett. B **224**, 373 (1989).
[30] A. M. Polyakov, *Gauge Fields and Strings*, Harwood Academic Publishers, Chur, Switzerland–New York (1987) [newer ed. of Russ. original, Inst. Teor. Fiz. im L. D. Landau, Moscow (1995)].
[31] E. Ogievetski and P. B. Wiegmann, Phys. Lett. B **168**, 360 (1986); C. Destri and H. J. de Vega, Preprint CERN-TH-4895/87 (1987).

Translated by P. Shelnitz

Cross-phase modulation of surface magnetostatic spin waves

A. O. Korotkevich*)

Moscow Physicotechnical Institute, 141700, Dolgoprudnyĭ, Moscow Region, Russia

S. A. Nikitov

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907, Moscow, Russia
(Submitted 14 July 1999)
Zh. Éksp. Teor. Fiz. **116**, 2058–2068 (December 1999)

This paper explains the modulation instability of two surface magnetostatic spin waves simultaneously propagating in a ferromagnetic film. Self-modulation of the spin waves appears when their power reaches a threshold, and this is a sign of cross-phase modulation. The parameters of the unstable process are calculated, and the gains of the perturbation amplitudes are determined. The results published earlier on the experimental detection of the cross-phase modulation of spin waves are explained. © *1999 American Institute of Physics.* [S1063-7761(99)01412-2]

1. INTRODUCTION

Magnetostatic spin waves (MSWs) propagating in magnetized films of yttrium iron garnet (YIG) are an extremely interesting object for research, since the nonlinear effects that appear when intense MSWs propagate begin to manifest themselves at relatively small powers.[1,2] Surface MSWs are especially interesting in this regard, since the energy of the wave is concentrated at the film surface in this case, and the losses are minimal when the signal is generated and detected. When a single surface MSW propagates, an increase in the power of the wave does not cause amplitude modulation of the envelope of the magnetostatic potential to appear spontaneously, and the wave is modulationally stable in this case.[3] However, recent experiments[4] with simultaneous propagation of two surface MSWs of different frequencies show that modulation instability appears under definite conditions. This effect is observed in the form of side frequencies on the peaks corresponding to the carrier frequencies on the output spectral characteristic. The cause of the instability is that the dispersion dependence for the MSWs changes, because the second wave propagates in a medium perturbed by the first wave. A phenomenon similar to that observed was theoretically explained for the first time in Ref. 5, which discussed the combined propagation of two waves of different polarizations in a plasma. Similar effects when signals propagated in optical waveguides were subsequently detected.[6] The instability causes the envelope of the MSWs to be modulated, since the ever-present noise serves as an initial perturbation.[7] This phenomenon has become known as modulation instability. The modulation instability that arises when two modulationally stable waves propagate simultaneously, due to cross-phase modulation,[5] is usually caused by induced modulation instability.[6] The derivation of the conditions for the appearance of modulation instability is one of the main problems in studying such processes. To do this, it is important to obtain the dispersion dependence of the amplitude perturbations of the magnetostatic potential of the carrier signal. In studying such processes, it is usual to neglect dissipative effects because they are small at the distances under consideration.[2] The output MSWs are attenuated if the dissipative terms are included, but the qualitative picture of the wave propagation does not change. This paper discusses a model of simultaneous nondissipative propagation of two surface MSWs in a ferromagnetic film. The dispersion dependences are derived in Sec. 2 in terms of this model. The equations for the evolution of the amplitudes of the coupled waves are obtained in Sec. 3. After this, an analysis of these equations is given in Sec. 4 in order to derive the conditions for the appearance of modulation instability. Finally, in Sec. 5, the results are used to explain the experimental results of Ref. 4.

2. DERIVATION OF THE DISPERSION DEPENDENCE

Let us consider the propagation of two surface MSWs in a thin ferromagnetic film of thickness d, placed in a saturating external magnetic field $\mathbf{H}$ (Fig. 1).

Under these conditions, it is possible to introduce into the discussion a magnetostatic potential that describes the propagating waves and has in our case the form[8]

$$\psi = A[\exp(k_1 x) + \alpha \exp(-k_1 x)]\exp(ik_1 y) + B[\exp(k_2 x) + \beta \exp(-k_2 x)]\exp(ik_2 y) + \text{c.c.}, \quad (1)$$

where A and B are the amplitudes of the two MSWs, k_1 and k_2 are the wave numbers, and α and β are factors that depend on the film parameters and the external magnetic field and determine the propagation of the waves on some surface or the other.[9] The time dependence of the potential will be introduced later. It is easy to obtain the high-frequency components of magnetic field $\mathbf{h} = \nabla \psi$ from the given expression:

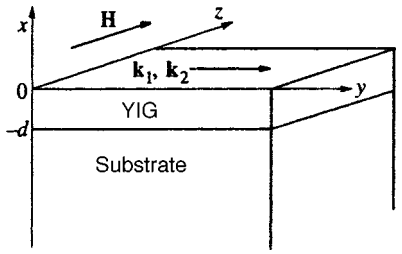

FIG. 1. Geometry of the problem. The MSWs propagate along the y axis, and the external magnetic field **H** is directed along the z axis.

$$h_x=\frac{\partial\psi}{\partial x}=Ak_1[\exp(k_1x)-\alpha\exp(-k_1x)]\exp(ik_1y)+Bk_2[\exp(k_2x)-\beta\exp(-k_2x)]\exp(ik_2y)+\text{c.c.}, \tag{2}$$

$$h_y=\frac{\partial\psi}{\partial y}=iAk_1[\exp(k_1x)+\alpha\exp(-k_1x)]\exp(ik_1y)+iBk_2[\exp(k_2x)+\beta\exp(-k_2x)]\exp(ik_2y)+\text{c.c.} \tag{3}$$

The high-frequency part of the magnetization is connected with **h**,

$$\mathbf{m}=\hat{\chi}\mathbf{h}, \tag{4}$$

by magnetic susceptibility tensor $\hat{\chi}$, which is written as

$$\hat{\chi}=\begin{pmatrix}\chi_1 & i\chi_2 & 0\\ -i\chi_2 & \chi_1 & 0\\ 0 & 0 & 1\end{pmatrix},\quad \mathbf{h}=\begin{pmatrix}h_x\\ h_y\\ h_z\end{pmatrix}. \tag{5}$$

As a result, we have

$$m_x=\chi_1\frac{\partial\psi}{\partial x}+i\chi_2\frac{\partial\psi}{\partial y},$$

$$m_y=-i\chi_2\frac{\partial\psi}{\partial x}+\chi_1\frac{\partial\psi}{\partial y}. \tag{6}$$

As will become clear later, we are interested only in the square of the modulus of the components of the magnetization. We consider $|\mathbf{m}|^2$ one component at a time:

$$|m_x|^2=\chi_1^2\left|\frac{\partial\psi}{\partial x}\right|^2+\chi_2^2\left|\frac{\partial\psi}{\partial y}\right|^2+i\chi_1\chi_2\left[\left(\frac{\partial\psi}{\partial y}\right)\left(\frac{\partial\psi}{\partial x}\right)^*-\left(\frac{\partial\psi}{\partial x}\right)\left(\frac{\partial\psi}{\partial y}\right)^*\right]. \tag{7}$$

Recalling that $\mathrm{Re}\,\psi=\psi$, we get

$$|m_x|^2=\chi_1^2\left|\frac{\partial\psi}{\partial x}\right|^2+\chi_2^2\left|\frac{\partial\psi}{\partial y}\right|^2. \tag{8}$$

Likewise, for m_y,

$$|m_y|^2=\chi_2^2\left|\frac{\partial\psi}{\partial x}\right|^2+\chi_1^2\left|\frac{\partial\psi}{\partial y}\right|^2. \tag{9}$$

Finally, we find

$$|\mathbf{m}|^2=|m_x|^2+|m_y|^2=(\chi_1^2+\chi_2^2)\left(\left|\frac{\partial\psi}{\partial x}\right|^2+\left|\frac{\partial\psi}{\partial y}\right|^2\right). \tag{10}$$

For the subsequent discussion, we need the quantity

$$\omega_M=4\pi\gamma\left[M_0-\frac{|m_x|^2+|m_y|^2}{2M_0}\right],\quad \omega_{M_0}=4\pi\gamma M_0, \tag{11}$$

which enters into the expression for the dispersion dependence[8] for a surface MSW:

$$\omega^2=\omega_H^2+\omega_H\omega_M+\frac{\omega_M^2}{4}[1-\exp(-2kd)]. \tag{12}$$

We introduce into the discussion

$$\omega_{0i}^2=\omega_H^2+\omega_H\omega_{M_0}+\frac{\omega_{M_0}^2}{4}[1-\exp(-2k_id)], \tag{13}$$

where $i=\{1,2\}$. Finally,

$$\omega_i^2=\omega_{0i}^2-\frac{\pi\gamma}{M_0}\{2\omega_H+\omega_{M_0}[1-\exp(-2k_id)]\}\times(|m_x|^2+|m_y|^2). \tag{14}$$

Simple but tedious algebraic formations can be used to obtain an expression for $|\partial\psi/\partial x|^2+|\partial\psi/\partial y|^2$, which contains terms proportional to $\exp(iky)$. In order to introduce the time dependence into the equations, it is necessary to make the substitution $\exp(iky)\to\exp[i(ky-\omega t)]$. This means that the terms containing t and y disappear after averaging over the period. These terms can play a role only if the waves are coherent. After this, we obtain

$$\left|\frac{\partial\psi}{\partial x}\right|^2+\left|\frac{\partial\psi}{\partial y}\right|^2=4\{|A|^2k_1^2[\exp(2k_1x)+|\alpha|^2\times\exp(-2k_1x)]+|B|^2k_2^2[\exp(2k_2x)+|\beta|^2\exp(-2k_2x)]\}. \tag{15}$$

We now have the following dispersion dependence:

$$\omega_i^2=\omega_{0i}^2-4\frac{\pi\gamma}{M_0}\{2\omega_H+\omega_{M_0}[1-\exp(-2k_id)]\}\times(\chi_1^2+\chi_2^2)\{|A|^2k_1^2[\exp(2k_1x)+|\alpha|^2\times\exp(-2k_1x)]+|B|^2k_2^2[\exp(2k_2x)+|\beta|^2\times\exp(-2k_2x)]\}. \tag{16}$$

The nonlinear dispersion Eqs. (16) for surface MSWs are derived in the limit of weak nonlinearity. Namely, nonlinear Eq. (11) describes the magnetization frequency, provided that the amplitude of the high-frequency magnetization is much less than the amplitude of the magnetization of the ferromagnetic film ($|m_x|,|m_y|\ll M_0$).

3. DERIVATION OF A SYSTEM OF EQUATIONS FOR THE EVOLUTION OF COUPLED WAVES

The dispersion dependence given by Eq. (16) can be written in general form as

$$G(k_1,k_2,\omega_1,\omega_2,|A|^2,|B|^2,|\alpha|^2,|\beta|^2)=0. \tag{17}$$

Let us introduce the dispersion dependence for an independently propagating wave:

$$G_{i0}=\omega_{0i}^2-\left\{\omega_H^2+\omega_H\omega_{M_0}+\frac{\omega_{M_0}^2}{4}[1-\exp(-2k_id)]\right\}. \tag{18}$$

We expand Eq. (16) up to terms corresponding to second order in amplitude:

$$\begin{aligned}G_i=&G_{i0}+\frac{\partial G_{i0}}{\partial\omega_i}\Delta\omega_i+\frac{\partial G_{i0}}{\partial k_i}\Delta k_i+\frac{\partial^2 G_{i0}}{\partial k_i\partial\omega_i}\Delta\omega_i\Delta k_i\\&+\frac{1}{2}\frac{\partial^2 G_{i0}}{\partial\omega_i^2}(\Delta\omega_i)^2+\frac{1}{2}\frac{\partial^2 G_{i0}}{\partial k_i^2}(\Delta k_i)^2+\frac{\partial G_i}{\partial|A|^2}|A|^2\\&+\frac{\partial G_i}{\partial|B|^2}|B|^2+\frac{\partial G_i}{\partial|\alpha|^2}|\alpha|^2+\frac{\partial G_i}{\partial|\beta|^2}|\beta|^2=0.\end{aligned} \tag{19}$$

Recalling that

$$\Delta\omega_i\leftrightarrow i\frac{\partial}{\partial t},\quad \Delta k_i\leftrightarrow -i\frac{\partial}{\partial y},\quad \frac{\partial^2}{\partial t^2}\leftrightarrow v_{gi}\frac{\partial^2}{\partial y},$$

we get

$$\begin{aligned}&i\frac{\partial A}{\partial t}+iv_{g1}\frac{\partial A}{\partial y}+\frac{1}{2}\beta_1\frac{\partial^2 A}{\partial y^2}=f_1(a^2|A|^2+b^2|B|^2)A,\\&i\frac{\partial B}{\partial t}+iv_{g2}\frac{\partial B}{\partial y}+\frac{1}{2}\beta_2\frac{\partial^2 B}{\partial y^2}=f_2(a^2|A|^2+b^2|B|^2)B,\end{aligned} \tag{20}$$

with

$$v_{gi}=2d\frac{\omega_{M0}^2}{4\omega_i}\exp(-2k_id), \tag{21}$$

$$\beta_i=-\frac{v_{gi}}{\omega_i}[v_{gi}+\omega_id], \tag{22}$$

$$\begin{aligned}f_i=&-\frac{\omega_{M_0}}{2M_0^2\omega_i}\{2\omega_H+\omega_{M_0}[1-\exp(-2k_id)]\}\\&\times(\chi_{1i}^2+\chi_{2i}^2),\end{aligned} \tag{23}$$

$$\begin{aligned}a^2&=k_1^2[\exp(2k_1x)+2|\alpha|^2\exp(-2k_1x)],\\b^2&=k_2^2[\exp(2k_2x)+2|\beta|^2\exp(-2k_2x)].\end{aligned} \tag{24}$$

Parameters a and b characterize the amplitude attenuation of the wave with distance from the surface. Equations (20) are in essence a system of equations of the type of the nonlinear Shrödinger equation and describe the evolution of the amplitudes of the coupled surface MSWs.

4. INVESTIGATION OF MODULATION INSTABILITY

Let us introduce the energy of the wave at depth x from the surface:

$$P_1=a^2|A|^2,\quad P_2=b^2|B|^2. \tag{25}$$

Then the steady-state equations have the form

$$iv_{g1}\frac{\partial A}{\partial y}+\frac{1}{2}\beta_1\frac{\partial^2 A}{\partial y^2}=f_1(P_1+P_2)A,$$

$$iv_{g2}\frac{\partial B}{\partial y}+\frac{1}{2}\beta_2\frac{\partial^2 B}{\partial y^2}=f_2(P_1+P_2)B. \tag{26}$$

Let us consider, for example, the first equation. We seek a solution in the form

$$A=C\exp(-i\alpha y). \tag{27}$$

After substitution, we find

$$\alpha_{1,2}=\frac{v_{g1}}{\beta_1}\left[1\mp\sqrt{1-\frac{2\beta_1f_1(P_1+P_2)}{v_{g1}^2}}\right]. \tag{28}$$

We have the following solution:

$$A=C_1\exp(-i\alpha_1y)+C_2\exp(-i\alpha_2y). \tag{29}$$

The amplitude close to the antenna is constant, and consequently the envelope is constant:

$$\begin{aligned}\frac{\partial A}{\partial y}=&C_1(-i\alpha_1)\exp(-i\alpha_1y)\\&+C_2(-i\alpha_2)\exp(-i\alpha_2y)=0.\end{aligned} \tag{30}$$

From this, we obtain the relationship between the amplitudes:

$$C_1=-\frac{\alpha_2}{\alpha_1}C_2. \tag{31}$$

Let us evaluate this relationship for a thin film ($k_id\ll 1$). In this case, the ratio under the radical in Eq. (28) equals

$$\epsilon_i=\frac{2\omega_{M_0}\omega_H}{M_0^2\omega_i^2}(\chi_{1i}^2+\chi_{2i}^2)(P_1+P_2). \tag{32}$$

We estimate the factors in Eq. (32) by the following approximations:

$$(\chi_{1i}^2+\chi_{2i}^2)\approx 3,\quad \frac{2\omega_{M_0}\omega_H}{\omega_i^2}\approx\frac{1}{2},$$

and consequently we get

$$\epsilon\approx 3(P_1+P_2)/M_0^2. \tag{33}$$

However, this ratio is small in the approximation considered here (in the experiment considered below, it equals 1/11). Consequently, $\epsilon\ll 1$. It follows from this that $C_1\gg C_2$. We shall neglect the quantity C_2 in the subsequent calculations. Then we can write

$$\alpha=\alpha_1=\frac{v_{g1}}{\beta_1}\left[1-\sqrt{1-\frac{2\beta_1f_1(P_1+P_2)}{v_{g1}^2}}\right]. \tag{34}$$

Finally,

$$A=\sqrt{P_1}e^{-i\alpha y}. \tag{35}$$

We now impose a perturbation on this solution:

$$A=[\sqrt{P_1}+\tilde{a}(y,t)]e^{-i\alpha y},\quad \tilde{a}\ll\sqrt{P_1}. \tag{36}$$

After substituting this expression into the equations, we carry out a similar procedure for B, and we linearize the resulting equations in terms of the perturbations:

$$i\frac{\partial\tilde{a}}{\partial t}+iv_{g1}\phi_1\frac{\partial\tilde{a}}{\partial y}+\frac{1}{2}\beta_1\frac{\partial^2\tilde{a}}{\partial y^2}=f_1[P_1(\tilde{a}+\tilde{a}^*)+\sqrt{P_1P_2}(\tilde{b}+\tilde{b}^*)],$$

$$i\frac{\partial\tilde{b}}{\partial t}+iv_{g2}\phi_2\frac{\partial\tilde{b}}{\partial y}+\frac{1}{2}\beta_2\frac{\partial^2\tilde{b}}{\partial y^2}=f_2[P_2(\tilde{b}+\tilde{b}^*)+\sqrt{P_1P_2}(\tilde{a}+\tilde{a}^*)], \quad (37)$$

where

$$\phi_i=v_{gi}\sqrt{1-\frac{2\beta_i f_i(P_1+P_2)}{v_{gi}^2}}. \quad (38)$$

We seek $\tilde{a}$ and $\tilde{b}$ in the following form:

$$\tilde{a}=u_1\cos[K_1(y-\phi_1 t)-\Omega t]+iv_1\times\sin[K_1(y-\phi_1 t)-\Omega t],$$

$$\tilde{b}=u_2\cos[K_2(y-\phi_2 t)-\Omega t]+iv_2\times\sin[K_2(y-\phi_2 t)-\Omega t], \quad (39)$$

where K_1 and K_2 are the wave numbers of the amplitude perturbations, while Ω is their frequency. After substitution, we must set the real and imaginary parts equal to zero. We obtain a system of linear equations with the matrix

$$M=\begin{pmatrix} m_1 & q_1 & l_1 & 0 \\ n_1 & p_1 & 0 & 0 \\ l_2 & 0 & m_2 & q_2 \\ 0 & 0 & n_2 & p_2 \end{pmatrix}, \quad (40)$$

where

$$m_i=-\frac{1}{2}\beta_i K_i^2\cos[K_i(y-\phi_i t)-\Omega t]-2f_i P_i\cos[K_i(y-\phi_i t)-\Omega t],$$

$$q_i=(K_i\phi_i+\Omega)\cos[K_i(y-\phi_i t)-\Omega t]-\phi_i K_i\cos[K_i(y-\phi_i t)-\Omega t],$$

$$l_i=-2f_i\sqrt{P_1P_2}\cos[K_{2-i}(y-\phi_{2-i}t)-\Omega t],$$

$$n_i=-(K_i\phi_i+\Omega)\sin(K_i(y-\phi_i t)-\Omega t)+\phi_i K_i\sin[K_i(y-\phi_i t)-\Omega t],$$

$$p_i=\frac{1}{2}\beta_i K_i^2\sin[K_i(y-\phi_i t)-\Omega t]. \quad (41)$$

For a nontrivial solution to exist in the system, it is necessary that the determinant of this matrix equal zero:

$$\det M=0, \quad (42)$$

or

$$(n_1q_1-p_1m_1)(n_2q_2-p_2m_2)-p_1p_2l_1l_2=0.$$

We obtain the following equation for Ω:

$$(\Omega^2-c_1^2)(\Omega^2-c_2^2)=\zeta, \quad (43)$$

where

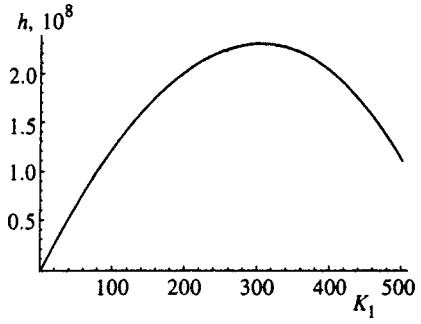

FIG. 2. Dependence of gain h on K_1 for fixed K_2.

$$c_i=\frac{1}{2}\beta_i K_i^2, \quad \zeta=\beta_1\beta_2 f_1 f_2 P_1P_2K_1^2K_2^2. \quad (44)$$

It is easy to express Ω^2 as

$$\Omega^2=\frac{c_1^2+c_2^2\pm\sqrt{(c_1^2+c_2^2)^2-4(c_1^2c_2^2-\zeta)}}{2}. \quad (45)$$

We find the condition for which $\Omega^2<0$:

$$\zeta>c_1^2c_2^2. \quad (46)$$

This is the condition for modulation instability to appear. In this case, it can be seen from Eqs. (44) and (46) that both waves are modulationally unstable regardless of the signs of the nonlinearity coefficients given in Eq. (23). These results also agree with the results published earlier in Refs. 10 and 11 concerning the nonlinear interaction between spin and acoustic waves and between spin and electromagnetic waves.

5. ANALYSIS OF EXPERIMENT

In Ref. 4, two MSWs with frequencies $\omega_1=6.55$ GHz and $\omega_2=6.75$ GHz, which correspond to wave numbers $k_1=52.97\,\mathrm{cm}^{-1}$ and $k_2=379.6\,\mathrm{cm}^{-1}$, were generated in a film of yttrium iron garnet of thickness $d=1.15\times10^{-3}$ cm. The film in this experiment had a saturation magnetization of $M_0=135.6$ G and was located in an external magnetic field of $H_0=1627$ Oe, with $\gamma=2.8$ MHz/Oe. The signal from the film was fed to a spectrum analyzer. The mistuning between the nearest side band and the carrier peak equalled 1.4 MHz, but the side band is caused by the interaction with another wave, and it is consequently necessary to consider the spacing not between the nearest peaks but between the farthest. This makes it possible to explain the presence of the asym-

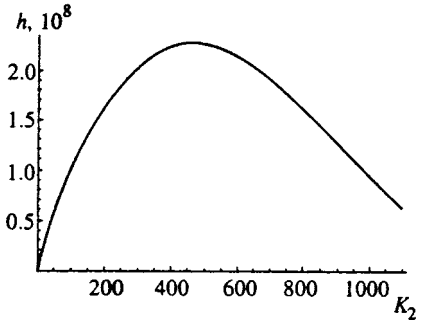

FIG. 3. Dependence of gain h on K_2 for fixed K_1.

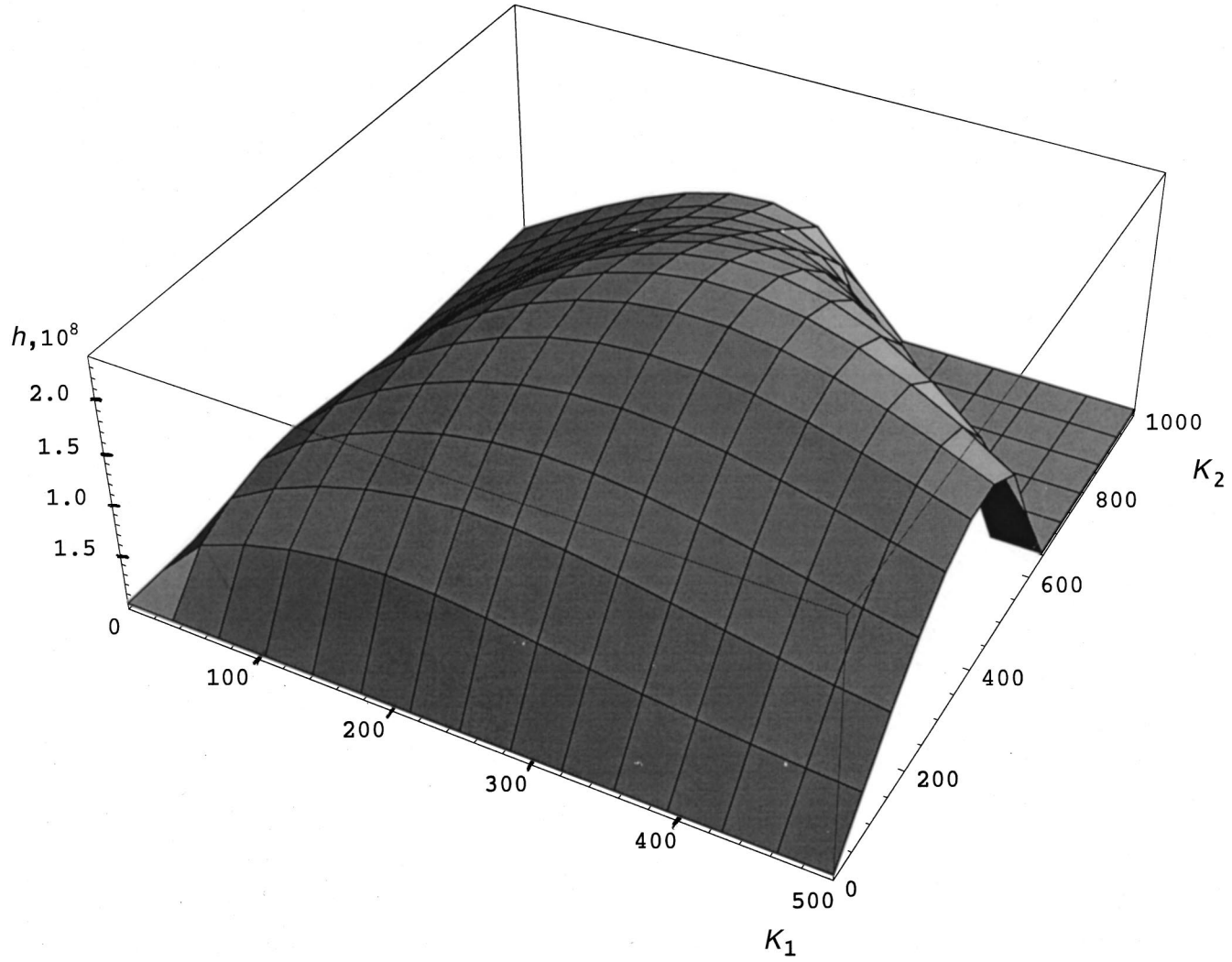

FIG. 4. Dependence of gain h on K_1 and K_2.

metry of the side bands as the result of the interaction of already modulated waves. Then the modulation frequency $\Omega=\Delta\omega$ is 201.4 MHz. Keeping in mind that the gain must have a maximum at this frequency, we get a system of equations for K_1 and K_2. Solving it, we find

$$K_1=283.3\ \text{cm}^{-1},\quad K_2=444.4\ \text{cm}^{-1}. \tag{47}$$

The power of the wave can be obtained from[3]

$$W\kappa\simeq\frac{1}{16\pi}Ld^2\omega P, \tag{48}$$

where $L=0.3$ cm is the length of the antenna, while $\kappa=0.25$ is a factor that characterizes the part of the supplied power W that goes into the generation of MSWs. From this we get

$$k=\frac{1}{16\pi}Ld^2=7.89\times10^{-9}\ \text{cm}^3,$$

$$P=\frac{W\kappa}{k\omega},\quad P_1=134\ \text{Oe}^2,\quad P_2=45\ \text{Oe}^2. \tag{49}$$

Let us check whether the conditions $\zeta>c_1^2c_2^2$ for modulation instability are satisfied:

$$4f_1f_2P_1P_2>\frac{1}{2}\beta_1K_1^2\cdot\frac{1}{2}\beta_2K_2^2,$$

$$\frac{16f_1f_2}{\beta_1\beta_2}P_1P_2>K_1^2K_2^2,$$

$$\kappa^2 1.38\times10^{12}\ \text{cm}^{-4}>K_1^2K_2^2,$$

$$8.63>1.59. \tag{50}$$

As can be seen, the condition for cross-phase modulation is satisfied. In order to understand the spectral content of the received signal, it is necessary to know how the gain of the modulation perturbations depends on K_1 and K_2. The gain of the perturbation amplitudes equals

$$h(K_1,K_2)=2\,\text{Im}(\Omega)=\sqrt{2[\sqrt{(c_1^2+c_2^2)^2+4(\zeta-c_1^2c_2^2)}-(c_1^2+c_2^2)]}. \tag{51}$$

Figure 2 shows a graph of the $h(K_1,K_2)$ dependence with K_2 fixed. Figure 3 shows the $h(K_1,K_2)$ dependence with K_1 fixed. The $h(K_1,K_2)$ surface is shown in Fig. 4.

It can be seen from the curves in the figures that the gain has a maximum at definite values of K_1 and K_2, corresponding to the frequencies Ω_1 and Ω_2 of the modulation instability observed in experiment.

6. CONCLUSION

Instability with respect to the amplitude of intense travelling surface MSWs in ferromagnetic films has been theoretically treated. Surface MSWs in the case of propagation of only one wave are modulationally stable even when the power of the wave varies within wide limits. When two waves simultaneously propagate in a ferromagnetic film, they become modulationally unstable when they reach a certain threshold power. Such instability cannot be explained in terms of a model of parametric instability. A model of the phase modulation of two intense waves is proposed here that explains the appearance of modulation instability of the waves. Calculations are carried out for the parameters of the waves, in particular, the threshold powers required for the phenomena discussed here. An explanation is given for the experimental results of Ref. 4, in which cross-phase modulation of surface MSWs was observed for the first time. Qualitative agreement is obtained between the theory developed here and the experimental data. Since this paper uses a model of nondissipative propagation of surface MSWs, while quenching of the waves plays a substantial role in the example studied, which involved a thick film, it is problematical to obtain quantitative agreement with experiment. Further development of the theory, taking into account wave dissipation, and additional experimental work should clarify the details of the process of cross-phase modulation of magnetostatic spin waves.

This work was carried out with the partial support of the Russian Fund for Fundamental Research (Project 99-02-17660) and the Ministry of Science and Education of the Russian Federation (the Atomic Surface Structures Program).

*)Current address: L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 117334, Moscow. E-mail: kao@landau.ac.ru

[1] A. G. Gurevich and G. A. Melkov, *Magnetic Vibrations and Waves* (Moscow, Nauka, 1994).
[2] D. D. Stancil, *Theory of Magnetostatic Waves* (Springer-Verlag, New York, 1993).
[3] A. K. Zvezdin and A. F. Popkov, Zh. Éksp. Teor. Fiz. **84**, 606, (1983) [Sov. Phys. JETP **57**, 350 (1983)].
[4] J. W. Boyle, S. A. Nikitov, A. D. Boardman, and K. Xie, J. Magn. Magn. Mater. **173**, 241 (1997).
[5] A. L. Berkhoer and V. E. Zakharov, Zh. Éksp. Teor. Fiz. **58**, 903 (1970) [Sov. Phys. JETP **31**, 486 (1970)].
[6] G. P. Agrawal, Phys. Rev. Lett. **59**, 880 (1987).
[7] L. L. Savchenko, S. A. Nikitov, A. F. Popkov, and M. V. Chetkin, Zh. Éksp. Teor. Fiz. **114**, 628 (1998) [JETP **87**, 342 (1998)].
[8] R. W. Damon and J. R. Eshbach, J. Phys. Chem. Solids **19**, 308 (1961).
[9] A. D. Boardman, Q. Wang, S. A. Nikitov, J. Shen, W. Chen, D. Mills, and J. S. Bao, IEEE Trans. Magn. **30**, 14 (1994).
[10] Yu. V. Gulyaev and S. A. Nikitov, Fiz. Tverd. Tela **26**, 2620 (1984) [Sov. Phys. Solid State **26**, 1589 (1984)].
[11] Yu. V. Gulyaev and S. A. Nikitov, Fiz. Tverd. Tela **27**, 2710 (1985) [Sov. Phys. Solid State **27**, 1624 (1985)].

Translated by W. J. Manthey

Multiphonon optical transitions in size-limited systems in a magnetic field

É. P. Sinyavskiĭ*) and E. I. Grebenshchikova

Institute of Applied Physics, Academy of Science of Moldova, 277028 Kishinev, Moldova
(Submitted 31 March 1999)
Zh. Éksp. Teor. Fiz. **116**, 2069–2078 (December 1999)

A study has been performed of optical multiphonon transitions in undoped size-limited systems in a magnetic field aligned with the spatial quantization axis. A theory is proposed which allows one to describe the half-width of the luminescence curve for isolated quantum wells and investigate the frequency and temperature dependence of the light absorption coefficient in the long-wavelength region. © *1999 American Institute of Physics.* [S1063-7761(99)01512-7]

1. INTRODUCTION

In a longitudinal magnetic field oriented perpendicular to the surface of a quantum well, the spectrum of a free electron is completely quantized (quasi-zero-dimensional) and for rectangular quantum wells it is given by

$$E_{n\nu}=\hbar\omega_c\left(n+\frac{1}{2}\right)+\varepsilon_0\nu^2, \quad \varepsilon_0=\frac{\hbar^2\pi^2}{2m_ca^2}.$$

Here ω_c is the cyclotron frequency, ε_0 is the size-quantization step in the quantum well, m_c is the effective electron mass, a is the width of the size-limited system, ν is the number of the size-quantized level, and n is the number of the Landau level.

The appearance of discrete levels should have a substantial effect on the optical properties of quantum wells. In the present paper we present a study of the optical properties of size-limited systems in a longitudinal magnetic field with multiphonon effects taken into account. The final results are valid for quantum wells of different shape (rectangular, parabolic), and also for heterostructures. To describe multiphonon optical processes in impurity systems, we use the model of shifted adiabatic potentials.[1] However, to study electron–vibrational optical transitions between discrete states of free holes (in the absence of a magnetic field B this is the valence band) and quasi-one-dimensional states of free electrons (for $B=0$ this is the conduction band) this model is inapplicable since the minima of adiabatic potentials for free carriers are not shifted. This is because the diagonal matrix elements of the electron–phonon interaction (and likewise the hole–phonon interaction) in the free-carrier wave functions are extremely small. In this case, a description of optical transitions with multiphonon effects taken into account is best constructed in the language of quasilevels,[2] as is done in the case of multiphonon interband transitions.[3] In what follows, we will consider strong, quantized magnetic fields, where the Coulomb interaction of the electron with a hole is small in comparison with the distance between the transverse quantization levels. In this case the internal motion of an electron–hole pair is finite, but a free state of the electron and hole does not, strictly speaking, exist.[4] As experimental studies on photoluminescence in quantum wells has shown, the binding energy of an exciton for $B>10$ T in InGaAs/GaAs (Ref. 5) and in GaAs/AlGaAs (Ref. 6) is proportional to B and, consequently, it is possible to use the above-indicated approximation. A detailed discussion and the criteria of this approximation are given in Ref. 7. In what follows we assume that the electron and hole interact with the phonons independently, as is done in the theory of large-radius excitons.[8] When an electron–hole pair is excited by light, the momentum of the exciton is equal to the momentum of the electromagnetic wave and is very small;[9] therefore, we will neglect the exciton bands arising in quasi-two-dimensional semiconductors in a strong magnetic field.[4]

The theory developed below allows one in a number of cases to describe the half-width of the luminescence curve for isolated quantum wells and in turn to study the frequency and temperature dependence of the light absorption coefficient in the long-wavelength region.

2. STATEMENT OF THE PROBLEM. GENERAL RELATIONS

In an intrinsic size-limited semiconductor the Hamiltonian of the system of electrons and holes interacting with phonons in a uniform magnetic field has the form

$$\hat{H}=\hat{H}_0+\hat{V}, \tag{1}$$

$$\hat{H}_0=\sum_\beta(\varepsilon_\beta^{(c)}-\xi)a_\beta^+a_\beta+\sum_\beta(\varepsilon_\beta^{(v)}-\xi_1)\alpha_\beta^+\alpha_\beta + \sum_{\mathbf{q}}\hbar\omega_q b_{\mathbf{q}}^+b_{\mathbf{q}}, \tag{2}$$

$$\hat{V}=\sum_{\mathbf{q},\beta,\beta_1}C_{\mathbf{q}}^{(c)}I_{\beta\beta_1}(\mathbf{q})(b_{\mathbf{q}}+b_{-\mathbf{q}}^+)a_\beta^+a_{\beta_1} + \sum_{\mathbf{q},\beta,\beta_1}C_{\mathbf{q}}^{(v)}I_{\beta\beta_1}(\mathbf{q})(b_{\mathbf{q}}+b_{-\mathbf{q}}^+)\alpha_\beta^+\alpha_{\beta_1}. \tag{3}$$

Here $\varepsilon_\beta^{(c)}$ and $\varepsilon_\beta^{(v)}$ are the energy of the electrons (holes) in a quantum well in a longitudinal magnetic field (see the Introduction), ξ is the chemical potential of the electron, $\xi_1=-\xi+E_g$, E_g is the band gap of the semiconductor, $\hbar\omega_q$ is the energy of a phonon with wave vector $\mathbf{q}$; a_β^+, a_β, α_β^+, α_β, $b_{\mathbf{q}}^+$, and $b_{\mathbf{q}}$ are creation and annihilation operators respectively for the electrons, holes, and phonons. $C_{\mathbf{q}}^{(c)}$ ($C_{\mathbf{q}}^{(v)}$) is the coefficient function describing the interaction of an

electron (hole) with lattice vibrations, $I_{\beta\beta_1}(\mathbf{q})=\langle\beta|\exp(i\mathbf{q}\times\mathbf{r})|\beta_1\rangle$ is the matrix element of the operator $\exp(i\mathbf{q}\cdot\mathbf{r})$ in the wave functions $|\beta\rangle$ of the free carriers of the size-limited system in a magnetic field, $\beta=(n,\nu,K_x)$, where K_x is a component of the carrier wave vector, n is the number of the Landau level, ν is the number of the size-quantized level. The Hamiltonian (1) does not include terms associated with the nonadiabaticity operator since we neglect nonradiative transitions in what follows.[3]

The absorption coefficient for light with frequency Ω associated with a transition of an electron from a completely discrete state (1) to a quasi-zero-dimensional state (2) is given by the Kubo formula[10]

$$K(\Omega)=\frac{4\pi e^2}{Vn_0c\hbar\Omega}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2\sum_{\beta\beta_1}\int_{-\infty}^{\infty}dt\exp(i\Omega t)\times\langle\alpha_\beta(t)a_\beta(t)\alpha^+_{\beta_1}(t)a^+_{\beta_1}(t)\rangle, \tag{4}$$

$$\hat{A}(t)=\exp\left(\frac{it}{\hbar}\hat{H}\right)\hat{A}\exp\left(-\frac{it}{\hbar}\hat{H}\right), \tag{5}$$

where V is the volume of the size-limited system, n_0 is the index of refraction, c is the speed of light, $\mathbf{P}_{cv}$ is the matrix element of the momentum operator in Bloch functions, m_0 is the mass of a free electron, $\boldsymbol{\xi}_0$ is the polarization vector of the electromagnetic wave, and the angular brackets denote averaging over the Hamiltonian (1).

The equation of motion for the operator $a_\beta(t)$ according to Eq. (5) can be written as

$$\dot{a}_\beta(t)=-\frac{i}{\hbar}\left\{(\varepsilon^{(c)}_\beta-\xi)a_\beta(t)+\sum_{\mathbf{q},\beta_1}C^{(c)}_{\mathbf{q}}\langle\beta|\times\exp(i\mathbf{q}\cdot\mathbf{r})|\beta_1\rangle(b_{\mathbf{q}}\exp(-i\omega_q t)+b^+_{-\mathbf{q}}\exp(i\omega_q t)a_{\beta_1}(t))\right\}. \tag{6}$$

Formula (6) neglects the influence of carriers on the phonon spectrum, i.e., it assumes that

$$b_{\mathbf{q}}(t)\simeq b_{\mathbf{q}}\exp(-i\omega_q t),\quad b^+_{\mathbf{q}}(t)\simeq b^+_{\mathbf{q}}\exp(i\omega_q t).$$

This approximation is valid for nondegenerate semiconductors since corrections to the free-phonon spectrum depend on the polarization operator, which to lowest order in the electron–phonon interaction is proportional to the concentration of charged particles.

For rectangular quantum wells of width a (the magnetic field is oriented along the spatial quantization axis)

$$\langle\beta|\exp(i\mathbf{q}\cdot\mathbf{r})|\beta_1\rangle=i\frac{4\pi^2\nu\nu_1(q_za)}{(q_za)^2-\pi^2(\nu+\nu_1)^2}\frac{1}{(q_za)^2-\pi^2(\nu-\nu_1)^2}\times\left[\frac{2^{n_1}n_1!}{2_{n_1}n!}\right]^{1/2}[\exp(iq_za)(-1)^{\nu+\nu_1}-1]\times\left[\exp\left\{-\frac{1}{4}R^2(q_x^2+q_y^2)+iq_y(K_x+K_{x_1})\frac{1}{2}R^2\right\}\right]\times\left[\frac{R}{2}(q_x-iq_y)\right]^{n-n_1}L^{n-n_1}_{n_1}\left(\frac{R^2}{2}(q_x^2+q_y^2)\right)\delta_{K_x-K_{x_1},q_x}\quad(n\geqslant n_1). \tag{7}$$

Here $R^2=\hbar/m_c\omega_c$, $L^{n-n_1}_{n_1}(z)$ are associated Laguerre polynomials.

$$V_{n\nu}(q)=i\frac{4\pi^2\nu^2}{(q_za)^2-(2\pi\nu)^2}\frac{\exp(iq_za)-1}{q_za}\times\exp\left[-\frac{R^2}{4}(q_x^2+q_y^2)\right]L_n\left[\frac{R^2}{2}(q_x^2+q_y^2)\right], \tag{8}$$

$\hat{P}_x=-i\hbar\partial/\partial x$ is the momentum operator, $|K_x\rangle$ is the wave function of a free electron along the x axis.

The terms with $n\neq n_1$, $\nu\neq\nu_1$ in expression (7), as will be shown below, make an insignificant contribution to optical multiphonon processes. Substituting expression (8) into Eq. (6) leads to the following equation of motion for $a_\beta(t)$:

$$\dot{a}_\beta(t)=-\frac{i}{\hbar}\left\{a_\beta(t)(\varepsilon^{(c)}_\beta-\xi)+\sum_{\mathbf{q},K_{x_1}}C^{(c)}_q V_{n\nu}(q)\times\langle K_x|\exp(i\hat{P})|K_{x_1}\rangle(b_{\mathbf{q}}\exp(-i\omega_q t)+b^+_{-\mathbf{q}}\exp(i\omega_q t))a_{n\nu K_{x_1}}(t)\right\}. \tag{9}$$

The solution of Eq. (9) has the form

$$a_{n\nu K_x}(t)=\sum_{K_{x_1}}\langle K_x|\exp\frac{it\hat{H}_f}{\hbar}\times\exp\left(-\frac{it(\hat{H}_f+W^{(c)}_{n\nu})}{\hbar}\right)|K_{x_1}\rangle\times\exp\left(-\frac{it}{\hbar}(\varepsilon^{(c)}_{n\nu K_x}-\xi)\right)a_{n\nu K_{x_1}}. \tag{10}$$

Here we have introduced the following notation:

$$\hat{H}_f=\sum_q\hbar\omega_q b^+_{\mathbf{q}}b_{\mathbf{q}},$$

$$W^{(c)}_{n\nu}=\sum_{\mathbf{q}}C^{(c)}_{\mathbf{q}}V_{n\nu}(q)\exp(i\hat{P})(b_{\mathbf{q}}+b^+_{-\mathbf{q}}),$$

$$\hat{P}=q_xx+\frac{1}{\hbar}q_yR^2\hat{P}_x, \tag{11}$$

We can calculate $\alpha_{n\nu K_x}(t)$ analogously.

If we substitute the values of $a_{n\nu K_x}$ and $\alpha_{n\nu K_x}$ into expression (4) and take account that for nondegenerate semiconductors

$$\langle a^+_\beta a_\beta\rangle=n^{(c)}_\beta\ll1,\quad\langle\alpha^+_\beta\alpha_\beta\rangle=n^{(v)}_\beta\ll1,$$

the light absorption coefficient takes the following form:

$$K(\Omega)=\frac{4\pi e^2}{Vcn_0\hbar\Omega}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2\sum_{n,\nu,K_x,K_{x_1}}\int_{-\infty}^{\infty}dt$$
$$\times\exp\left\{\frac{it}{\hbar}(\hbar\Omega-\varepsilon^{(c)}_{n\nu K_x}-\varepsilon^{(\nu)}_{n\nu K_x}-E_g)\right\}$$
$$\times\left\langle\langle K_x|\exp\left(\frac{it}{\hbar}H_f\right)\exp\left(-\frac{it}{\hbar}(H_f+W^{(c)}_{n\nu})\right)|K_{x_1}\rangle\langle K_x|\exp\left(\frac{it}{\hbar}H_f\right)\right.$$
$$\left.\times\exp\left(-\frac{it}{\hbar}(H_f+W^{(\nu)}_{n\nu})\right)|K_{x_1}\rangle\right\rangle_{\text{phon}}. \qquad (12)$$

In expression (12) the average $\langle\ldots\rangle_{\text{phon}}$ is taken over the system of free phonons since the influence of the interaction of carriers with vibrations on the phonon spectrum is neglected. The average in expression (12) can be found by the usual methods of the theory of multiphonon transitions[1] using, for example, the algebra of Bose operators.[11] As a result, for the light absorption coefficient we obtain the expression

$$K(\Omega)=\frac{2e^2}{acn_0R^2\Omega}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2\sum_{n,\nu}\int_{-\infty}^{\infty}dt$$
$$\times\exp\left(it\Omega-\frac{it}{\hbar}\varepsilon_{n\nu}\right)\exp(-g_{n\nu}(t)). \qquad (13)$$

Here

$$g_{n\nu}(t)=\sum_{\mathbf{q}}\frac{1}{(\hbar\omega_q)^2}(|C^{(c)}_{\mathbf{q}}|^2+|C^{(\nu)}_{\mathbf{q}}|^2)|V_{n\nu}(q)|^2$$
$$\times\{it\omega_q+(2N_q+1)-(2N_q+1)$$
$$\times\cos(\omega_q t)-i\sin(\omega_q t)\},$$
$$\varepsilon_{n\nu}=\hbar\omega^*_c\left(n+\frac{1}{2}\right)+\varepsilon^*_0\nu^2+E_g,\quad \hbar\omega^*_c=\frac{\hbar eH}{\mu c},$$
$$\varepsilon^*_0=\frac{\hbar^2\pi^2}{2a^2\mu},\quad \frac{1}{\mu}=\frac{1}{m_c}+\frac{1}{m_\nu}, \qquad (14)$$

$N_q=[\exp(\hbar\omega_q/k_0T)-1]^{-1}$ is the distribution function of equilibrium phonons at the temperature T.

As follows directly from Eq. (14), the average over the system of free phonons is in fact taken independently for the electrons and the holes. Terms of the type $\Sigma_{\mathbf{q}}C^{(\nu)}_{\mathbf{q}}C^{(c)}_{\mathbf{q}}|V_{n\nu}(q)|^2$ make extremely small corrections ($\sim 1/L_x$) to the light absorption coefficient.

The spectral radiation intensity is related in a simply way to the transition probability per unit time[12] and is given by

$$\Phi(\Omega)=\frac{2\Omega^2e^2n_0}{\pi\hbar Vc^3}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2\sum_{\beta\beta'}\int_{-\infty}^{\infty}dt\exp(-it\Omega)$$
$$\times\langle a^+_\beta(t)\alpha^+_\beta(t)a_{\beta'}\alpha_{\beta'}\rangle. \qquad (15)$$

The average in expression (15) is calculated in the same way as for the light absorption coefficient. As a result, we obtain

$$\Phi(\Omega)=\frac{\Omega^2e^2n_0}{a\pi^2c^3\hbar R^2}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2\sum_{n\nu}n^{(c)}_{n\nu}n^{(\nu)}_{n\nu}\int_{-\infty}^{\infty}dt$$
$$\times\exp(-it\Omega)\exp\left(\frac{it}{\hbar}\varepsilon_{n\nu}\right)\exp(-g_{n\nu}(t)), \qquad (16)$$

where $n^{(c)}_{n\nu}=\langle a^+_\beta a_\beta\rangle$ and $n^{(\nu)}_{n\nu}=\langle\alpha^+_\beta\alpha_\beta\rangle$ are the distribution functions for the electrons and the holes, respectively. For rectangular quantum wells neglecting the polaron effect

$$n^{(c)}_{n\nu}=\frac{n_e\sinh(\hbar\omega_c/2k_0T)a2\pi R^2}{\sum_\nu\exp(-\varepsilon_0\nu^2/k_0T)}$$
$$\times\exp\left\{-\frac{1}{k_0T}\left[\left(n+\frac{1}{2}\right)\hbar\omega_c+\varepsilon_0\nu^2\right]\right\}, \qquad (17)$$

where n_e is the electron density. An analogous expression follows for $n^{(\nu)}_{n\nu}$.

3. DISCUSSION OF RESULTS

Relations (13) and (16) describe processes of absorption and emission of an electromagnetic wave in intrinsic size-limited systems in a longitudinal magnetic field with allowance for multiphonon effects. For carriers interacting with optical vibrations of frequency ω_0 (we neglect the small dispersion) we can represent expression (14) in the form

$$g_{n\nu}(t)=a_{n\nu}\{it\omega_0+(2N+1)\cos(\omega_0t-\varphi)\}. \qquad (18)$$

Here

$$a_{n\nu}=\sum_{\mathbf{q}}[|C^{(c)}_{\mathbf{q}}|^2+|C^{(\nu)}_{\mathbf{q}}|^2]\frac{|V_{n\nu}(q)|^2}{(\hbar\omega_0)^2},$$
$$\tan\varphi=\frac{i}{2N+1},\quad N=\frac{1}{\exp(\hbar\omega_0/k_0T)-1}. \qquad (19)$$

Let us consider the region of temperatures ($3\varepsilon^*_0/k_0T>1$, $\hbar\omega_c/k_0T>1$) where the electrons are found in the lowest completely discrete state (2) ($n=0$, $\nu=1$). If we invoke the relation[13]

$$\exp(z\cos\varphi)=\sum_{m=-\infty}^{\infty}I_m(z)\exp(im\varphi)$$

($I_m(z)$ are the modified Bessel functions), then the spectral radiation intensity can be written as follows:

$$\Phi(\Omega)=\frac{2\pi e^2n_0aR^2\Omega^2}{c^3}n_en_h\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2$$
$$\times\exp[-a_{01}(2N+1)]\sum_{m=-\infty}^{\infty}I_m(z)$$
$$\times\left[\frac{1+N}{N}\right]^{m/2}\delta\{\hbar\Omega-\tilde{\varepsilon}_g+m\hbar\omega_0\}, \qquad (20)$$

$$z=2a_0\sqrt{N(N+1)}, \quad \tilde{\varepsilon}_g=E_g+\frac{1}{2}\hbar\omega_c^*+\varepsilon_0^*-\hbar\omega_0 a_{01},$$

where n_h is the hole density.

As follows directly from Eq. (20), for $z<1$ the frequency dependence of the luminescence is given by a narrow δ-shaped curve ($m=0$) with phonon satellites ($m\neq 0$) set off from one another by the distance $\hbar\omega_0$. If we take into account the nonstationarity of the electron states, the luminescence lines are described by a Lorentzian. The half-width of the emission lines, determined by inelastic scattering on acoustic vibrations (single-phonon transitions between Landau levels), has the form

$$\gamma=\frac{(2\pi)^3 k_0 T E_c^2 m_c}{\rho w^2 \hbar^2 a}\left(\frac{w}{a\omega_c}\right)^5. \tag{21}$$

Here ρ is the density of the semiconductor quantum well, E_c is the constant of the deformation potential for the electron, and w is the speed of sound. For typical parameters of a GaAs/AlGaAs quantum well ($\rho=5.4\,\mathrm{g/cm^3}$, $w=2\times10^5$ cm/s, and $E_c=9$ eV) for $T=100$ K, $a=50$ Å, and $\hbar\omega_c=10^{-2}$ eV we have for the half-width $\gamma\sim 6\times10^{-4}$ meV. However, in strong magnetic fields the half-width Δ of the photoluminescence line, as experimental studies have shown, reaches several meV and its shape is approximated by a Gaussian.[6] For example, in the quasi-two-dimensional systems InGaAs/GaAs $\Delta=5$ meV ($B<11.8$ T, Ref. 5), InP/$In_{0.53}Ga_{0.47}As$ $\Delta\leqslant 8$ meV ($B=4$ T, Ref. 14), InAs/$InAs_{0.09}Sb_{0.91}$ $\Delta\leqslant 7$ meV ($B=7.6$ T, Ref. 15), GaAs/$Al_xGa_{1-x}As$ $\Delta\approx 3$ meV ($B=9$ T, Ref. 16).

In a semiclassical description of vibrations of the crystal lattice $g_{n\nu}(t)$ can be expanded in its argument to terms $\sim t^2$ inclusive:

$$g_{n\nu}(t)\approx\frac{B_{n\nu}t^2}{2},$$

$$B_{n\nu}=\sum_{\mathbf{q}}\frac{1}{\hbar^2}[(C_{\mathbf{q}}^{(c)})^2+(C_{\mathbf{q}}^{(v)})^2]|V_{n\nu}(q)|^2(2N_q+1). \tag{22}$$

Substituting expression (22) into formula (16) ($n=0$, $\nu=1$) leads to the following expression for the spectral radiation intensity:

$$\Phi(\Omega)=\frac{n_0 e^2 a R^2\Omega^2}{c^3}n_e n_h\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|\sqrt{\frac{2\pi}{\Delta}} \times\exp\left\{-\frac{(\hbar\Omega-E_g-\hbar\omega_c^*/2-\varepsilon_0^*)^2}{2\Delta}\right\}, \tag{23}$$

$$\Delta=\hbar^2 B_{01}.$$

According to formula (23), the frequency dependence $\Phi(\Omega)$ is described by a Gaussian curve with its maximum at $\hbar\Omega_m=E_g+\hbar\omega_c^*/2+\varepsilon_0^*$ and half-width

$$\delta\Omega=2\sqrt{2\Delta\ln 2}. \tag{24}$$

If the frequency dependence of the photoluminescence is nearly Gaussian, then the position of the maximum, and the half-width, can be calculated using the method of moments,[17] which, in particular, allows one to formulate a criterion for the semiclassical approximation.

For carriers interacting with long-wavelength acoustic vibrations ($\omega_q=wq$) for $N_q\simeq k_0T/\hbar\omega_q$, the half-width of the emission line according to formulas (24) and (22) is given by

$$\delta\Omega=2\sqrt{k_0 T a_0\ln 2}, \quad a_0=\frac{3(E_c^2+E_v^2)}{\pi\rho w^2 R^2 a}, \tag{25}$$

where E_v is the constant of the deformation potential for the holes.

Note that the criterion for use of the ''high-temperature'' approximation in the calculation of formula (25) is met for $k_0T\gg\hbar w/R\sqrt{6}$ (for $\hbar\omega_c\simeq 10^{-2}$ eV, $T\gg 0.5$ K).

For typical parameters of GaAs/$Al_xGa_{1-x}As$ quantum wells ($E_v=7$ eV) for $a=50$ Å, $T=10$ K, and $B=10$ T we have $\delta\Omega\simeq 2$ meV, which agrees in order of magnitude with the experimental results of Ref. 16. For carriers interacting with optical dispersionless vibrations of frequency ω_0 the heat liberation parameter a_{01} according to formulas (19) is given by

$$a_{10}=\tilde{a}_0 I(\tilde{\xi}_0), \tag{26}$$

$$\tilde{a}_0=\frac{e^2 c_0}{a\hbar\omega_0},$$

$$I(\tilde{\xi})=\int_0^\infty d\tau\exp(-\tau)\left\{\frac{1}{2(\tau+\tilde{\xi}4\pi^2)}+\frac{1}{\tau} -\frac{(\tilde{\xi}4\pi^2)^2}{\tau(\tau+4\pi^2\tilde{\xi})^2}\sqrt{\frac{\tilde{\xi}}{\tau}}\left[1-\exp\left(-\sqrt{\frac{\tau}{\tilde{\xi}}}\right)\right]\right\},$$

$$\tilde{\xi}=\left(\frac{R}{a}\right)^2, \quad c_0=\frac{1}{\tilde{\varepsilon}_0}-\frac{1}{\varepsilon_\infty},$$

where $\tilde{\varepsilon}_0$ and ε_∞ are the low-frequency and high-frequency dielectric constant, respectively.

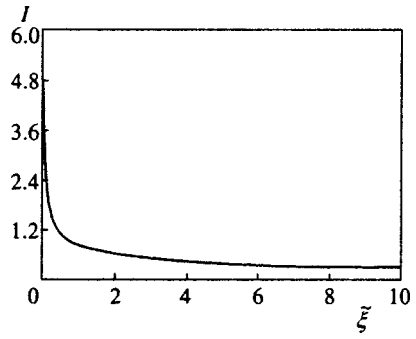

FIG. 1. Dependence of the normalized heat liberation parameter I on $\tilde{\xi}$.

Figure 1 plots the dependence of $a_{10}/\tilde{a}_0$ on $\tilde{\xi}$.

The parameter z entering into the argument of the Bessel function in relation (20) determines the intensity of the phononless line of the intrinsic luminescence ($m=0$) and also the phonon satellites ($m\neq 0$). At high temperatures T

≈ 190 K ($N=1$) for $\tilde{\xi}=1$, $\hbar\omega_0=20$ meV, and $a=50$ Å, the parameter z for different quantum wells varies over quite wide limits. It is equal to 0.3 for GaAs/AlGaAs quantum wells, 0.34 for InP/InPAs quantum wells, and 1.8 for GaN/AlGaN quantum wells. Consequently, in quasi-two-dimensional systems with $z<1$ luminescence lines should be observed which are separated from one another by $\hbar\omega_0$ and whose half-width is given by formula (25) if we take into account the interaction of the carriers with the acoustic vibrations, and the line shape is described by a Gaussian curve. If the criterion of ''strong heat liberation'' is met,[17] $a_{01}\tanh(\hbar\omega_0/2k_0T)\gg 1$ ($z>1$), then the luminescence line is described by a Gaussian curve with half-width

$$\delta\Omega=2\hbar\omega_0\sqrt{2a_{01}\tanh(\hbar\omega_0/2k_0T)\ln 2}\,. \tag{27}$$

For example, for a GaN/AlGaN quantum well at high temperatures ($T=200$ K) for $\tilde{\xi}=1$ and $\hbar\omega_0=0.05$ eV, we have $\delta\Omega=100$ meV.

The light absorption coefficient given by relation (13), which takes account of the interaction of the carriers with the optical phonons, can be represented in the form

$$K(\Omega)=K_0\sum_{n,\nu,p} I_p(z)\left(\frac{N}{N+1}\right)^{p/2}\exp[-a_{n\nu}(2N+1)]$$
$$\times\delta\{\hbar\Omega-\varepsilon_{n\nu}+a_{n\nu}\hbar\omega_0+p\hbar\omega_0\}, \tag{28}$$
$$K_0=\frac{4\pi e^2}{an_0c\Omega R^2}\left|\frac{\mathbf{P}_{cv}\cdot\boldsymbol{\xi}_0}{m_0}\right|^2.$$

If we take account of the interaction of the carriers with the long-wavelength acoustic vibrations, then $K(\Omega)$ is described by a system of equidistant Gaussian curves, the position of the maxima of which corresponds to the frequencies $\Omega=\Omega_0-p\omega_0$:

$$K(\Omega)=K_0\sqrt{\frac{1}{2\pi\Delta}}\sum_p\left(\frac{N}{N+1}\right)^{p/2}I_p(z)$$
$$\times\exp\left\{-\frac{(\hbar\Omega-\hbar\Omega_0+p\hbar\omega_0)^2}{2\Delta}\right\}. \tag{29}$$

Here $\hbar\Omega_0=E_g+\hbar\omega_c^*/2+\varepsilon_0^*-\hbar\omega_0a_{01}$.

Let us investigate the behavior of $K(\Omega)$ for $p=(\Omega_0-\Omega)/\omega_0\gg 1$, i.e., in the long-wavelength absorption region. Following Ref. 18, it can be shown that for a fixed value of p the light absorption coefficient is given by

$$K(\Omega)\simeq K_0\frac{1}{\sqrt{2\pi\Delta}}\exp\left\{-\frac{\hbar(\Omega-\Omega_0)}{k_0T}\sigma\right\}, \tag{30}$$
$$\sigma=1+\ln\left(\frac{\Omega-\Omega_0}{e_0\omega_0a_{01}(N+1)}\right),\quad e_0=2.7182.$$

Consequently, in the long-wavelength region $K(\Omega)$ is described by the Urbach rule. The parameter σ increases with growth of the width of the quantum well and depends weakly on the strength of the magnetic field (for $\tilde{\xi}>1$) and the frequency of the light. For example, for a GaN/AlGaN quantum well for $T=190$ K and $p=5$ we have $\sigma=1.87$. As was shown in Ref. 18, relation (30) holds for $z\ll 1$ as well as $z>1$, i.e., in the low-temperature region.

In the calculation of optical multiphonon processes, we used the diagonal approximation in the quantum numbers n, ν [see relation (8)]. It is precisely this approximation that has allowed us to average over the phonon subsystem exactly. If $n\neq n_1$ and $\nu\neq\nu_1$, then the average in expression (12) can be found approximately by using the cumulant expansion[19] limited to the second cumulant. This approximation in the theory of magneto-optical effects, as was shown in Ref. 20, corresponds in the language of the diagrammatic technique of Konstantinov and Perel' to summing over graphs without intersecting phonon lines[21] and to the usual splitting of the chain of Green's functions.[22] As calculations show, the contribution of the terms $n\neq n_1$, $\nu\neq\nu_1$ to the heat liberation parameter (19), both for the optical and for the acoustic vibrations for $\tilde{\xi}>1$ is less than 10%. Consequently, the diagonal approximation turns out to be completely reasonable in the region of large magnetic fields for studying the effects of the electron–vibrational interaction in the optical spectra in size-limited systems.

*)E-mail: exciton@phys.asm.md

[1] Yu. E. Perlin, Usp. Fiz. Nauk **80**, 553 (1963) [Sov. Phys. Usp. **6**, 542 (1964)].
[2] Ya. B. Zel'dovich, Zh. Éksp. Teor. Fiz. **51**, 1492 (1966) [Sov. Phys. JETP **24**, 1006 (1967)].
[3] V. A. Kovarskiĭ, *Multiquantum Transitions* [in Russian] (Shtiintsa, Kishinev, 1974).
[4] I. V. Lerner and Yu. E. Lozovik, Zh. Éksp. Teor. Fiz. **78**, 1167 (1980) [Sov. Phys. JETP **51**, 588 (1980)].
[5] H. Q. Hou, W. Staguhn, N. Miura, Y. Segawa, S. Takeyama, Y. Aoyagi, and J. M. Zhou, Solid State Commun. **74**, 687 (1990).
[6] L. V. Butov, A. Zrenner, M. Shayegan *et al.*, Phys. Rev. B **49**, 14054 (1994).
[7] W. Edelstem, H. N. Spector, and R. Marasas, Phys. Rev. B **39**, 7697 (1989).
[8] R. S. Knox, ''Theory of Excitons,'' in *Solid State Physics*, Suppl. 5 (Academic Press, New York, 1963).
[9] R. J. Elliot and R. London, J. Phys. Chem. Solids **15**, 196 (1960).
[10] R. Kubo, J. Phys. Soc. Japan **12**, 570 (1957) (in *Thermodynamics of Irreversible Processes* [Russian translation] (IIL, Moscow, 1962), p. 345).
[11] W. H. Louisell, *Radiation and Noise in Quantum Electronics* (McGraw-Hill, New York, 1964).
[12] A. P. Levanyuk and V. V. Osipov, Usp. Fiz. Nauk **133**, 427 (1981) [Sov. Phys. Usp. **24**, 187 (1981)].
[13] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, transl. of 4th Russ. ed. (Academic Press, New York, 1980).
[14] Q. X. Zhao, P. O. Holtz, B. Monemar, *et al.*, Phys. Rev. B **47**, 11890 (1993).
[15] S. R. Kurtz and R. M. Biefeld, Appl. Phys. Lett. **66**, 364 (1995).
[16] H. Buhman, W. Joss, K. v. Klitzing *et al.*, Phys. Rev. Lett. **66**, 926 (1991).
[17] Yu. E. Perlin, B. S. Tsukerblat, *Effects of Electron–Vibrational Interactions in the Optical Spectra of Paramagnetic Impurity Ions* [in Russian] (Shtiintsa, Kishinev, 1974).
[18] A. S. Davydov and A. F. Lubchenko, Dokl. Akad. Nauk SSSR **179**, 1301 (1968) [Sov. Phys. Dokl. **13**, 325 (1968)].
[19] R. Kubo, J. Phys. Soc. Jpn. **17**, 1100 (1962).
[20] É. P. Sinyavskiĭ, *Kinetic Effects in Electron–Phonon Systems in a Laser Field* [in Russian] (Shtiintsa, Kishinev, 1976).
[21] L. I. Korovin, E. V. Kharitonov, Fiz. Tverd. Tela (Leningrad) **7**, 2162 (1965) [Sov. Phys. Solid State **7**, 1740 (1966)].
[22] G. Giobanu and L. Banyani, Phys. Status Solidi **3**, 2299 (1963).

Translated by Paul F. Schippnick

Universal description of crossover between the Mott regime and the Coloumb-gap regime in hopping conductivity: application to compensated CdTe

N. V. Agrinskaya*) and V. I. Kozub†)

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia
(Submitted 31 March 1999)
Zh. Éksp. Teor. Fiz. **116**, 2079–2086 (December 1999)

A simple approach to describing crossover between the Mott regime and the Coulomb-gap regime in hopping conductivity with variable-range hopping is described, based on notions of percolation theory and utilizing an interpolation expression for the density of states. The theoretically derived equation for the argument of the exponential of the resistance ($\ln R$) can be solved numerically. The universal function $\ln(R(T))$ found in this way provides a good description of the experimental curves for crystals of CdTe with varying degree of compensation; however, for samples near the metal–insulator transition it is necessary to take into account the temperature dependence of the pre-exponential factor. Both the form of the obtained function and its comparison with the experimental data show that the crossover region is indeed very wide and, as a rule, neither in the low-temperature region nor in the high-temperature region, are the limiting values of the exponent characterizing the Mott law or the Efros–Shklovskiĭ law reached. With the help of the universal function we obtain values of the parameters T_0 and T_1 characterizing the density of states and the width of the Coulomb gap, for CdTe samples with varying degree of compensation and compare them with their theoretical values. Despite taking the crossover function into account, the values of the parameter T_1 turn out to be almost an order of magnitude lower than those calculated theoretically by Efros and Shklovskiĭ.
 [S1063-7761(99)01612-1]

1. INTRODUCTION

Despite the large number of publications that have addressed the question of the mechanisms of hopping conductivity with variable-range hopping in doped semiconductors, some questions remain unsolved. These include questions of the character of hopping conductivity in the Coulomb-gap regime and the role of many-particle effects (correlated hops), etc.

A study of the corresponding details of the mechanism of hopping conductivity obviously requires a comparison of the experimental data, in particular, on the temperature dependence of the resistance, with the results of theoretical calculations over a wide range of temperatures. Here it is necessary to note the following fact (to which attention has been drawn recently, see, e.g., Refs. 1–3). As it turns out, for materials with typical parameters (intermediate impurity concentration and intermediate compensation) in the temperature region that is typical for the experiments (~ 0.01–10 K) crossover is observed from conductivity with variable-range hopping (VRH) of Mott type [$\rho(T) \propto \exp(T_0/T)^{1/4}$] to conductivity over states in the Coulomb gap—the Efros–Shklovskiĭ law [$\rho(T) \propto \exp(T_1/T)^{1/2}$]. On the one hand, the nature of the conductivity in the two indicated regimes differs both in its numerical parameters and, possibly, in its very physical nature, and on the other hand, the crossover region turns out to be quite wide; therefore, a comparison of the predictions of the theory with experiment is quite complicated. Thus, in order to be able to compare the experimental values of the parameters T_0 and T_1 with their calculated values

$$T_0 = \beta_0 / g(\varepsilon_f) a^3, \quad T_1 = \beta_1 e^2 / \kappa a$$

(where $\beta_0 = 21$ and $\beta_1 = 2.8$ are numerical coefficients, a is the localization radius, and κ is the dielectric constant), it is important to have a valid description of the crossover region. Earlier, using a simplified crossover model (based on matching of the effective energy bands corresponding to conductivity for the two indicated regimes), we drew attention to the possible role of the pre-exponential factors in the expression for the hopping conductivity. The latter, obviously, is especially important for samples hear the metal–insulator transition, for which the interval of variation of the resistance is not too large.[4] At the same time, for samples far from this transition, in a number of cases the pre-exponential factors can be neglected.

The first attempt to describe the temperature behavior of the conductivity in the crossover region with the aid of a universal expression was made in Ref. 2, where the authors used an interpolation expression for the effective energy bands, obtained by simply summing the corresponding expressions for the two regimes. As was shown later,[3] such an approach is not entirely valid and smears out the crossover singularities. The latter work proposes a more general approach to the description of crossover, which leads to a quite complicated multiparameter integral expression, not accessible to analytical treatment and requiring quite cumbersome numerical analysis. Note also Ref. 5, in which to describe

crossover in the two-dimensional case the authors proposed a procedure based on optimization of the argument of the exponential in the expression for the hopping resistance which made use of an interpolation expression for the density of states. Although this approach leads to a comparatively simple analytical expression, it is less rigorous than an analysis of the percolation problem. As was shown in Ref. 3, it leads to a noticeable overestimate of the width of the crossover region.

Therefore, the present work proposes a simpler approach to the description of crossover, based (as in the approach of Ref. 3) on notions of percolation theory analogous to those used by Efros and Shklovskiĭ,[6] however, enabling a simpler analysis of the condition of connectedness. The theoretically obtained equation is applied to determine the temperature dependences in the crossover region, which are compared with the experimental dependences for CdTe crystals with varying degree of compensation.

2. THEORY

We express the dimensionless concentration of sites satisfying the condition of connectedness for the argument of the exponential less than some value ξ, as

$$n(\xi)=2\int_0^{\varepsilon_{\max}} r_{\max}^3 g(\varepsilon)d\varepsilon. \tag{1}$$

Here $\varepsilon_{\max}$ and $r_{\max}$ are the maximum values of the energy and interstitial distance allowing connectedness, and $g(\varepsilon)$ is the density of states. Equating the corresponding value to the critical concentration for the percolation threshold n_c (for the three-dimensional situation the corresponding value is equal to ~ 5.3), we obtain an equation for the critical value of ξ determining the resistance.

Note that the dimensionless concentration is defined as the product of the concentration of sites with $\varepsilon<\varepsilon_{\max}$ and the volume accessible to these sites. Correspondingly, for the density of states we have used the well-known interpolation expression (see, e.g., Ref. 3)

$$g=g_0\frac{(\varepsilon/\varepsilon_0)^2}{1+(\varepsilon/\varepsilon_0)^2}, \tag{2}$$

where g_0 is the density of states in the absence of the Coulomb gap, and

$$\varepsilon_0\simeq 3^{-1/2}\left(\frac{e^6 g_0}{\kappa}\right)^{1/2}. \tag{3}$$

Note that the given procedure (used by Efros and Shklovskiĭ[6] in application to the Mott regime) for an energy-dependent density of states is not entirely rigorous, and the condition of connectedness must be considered separately for each value of the energy, which leads to an integral equation.[3] We, however, make use of a simplified procedure which makes it possible to simplify the calculations substantially. Note that the coefficient $3^{-1/2}$ in the expression for ε_0 is connected with this specific procedure and provides the correct asymptotic behavior in the limit of low temperatures. Direct calculation of the integral on the right-hand side gives

$$2(\varepsilon_{\max}-\varepsilon_0\arctan(\varepsilon_{\max}/\varepsilon_0))r_{\max}^3. \tag{4}$$

Taking into account that $r_{\max}=(a\xi/2)^3$ and $\varepsilon_{\max}=T\xi$, where ξ is the critical value of the argument of the exponential in the temperature dependence of the hopping resistance $\ln R(T)=A\xi(T)$, where A is a coefficient and a is the localization radius, we finally obtain the following equation for ξ:

$$\left(T\xi-\varepsilon_0\arctan\left(\frac{\xi T}{\varepsilon_0}\right)\right)(a\xi)^3\simeq 21. \tag{5}$$

Expressing the parameters a and g_0 in terms of the characteristic temperatures T_0 and T_1, we obtain for the Mott and Efros–Shklovskiĭ laws

$$\left(T\xi-0.6\,\frac{T_1^{3/2}}{T_0^{1/2}}\arctan\left(\frac{\xi T T_0^{1/2}}{0.6T_1^{3/2}}\right)\right)\xi^3=T_0. \tag{6}$$

For comparison we have generalized the results of the calculation in Ref. 5, which also uses an interpolation expression for the density of states, but based on optimization of the argument of the exponential in the expression for the elementary hopping resistance, to the three-dimensional case. The given approach leads to a somewhat different equation for the critical value ξ governing the magnitude of the hopping resistance:

$$\left(T\xi-\frac{1}{9}\,\frac{T_1^{3/2}}{T_0^{1/2}}\arctan\left(\frac{\xi T T_0^{1/2}}{(1/9)T_1^{3/2}}\right)\right)\times\left(\frac{\xi^2+(TT_0^{1/2}/(1/9)T_1^{3/2})^{-2}}{\xi^2}\right)^{3/4}=(T^3T_0)^{1/4}. \tag{7}$$

3. EXPERIMENT

We investigated CdTe crystals doped with shallow donors. In two of the investigated samples the donor concentration was roughly the same, $N_d\geqslant 2\times 10^{17}\,\mathrm{cm}^{-3}$, and the degree of compensation by native lattice defects was different: $\geqslant 0.4$ and $\geqslant 0.6$, and the electron density at 300 K was respectively 1.2×10^{17} and $8\times 10^{16}\,\mathrm{cm}^{-3}$ (the exact values of N_d and the degrees of compensation in the give case are impossible to determine due to the absence of a temperature dependence of the Hall constant). Sample 3 was doped lightly, and from the temperature dependence of its Hall constant we were able to determine separately the total donor concentration $N_d=1.1\times 10^{17}\,\mathrm{cm}^{-3}$ and total acceptor concentration $N_a=8.5\times 10^{16}\,\mathrm{cm}^{-3}$, the concentration of filled donors $(N_d-N_a)=n_{300\,\mathrm{K}}=2.5\times 10^{16}\,\mathrm{cm}^{-3}$, and the degree of compensation $\mathcal{K}=0.77$. All samples were on the insulating side of the metal–insulator transition associated with compensation.

Figures 1 and 2 plot the temperature dependence of the resistance $R(T)$ for the three samples together with curves obtained by solving Eq. (6), which has the form $\xi=A\ln R$ for optimally chosen values of the parameters T_0 and T_1. (As a consequence of the high values of the resistance at low temperatures, for sample 1 the dependence $R(T)$ could be tracked down to only 0.5 K.) It can be seen that for the two highest-resistance samples the experimental curves give a

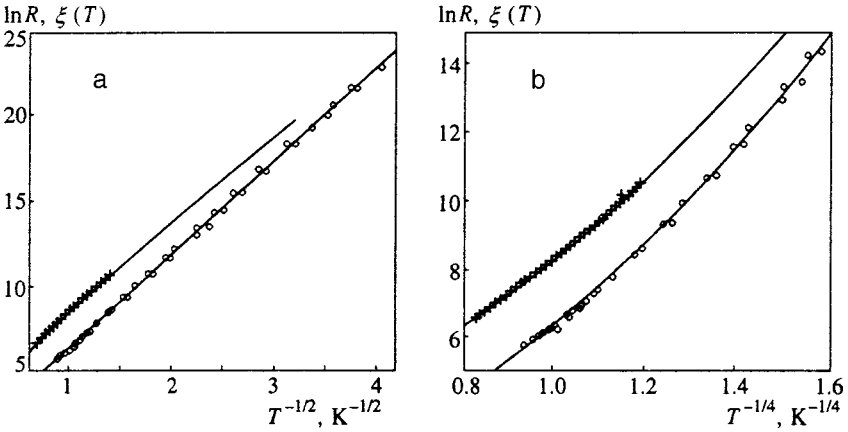

FIG. 1. Experimental temperature dependence of the resistance for samples 1 (+) and 2 (○); the curves were calculated by solving Eq. (6) for the function $\xi(T)$ for the parameter values $T_0=6500$ K and $T_1=33$ K for sample 1 and $T_0=5900$ K and $T_1=36$ K for sample 2.

good description of the function $\xi(T)$, which is the solution of Eq. (6) with optimally chosen values of the parameters T_0 and T_1, and that in these two cases the Mott and Efros–Shklovskiĭ laws become valid in the low-temperature and high-temperature limits. The parameter values are given in Table I. For sample 3 the experimental and theoretical curves are seen to diverge in the high- and low-temperature regions. From our point of view, this is explained by the fact that we have neglected the temperature dependence of the pre-exponential factors in the expression for $R(T)$. The corresponding factor proves to be important for samples near the transition, for which the interval of variation of the resistance is not too large.[4] In Fig. 2 the dependence $R(T)$ for sample 3 is constructed on scales that take the temperature dependence of the pre-exponential factor in the Mott law into account: $\ln(RT^{-1/4})$ versus $T^{-1/4}$. The agreement in the high- and low-temperature regions is noticeably improved. The values of T_0 and T_1 obtained in this way are given in the table.

Note that in the region of applicability of the Efros–Shklovskiĭ law a falloff of the negative magnetoresistance is observed. This fact can also serve as an indicator of a change in the character of the hopping resistance (in our case, of the transition to the Efros–Shklovskiĭ law). We discussed mechanisms leading to such behavior previously in Refs. 7 and 8.

4. DISCUSSION

The picture of the temperature dependence of the conductivity of doped semiconductors in the region of hopping conductivity with variable-range hopping (0.01–10 K) is often quite complicated and does not obey in detail the well-known Mott and Efros–Shklovskiĭ laws; here we note in particular the quantitative discrepancies observed, as a rule, for the parameters T_0 and T_1 in comparison with the theoretical estimates of Mott and Efros–Shklovskiĭ. The problem consists, in particular, in the fact that the generally used range of temperatures turns out to lie in the transition region between the Mott and Efros–Shklovskiĭ laws. Since there is a competition going on in the high-temperature region with the activation temperature dependence for the conductivity via nearest neighbors, the high-temperature region for the VRH conductivity is small enough to observe the pure Mott law, but the low-temperature region (usually down to 0.01 K) is not large enough to observe the pure Efros–Shklovskiĭ law (the impossibility of going lower into the region of superlow temperatures is often connected with the large resistance of the sample and heating effects). In light of what has been said, it is important to derive a universal expression for the transitional region. As was already noted, such efforts have been made; however, they are marred by an inadequate de-

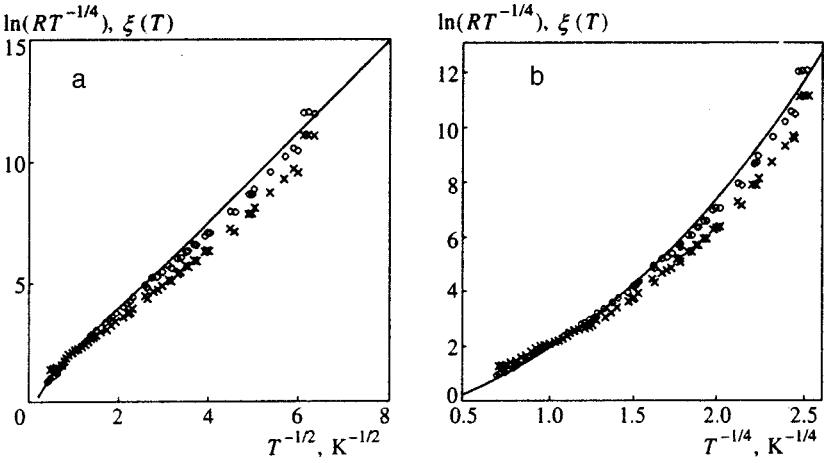

FIG. 2. Experimental temperature dependence of the resistance for sample 3; the symbols × represent the functional dependence $\ln(R(T))$; the symbols ○ represent the functional dependence $\ln(R(T)T^{-1/4})$; the curves were calculated by solving Eq. (6) for the function $\xi(T)$ for the parameter values $T_0=110$ K and $T_1=3.8$ K.

TABLE I.

N	$n_{300\,\mathrm{K}}$, cm^{-3}	N_d, cm^{-3}	$\mathcal{K}$	T_0, K	T_1, K
1	2.5×10^{16}	1.1×10^{17}	0.77	6500	33
2	8×10^{16}	$\geqslant2\times10^{17}$	$\geqslant0.6$	5900	36
3	1.2×10^{17}	$\geqslant2\times10^{17}$	$\geqslant0.4$	110	3.8

Remark. $n_{300\,\mathrm{K}}$ — electron density at 300 K, obtained from Hall measurements, N_d — donor concentration, $\mathcal{K}$ — degree of compensation, T_0, T_1 — parameters obtained from the temperature dependence of the resistance with the help of Eq. (6).

scription of the experiment or by cumbersomeness of the obtained expressions. The approach proposed above leads to a simple analytical expression, whose solution $\ln R(T) = A\xi(T)$, as can be seen from Figs. 1 and 2, provides a good description of the experimental data.

In order to understand how fast the transition to the corresponding power laws goes as the temperature is decreased or increased, relative to the crossover region, it is customary (see, e.g., Ref. 9) to use a graph of the dependence $d\xi/(d\log T)$ in which the changeover to the corresponding power-law asymptotic limit corresponds to a straightening of the curve with slope equal to the exponent of the power law. Note, however, that determining both the degree of straightening and the corresponding slope by purely graphical methods is fraught with some difficulty. Therefore, we propose a somewhat different procedure which allows one to identify the character of the approach to the asymptotic limit in a purely quantitative way. Toward this end, we use the expression

$$p(\log T)=\frac{d\log\xi}{d\log T}=\frac{T}{\xi}\frac{d\xi}{dT}.$$

As can be easily seen, for a pure power law this expression is equal to the exponent of the corresponding dependence and thus the asymptotic behavior is determined by the approach of the dependence $p(\log T)$ to the corresponding constant. In Fig. 3 the function $p(\log T)$ is constructed for $\xi(T)$ found with the help of Eqs. (6) and (7). It can be seen that the proposed procedure leading to Eq. (6) provides a comparatively rapid approach to the asymptotic limit whereas the

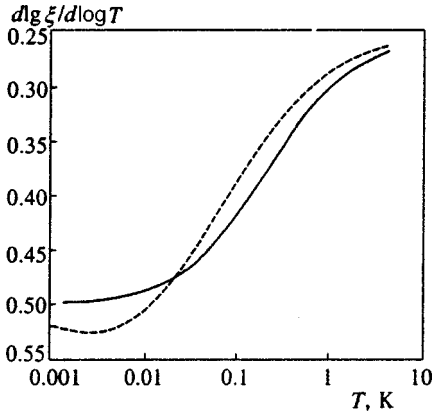

FIG. 3. Asymptotic behavior of the function $\xi(T)$, calculated by solving Eq. (6) — solid curve, and Eq. (7) — dashed curve. Parameters $T_0=5900$ K and $T_1=36$ K.

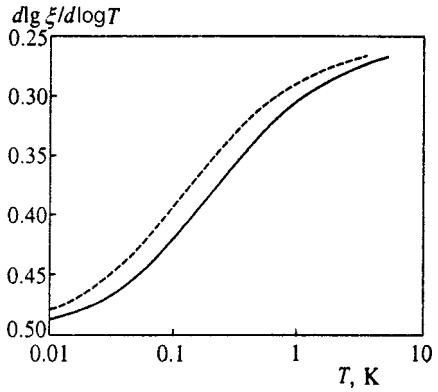

FIG. 4. Asymptotic behavior of the function $\xi(T)$, calculated by solving Eq. (6) for sample 2 with parameter values $T_0=5900$ K and $T_1=36$ K (solid curve) and sample 3 with parameter values $T_0=110$ K and $T_1=3.8$ K.

procedure leading to Eq. (7) not only substantially overestimates the width of the crossover region, but is also capable of leading to an erroneous identification of the crossover. Indeed, as can be seen from Fig. 3, the value of the exponent p in the intermediate region even exceeds 1/2 and approaches the asymptotic value only at very low temperatures. Thus, the use of the standard procedure for identifying the asymptotic behavior by estimating the slope of the curve of the logarithmic derivative can lead in this case to an underestimate of the width of the crossover region [since the value $p=1/2$ is reached not only in the real asymptotic region but also, by virtue of the nonmonotonic behavior of $p(\log T)$, in the crossover region].

In Fig. 4 we show curves of $p(\log T)$ constructed for functions $\xi(T)$ determined from the experimental curves for samples 2 and 3. It can be seen that the crossover region is quite wide (variation of the temperature over three orders of magnitude); however, at the edges of this region the function approaches asymptotically the exponents characteristic of the Mott and Efros–Shklovskiĭ laws (0.25 and 0.5). The temperature characterizing the crossover, T_c, is shifted (e.g., for $p=0.375$) for sample 3 into the low-temperature region, $T_c=0.1$ K, in comparison with the corresponding temperature for sample 2, $T_c=0.2$ K. Besides, these values are in good agreement with those calculated using the expression $T_c=T_1^2/T_0$, which also demonstrates the validity of describing crossover on the basis of solution of Eq. (6).

Let us compare the obtained values of T_0 and T_1 with their theoretical values for the most compensated samples 1 and 2, which are situated quite far from the insulator–metal transition (for which reason it is possible to neglect the divergence of κ and a. The parameter T_0 can be calculated, knowing the donor concentration and the width of the impurity band and assuming the localization radius to be equal to the radius of an isolated donor, $a_0=50$ Å. These parameters are known most accurately for sample 1: $N_d=1.1\times10^{17}$ cm^{-3}, width of the impurity band $\varepsilon_3=0.002$ eV. Thus, setting $g(\varepsilon_f)=N_d/\varepsilon_3$, for this case we obtain the calculated value $T_{0cal}=\beta_0/g(\varepsilon_f)a^3=10000$ K, which is not much higher than the value obtained from the results of an analysis of crossover: $T_0=6500$ K. (The indicated discrepancy can be explained by corrections to the localization radius associated

with scattering by neutral and charged impurities;[10] the maximum value of the correction is $\sim a_0$, and in this case the maximum value of the ratio $T_{0\mathrm{cal}}/T_0 \sim 8$). Thus we can conclude that the one-particle density of states provides a good description of the Mott regime.

We calculate the parameter T_1 by setting $\kappa = 11$; for the insulating limit it should be equal to 1000 K. The maximum value of T_1 obtained from the temperature data for samples 1 and 2 located far into the insulating side of the metal–insulator transition is 30–40, i.e., it is 25 times smaller than the theoretical value. (Corrections to the localization radius can explain only a discrepancy by a factor of two.)

This latter circumstance is in agreement with the conclusions of a number of earlier works on doped semiconductors with intermediate degree of compensation[11,12] and is evidence of the role of many-particle correlated hops[13] or of a sequence of assisting hops[14] in the conductivity via states of the Coulomb gap. Note that the role of Coulomb correlations should be greatest in the Coulomb-gap regime, where Coulomb interactions determine the character of the energy scale.

The authors are grateful to R. V. Parfen'ev for consultation in regard to low-temperature measurements, and A. N. Chernyaev for assistance with the measurements. This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project No. 97-02-18280).

*)E-mail: nina.agrins@pop.ioffe.rssi.ru
†)E-mail: ven.kozub@pop.ioffe.rssi.ru

[1] N. V. Agrinskaya and A. N. Aleshin, Fiz. Tverd. Tela (Leningrad) **31**, 277 (1989) [Sov. Phys. Solid State **31**, 1996 (1989)].
[2] A. Aharony, Y. Zhang, and M. P. Sarachik, Phys. Rev. Lett. **68**, 3900 (1992).
[3] Y. Meir, Phys. Rev. Lett. **77**, 5265 (1996).
[4] N. V. Agrinskaya and V. I. Kozub, Zh. Éksp. Teor. Fiz. **106**, 848 (1994) [JETP **79**, 466 (1994)].
[5] Nguen Van Lien, Phys. Lett. A **207**, 379 (1995); R. Rosenbaum and Nguen Van Lien, J. Phys. C **9**, 6247 (1997).
[6] A. L. Efros and B. I. Shklovskii, *Electronic Properties of Doped Semiconductors* (Elsevier, 1985).
[7] N. V. Agrinskaya, V. I. Kozub, and D. V. Shampur, Zh. Éksp. Teor. Fiz. **107**, 2063 (1995) [JETP **80**, 1142 (1995)].
[8] N. V. Agrinskaya, V. I. Kozub, R. Rench, P. Fozoni, and M. D. Li, Zh. Éksp. Teor. Fiz. **111**, 1477 (1997) [JETP **84**, 814 (1997)].
[9] A. G. Zabrodskiĭ and K. N. Zinov'eva, Zh. Éksp. Teor. Fiz. **86**, 727 (1984) [Sov. Phys. JETP **59**, 425 (1984)].
[10] B. I. Shklovskii and B. Z. Spivak, in *Hopping Transport in Solids*, edited by M. Pollak and B. Shklovskii (Elsevier, 1991), p. 271.
[11] A. G. Zabrodskiĭ and A. G. Andreev, JETP Lett. **58**, 756 (1993).
[12] J. G. Massey and M. Lea, Phys. Rev. Lett. **75**, 4266 (1995).
[13] M. Pollak and M. Ortuno, in *Electron–Electron Interactions in Disordered Systems*, edited by A. L. Efros and M. Pollak (Amsterdam, 1985).
[14] N. V. Agrinskaya and V. I. Kozub, Fiz. Tekh. Poluprovodn. **32**, 703 (1998) [Semiconductors **32**, 631 (1998)].

Translated by Paul F. Schippnick

Influence of the interconfigurational interaction on the crystal field of Ln^{3+} ions

A. A. Kornienko and E. B. Dunina

Vishebsk State University, 210036 Vishebsk, Belarus

A. A. Kaminskiĭ*)

Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia

(Submitted 7 April 1999)
Zh. Éksp. Teor. Fiz. **116**, 2087–2102 (December 1999)

A theory of the crystal field for Ln^{3+} ions is proposed which takes account of the difference in the effect of excited configurations on high-lying and low-lying multiplets. The effective-operator method in third-order perturbation theory is used to obtain the Hamiltonian of the crystal field, which in addition to the usual terms contains energy-dependent operators. Their role is discussed in detail. For the new operators we have obtained convenient expressions which make it possible for the first time to determine the parameters of an odd crystal field on the basis of an analysis of the structure of the energy spectrum. Theory is compared with experiment for the laser crystals $Y_3Al_5O_{12}:Tm^{3+}$ and $LiYF_4:Pr^{3+}$. Taking the new terms of the crystal-field Hamiltonian into account produces an additional shift of individual levels within the limits from $-40\,cm^{-1}$ to $40\,cm^{-1}$ and makes it possible in a number of cases not only to substantially reduce the value of the standard deviation, but also to obtain the correct arrangement of levels.
 [S1063-7761(99)01712-6]

1. INTRODUCTION

Crystals with trivalent lanthanide ions (Ln^{3+}) have found wide application as active laser media. Therefore, a theoretical and experimental study of the energy spectrum and intensity characteristics of the absorption and luminescence of such crystals is of great practical importance. Although these crystals are suitable objects of theoretical study (relatively small influence of the crystal field, narrow lines, large number of observable transitions), it is nevertheless not possible in the one-electron approximation to obtain the desired accuracy of description, and sometimes the theoretical results even contradict the experimental data.

Efforts to improve the description of the experimental data by adding electrical dipole transitions of two-particle operators representing different electron-correlation effects to the one-electron crystal-field Hamiltonian and the one-electron line-strength operator have had little success.[1–7] In this regard, studies of the influence of the interconfigurational interaction on the states of the f^N configuration seem to be more promising.

The point is that the one-electron crystal-field Hamiltonian usually used to interpret optical spectra

$$H_{cf}=\sum_{k,q} B_q^k C_q^k \tag{1}$$

and the one-electron line-strength operator

$$S_{JJ'}=e^2\sum_{k=2,4,6}\Omega_k\langle \gamma J\|U^k\|\gamma' J'\rangle^2 \tag{2}$$

were derived in the weak-configurational interaction approximation,[8–10] where the excited configurations act to the same degree on different multiplets. Here B_q^k are the crystal-field parameters, C_q^k is the spherical tensor, Ω_k are the intensity parameters, and $\langle \gamma J\|U^k\|\gamma' J'\rangle$ are the reduced matrix elements of the unit tensor U^k. In this approximation the sets of parameters B_q^k and Ω_k should be the same for all multiplets of the f^N configuration.

In reality, the energies of the multiplets of rare-earth ions have the same order of magnitude as the energies of the lower excited configurations. Thus, fulfillment of the conditions for realization of the weak-configurational interaction approximation is improbable, and use of Eqs. (1) and (2) should encounter contradiction even more often than actually is the case. Since the energy intervals between the excited configuration and the high- and low-lying multiplets differ significantly from one another, the excited configurations will influence different multiplets to a substantially different extent. If we take this effect into account in third-order perturbation theory, we obtain the following crystal-field Hamiltonian:

$$H_{cf}=\sum_{k,q}\underbrace{[B_q^k+(E_J+E_{J'}-2E_f^0)G_q^k]}_{\tilde{B}_q^k}C_q^k \tag{3}$$

and effective line-strength operator[12]

$$S_{JJ'}=e^2\sum_{k=2,4,6}\underbrace{\Omega_k[1+2R_k(E_J+E_{J'}-2E_f^0)]}_{\tilde{\Omega}_k}\times\langle \gamma J\|U^k\|\gamma' J'\rangle^2, \tag{4}$$

in which the parameters $\tilde{B}_q^k$ and $\tilde{\Omega}_k$ depend linearly on the energy of the multiplets E_J and $E_{J'}$. Here E_f^0 is the energy of the center of mass of the f^N configuration, G_q^k and R_k are

additional parameters prescribing the amplitude of the interconfigurational interaction. The approximation of such an intermediate-in-strength interconfigurational interaction will obviously be more adequate for Ln^{3+} ions.

The excited configurations of actinides have a lower energy than the corresponding configurations of lanthanides. Therefore, for the actinides the interconfigurational interaction should be stronger. The following crystal-field Hamiltonian was obtained in the strong configurational interaction approximation in first-order perturbation theory[13]:

$$H_{cf}=\sum_{k,q}\underbrace{\left[B_q^k+\left(\frac{\Delta}{\Delta-E_J}+\frac{\Delta}{\Delta-E_{J'}}\right)\tilde{G}_q^k\right]}_{\overline{B_q^k}}C_q^k \tag{5}$$

with the following effective line-strength operator of the electric dipole transitions[14,15]

$$S_{JJ'}=e^2\sum_{k=2,4,6}\underbrace{\Omega_k\left[\frac{\Delta}{\Delta-E_J}+\frac{\Delta}{\Delta-E_{J'}}\right]^2}_{\overline{\Omega_k}}\langle\gamma J\|U^k\|\gamma' J'\rangle^2. \tag{6}$$

Here Δ is the energy of the excited configuration. In this case the crystal-field parameters $\overline{B_q^k}$ and intensity parameters $\overline{\Omega_k}$ depend on the energy of the multiplets according to a law that is more complicated than linear.

Thus, the effective operators (3)–(6) in the intermediate and strong configurational interaction approximation were recently proposed to describe the optical spectra of crystals activated by rare-earth ions. Only under certain conditions is it possible to limit the treatment to such a simple tensor form of the effective operators as (3)–(6). Despite the successful application of expressions (3)–(6) to the description of the spectral properties of a number of systems,[12–14,16] the conditions of their applicability have not yet been sufficiently investigated. This especially holds true in regard to the Hamiltonian (3), for which there is only a preliminary communication.[11]

Therefore, the main goal of this paper is to present a detailed study of the conditions of derivation and applicability of the crystal-field Hamiltonian (3) in the intermediate configurational interaction approximation together with comprehensive tests. To test the approximation, we chose the optical spectra of the typical generating ions Pr^{3+} and Tm^{3+} with experimentally well-established Stark structure in fields with symmetry S_4 and D_2.

2. EFFECTIVE HAMILTONIAN

Accurate values of the Stark levels can be obtained by diagonalizing the matrix of the Hamiltonian (1) in a basis consisting of the wave functions of the ground configuration and all excited configurations. Since this is hard to realize in practice, the method of the effective Hamiltonian seems more acceptable. The effective Hamiltonian, acting in a model space of significantly lower dimensionality, has the same eigenvalues as the real Hamiltonian.[17] The effective Hamiltonian can be easily constructed with the help of a method developed in Ref. 18 for a basis of non-orthogonal functions. Toward this end, we write out the most important terms from expression (25) of Ref. 18:

$$\begin{aligned}\langle n|H_{\rm eff}|n'\rangle=&\langle n|H|n'\rangle+\sum_b\frac{1}{\Delta_{nb}}\langle n|\mathbf{H}_n|b\rangle\langle b|\mathbf{H}_n|n'\rangle\\&-\frac{1}{2}\sum_{b,n''}\frac{1}{\Delta_{nb}^2}[\langle n|\mathbf{H}_n|b\rangle\langle b|\mathbf{H}_n|n''\rangle\\&\times\langle n''|W|n'\rangle+\langle n|W|n''\rangle\langle n''|\mathbf{H}_n|b\rangle\\&\times\langle b|\mathbf{H}_n|n'\rangle]-\sum_{b,n''}\frac{1}{\Delta_{nb}}[\langle n|\mathbf{H}_n|b\rangle\langle b|n''\rangle\\&\times\langle n''|W|n'\rangle+\langle n|W|n''\rangle\langle n''|b\rangle\\&\times\langle b|\mathbf{H}_n|n'\rangle],\end{aligned} \tag{7}$$

where

$$\langle n|\mathbf{H}_n|b\rangle=\langle n|H|b\rangle-\langle n|H^0|n\rangle\langle n|b\rangle,$$

W is the perturbation potential, H^0 is the unperturbed Hamiltonian, n, n', n'' and b denote the states of the ground configuration and the excited configurations, respectively, and $\Delta_{nb}=E_n-E_b$ is the excitation energy.

The method developed in Refs. 19 and 20 makes it convenient to write out expression (7) in terms of the spherical tensors C_q^k. The main difficulty here is to make the correct choice of the unperturbed Hamiltonian, which affects the interpretation of the matrix elements of the perturbation potential $\langle n|W|n'\rangle$. There are several ways of choosing the unperturbed potential H^0 (Refs. 21–23). Here we have applied a variant of perturbation theory, based on the formalism of projection operators,[24] and the following definition of H^0 (Ref. 23) in terms of the projection operator

$$H^0=\sum_n\overline{\langle n|H|n\rangle}|n\rangle\langle n| \tag{8}$$

is preferable. In expression (8) the sum over n and the mean value $\overline{\langle n|H|n\rangle}$ are calculated over all states of the f^N configuration, wherefore

$$\overline{\langle n|H|n\rangle}\approx\frac{\Sigma_J(2J+1)E_J}{\Sigma_J(2J+1)}=E_f^0. \tag{9}$$

In such a choice of H^0 the energies of all unperturbed states are equal to the center-of-mass energy E_f^0, and in the zeroth approximation the f^N configuration is completely degenerate. This is true only in central fields. Consequently, the perturbation potential $W=H-H^0$ should contain all noncentral interactions, i.e.,

a) the noncentral part of the Coulomb interaction of the electrons with each other and the spin–orbit interaction (these interactions give the main contribution to the energy of the multiplets);

b) the Coulomb interaction of the electrons of the impurity ion with the electrons of the ligands and their nuclei (these interactions are responsible for formation of the Stark structure of the multiplets).

From this point of view it seems reasonable to replace the matrix element $\langle n|W|n'\rangle$ by the expression

$$\langle n|W|n'\rangle=(E_J-E_f^0)\,\delta_{nn'}+\langle n|\sum_{k=2}^{6}\sum_{q=-k}^{k}F_q^kC_q^k + \text{two-particle operators}|n'\rangle, \tag{10}$$

where E_J is the energy of the state $|n\rangle$. The crystal-field parameters F_q^k are due mainly to interactions enumerated in item (b) up above, and wave-function overlap effects. It is well known that these interactions give the defining contribution to the usual crystal-field parameters B_q^k. Consequently, the crystal-field parameters F_q^k should be similar in magnitude to the corresponding parameters B_q^k.

Using an analogous method for the fragment $\langle n|\mathbf{H}_n|b\rangle\times\langle b|\mathbf{H}_n|n''\rangle$ from expression (7), it is easy to obtain

$$-\frac{1}{2}\langle n|\mathbf{H}_n|b\rangle\langle b|\mathbf{H}_n|n''\rangle/\Delta_{bn}^2=\langle n|\sum_{k=2}^{6}\sum_{q=-k}^{k}G_q^kC_q^k + \text{two-particle operators}\ |n''\rangle. \tag{11}$$

Here the parameters G_q^k are due to effects of covalency and the interconfigurational interaction.

After substituting expressions (10) and (11) into expression (7) followed by some simple transformations we obtain the following effective crystal-field Hamiltonian:

$$\begin{aligned}\langle n|H_{\text{eff}}|n'\rangle&=E_J\delta_{nn'}\\&+\langle n|\sum_{k=2}^{6}\sum_{q=-k}^{k}\underbrace{[B_q^k+(E_J+E_{J'}-2E_f^0)G_q^k]}_{\tilde{B}_q^k}\,C_q^k|n'\rangle\\&+\sum_{n''}\langle n|\sum_{k=2}^{6}\sum_{q=-k}^{k}F_q^kC_q^k|n''\rangle\langle n''|\sum_{k=2}^{6}\sum_{q=-k}^{k}G_q^kC_q^k|n'\rangle\\&+\sum_{n''}\langle n|\sum_{k=2}^{6}\sum_{q=-k}^{k}G_q^kC_q^k|n''\rangle\langle n''|\sum_{k=2}^{6}\sum_{q=-k}^{k}F_q^kC_q^k|n'\rangle\\&+\ldots\end{aligned} \tag{12}$$

In the description of the experimental data, the parameters B_q^k, G_q^k, and F_q^k can be considered as variable parameters. Here to reduce the number of fitting parameters it is expedient to set $B_q^k=F_q^k$.

The Hamiltonian (12) is more complicated than the one-electron Hamiltonian (1): its parameters $\tilde{B}_q^k$ are linear functions of the energy of the multiplets, and it contains the ''quadratic crystal field'' operator [the last two lines of Eq. (12)]. Although energy-dependent terms have been added to the Hamiltonian (12) from third-order perturbation theory, their role can be substantial since they are multiplied by the energy of a multiplet, whose order of magnitude is approximately $10000\,\text{cm}^{-1}$. These terms have the sense of energy-dependent corrections to the usual crystal-field parameters. That is to say, these corrections depend on the energy of the multiplets.

The center-of-mass energy E_f^0 is responsible for a uniform shift of the parameters $\tilde{B}_q^k$ by the amount $-2E_f^0G_q^k$. This shift can always be compensated by the corresponding choice of the parameters B_q^k. Therefore, in the description of the experiments, the parameter E_f^0 can be assigned any convenient value. In the present work we set $E_f^0=0$.

The energy-dependent contributions to the crystal-field parameters are determined by all the lines of expression (7) with the exception of the first. Consequently, their amplitude is inversely proportional to the square of the energy difference between the ground configuration and the excited configuration. Such a dependence gives rise to a rapid falloff of the amplitude of the contribution with growth of the excitation energy. Therefore, only the low-lying excited configurations will give the defining contribution to the parameters G_q^k. In the case of f^N systems, these include configurations of the types $nf^{N-1}(n+1)d$, $nf^{N-1}(n+1)g$, $(n+1)p^5nf^{N+1}$, and configurations with electron transfer from the ligand to the f shell. An analytical expression for the contributions to G_q^k due to admixture of the configurations $nf^{N-1}(n+1)d$, $nf^{N-1}(n+1)g$, and $(n+1)p^5nf^{N+1}$ have the form

$$\begin{aligned}G_q^k(l)=(-)^{l+1}&\frac{2k+1}{2\Delta_{fl}^2\langle f\|C^k\|f\rangle}\sum_{\substack{k',k''\\q',q''}}(-1)^q\\&\times\begin{pmatrix}k'&k''&k\\q'&q''&-q\end{pmatrix}\begin{Bmatrix}k'&k''&k\\f&f&l\end{Bmatrix}\langle f\|C^{k'}\|l\rangle\\&\times\langle l\|C^{k''}\|f\rangle B_{q'}^{k'}(l)B_{q''}^{k''}(l),\end{aligned} \tag{13}$$

where $l=d,g,p$; $\langle f\|C^k\|l\rangle$ is the reduced matrix element of the spherical tensor C^k, and we have used the usual notation for the crystal field:

$$B_q^k(l)=\langle f|r^k|l\rangle A_{kq}. \tag{14}$$

The expression for estimating the contributions to G_q^k due to admixture of the configurations with electron transfer can be easily obtained with the aid of methods described in Refs. 19 and 25:

$$\begin{aligned}G_q^k(cov)=&\frac{2k+1}{2\langle f\|C^k\|f\rangle}\sum_b C_q^{k*}(\Theta_b,\Phi_b)\sum_{m,\zeta}(-1)^{f-m}\\&\times\begin{pmatrix}f&k&f\\-m&0&m\end{pmatrix}|\lambda_{f\zeta m}|^2.\end{aligned} \tag{15}$$

Here $C_q^k(\Theta_b,\Phi_b)$ is the spherical tensor of the angles Θ_b and Φ_b defining the direction from the activator ion to the ligand b. The covalency parameter $\lambda_{f\zeta m}$ corresponds to transfer of an electron from the orbit ζ of the ligand to the f shell of the activator. Thus, the parameter λ contains information about delocalization of electrons. An analogous mechanism of electron delocalization was probably investigated in Ref. 26, but without an account of the spin–orbit interaction and only for the case of crystal fields of cubic symmetry.

3. COMPARISON WITH EXPERIMENT

3.1. $LiYF_4:Pr^{3+}$

A detailed study of the system $LiYF_4:Pr^{3+}$ (with symmetry S_4) in the weak configurational interaction approxima-

TABLE I. Comparison of experimental[27] and calculated energy levels in the irreducible representations Γ_i of the Pr^{3+} ion in $LiYF_4$.

Multiplet	Experiment[27]		Calculated in the interconfigurational interaction approximation			
			a) weak (1)		*b*) intermediate (12)	
SLJ	Γ	E, cm^{-1}	Γ	E, cm^{-1}	Γ	E, cm^{-1}
$^3H_4$	2	0	2	8.2	2	−4.7
	3,4	79	3,4	85.3	3,4	79.9
	–	–	1	217.9	1	236.3
	1	220	1	219.0	1	237.2
	3,4	496	3,4	487.8	3,4	500.7
	–	–	1	512.5	1	512.5
	–	–	2	514.5	2	540.5
$^3H_5$	1	2253	1	2244.9	1	2236.3
	3,4	2272	3,4	2253.8	3,4	2245.1
	2	2280	2	2275.6	2	2252.8
	1	2297	1	2276.7	1	2278.0
	3,4	2341	3,4	2326.3	3,4	2327.1
	2	2549	2	2557.1	2	2565.7
	–	–	1	2578.7	1	2566.7
	–	–	3,4	2597.8	3,4	2597.8
$^3H_6$	2	4314	2	4321.4	2	4301.5
	3,4	4394	3,4	4421.0	3,4	4410.6
	–	–	1	4441.7	1	4433.8
	–	–	2	4470.5	2	4479.6
	3,4	4454	3,4	4487.9	3,4	4479.6
	1	4486	1	4523.0	1	4539.8
	2	4557	2	4570.2	2	4562.0
	–	–	1	4891.2	1	4890.6
	3,4	4907	3,4	4894.1	3,4	4900.7
	2	4945	2	4937.6	2	4957.5
$^3F_2$	–	–	1	5171.3	1	5162.0
	2	5201	2	5247.4[b]	2	5227.7
	3,4	5221	3,4	5230.2[b]	3,4	5227.7
	2	5342	2	5332.8	2	5315.3
$^3F_3$	3,4	6481	3,4	6463.9	3,4	6447.1
	2	6521	2	6512.6	2	6499.2
	1	6586	1	6547.9	1	6553.9
	3,4	6671	3,4	6659.0	3,4	6652.5
	2	6686	2	6703.1	2	6719.9
$^3F_4$	1	6920	1	6898.3	1	6883.1
	3,4	6942	3,4	6905.1	3,4	6904.1
	2	6983	2	6943.0	2	6950.7
	1	7105	1	7120.1[b]	1	7119.7
	2	7116	2	7109.7[b]	2	7136.5
	3,4	7142	3,4	7129.3	3,4	7154.2
	1	7220	1	7241.7	1	7256.9
$^1G_4$	1	9699	1	9702.5	1	9679.5
	3,4	9832	3,4	9802.2	3,4	9809.0
	2	9930	2	9918.0	2	9899.4
	2	10011	2	10007.5	2	10030.5
	3,4	10112[a]	3,4	10157.2[b]	3,4	10198.1[b]
	1	10217[a]	1	10126.7[b]	1	10136.4[b]
	1	10313[a]	1	10578.8	1	10643.1
$^1D_2$	2	16740	2	16865.6[b]	2	16757.2
	1	16810	1	16814.5[b]	1	16814.0
	3,4	17083	3,4	17077.7	3,4	17033.3
	2	17406	2	17401.5	2	17388.8
$^3P_0$	1	20860	1	20860.0	1	20860.0
$^1I_6$	–	–	2	21083.0	2	21175.1
	–	–	2	21083.3	2	21180.2
	–	–	3,4	21400.2	2	21306.2
	–	–	1	21413.1	3,4	21307.3
	–	–	2	21414.4	1	21331.1
$^3P_1$	–	–	3,4	21442.9	3,4	21464.8
	–	–	1	21610.8	1	21591.9
$^1I_6$	–	–	3,4	21622.1	3,4	21665.7
	–	–	1	21758.6	1	21745.2
	–	–	1	22032.2	3,4	22078.7
	–	–	3,4	22043.7	2	22084.7
	–	–	2	22054.5	1	22090.5

tion (1) was presented in Ref. 27 where it was established that the ordering of the calculated energy levels corresponding to the experimental levels with $E=5201$ and $5221\,cm^{-1}$ of the $^3F_2$ multiplet, 7105 and $7116\,cm^{-1}$ of the $^3F_4$ multiplet, 10112 and $10217\,cm^{-1}$ of the $^1G_4$ multiplet, and 16740 and $16810\,cm^{-1}$ of the $^1D_2$ multiplet is inverted. In addition, the calculated splitting of the $^1G_4$ multiplet obtained is 1.5 times smaller than the experimentally measured value.

Results of our calculations based on this approximation are listed in column *a* of Table I. The value of the standard deviation, which is somewhat different than in Ref. 27, is due to the different manner of choosing the centroids (center-of-mass energies) for the multiplets. The centroids were chosen such that the sum of the standard deviations of the theoretical values from the experimental values for the highest and the lowest level of each multiplet was equal to zero.

The results obtained in the intermediate configurational interaction approximation with allowance for the ''quadratic crystal field'' (12) are listed in column *b* of Table I. In this approximation the theory fails to give the correct arrangement of levels for the $^1G_4$ multiplet. Theory *b*) provides a significantly better description of splitting of the $^1D_2$ multiplet. However, on the whole over all levels the value of the standard deviation is only slightly less than in theory *a*). If we ignore information about the irreducible representations of each level, then theory *b*) leads to a substantially smaller value of the standard deviation.[11] Comparison of the results of *b*) in Table I with the calculation of the energy spectrum based on the Hamiltonian (12), but without the ''quadratic crystal field,'' shows that the ''quadratic crystal field'' creates additional splitting of the multiplets on the order of $2\,cm^{-1}$. This is much less than the observed multiplet splitting, and for rare-earth ions it is entirely possible to use expression (3) as a simplified variant of the Hamiltonian (12).

Theories *a*) and *b*) predict a different arrangement of some levels of the $^1I_6$ multiplet. However, an experimental determination of the energies of these levels is hindered. Therefore, it is hard to conclude which theory gives the more accurate predictions.

It should be noted that a similar effect of the excited configurations on the multiplet splitting was investigated in Refs. 28 and 29 by diagonalization of the matrix of the Hamiltonian (1) in the basis of states of the configurations

TABLE I. (*continued*)

Multiplet	Experiment[27]		Calculated in the interconfigurational interaction approximation			
			a) weak (1)		*b*) intermediate (12)	
SLJ	Γ	E, cm^{-1}	Γ	E, cm^{-1}	Γ	E, cm^{-1}
$^3P_2$	1	22498	1	22507.3	1	22512.9
	3,4	22645	3,4	22635.7	3,4	22630.1
	–	–	2	22679.5	2	22630.9
	–	–	2	22776.5	2	22760.3
$^1S_0$	1	–	1	48831.0	1	46635.0
σ[c]				28.4		27.6
Parameters B_q^k (in cm^{-1}) and G_q^k (in 10^{-4}, dimensionless)						
B_0^2				488.9		456.4
B_0^4				−1043		−1087
B_4^4				1242		1348
B_0^6				−42		−9.4
B_4^6				1213		1284
Im B_4^6				22.5		53.2
G_0^2						−22.3
G_0^4						35.2
G_4^4						43.7
G_0^6						53.3
G_4^6						7.8
Im G_4^6						−25.0
						$F_q^k = B_q^k$

a—levels not included in fitting procedure.
[b]—inverted level.
[c]—$\sigma = (\Sigma_{i=1}^{n}[E_{\text{exp}}(i) - E_{\text{calc}}(i)]^2/(n-p))^{1/2}$ where E_{exp} and E_{calc} are respectively the experimental and calculated levels; p is the nuber of fitting parameters.

$4f^2$, $4f^15d$, $4f^16s$, and $4f^16p$. However, the application of the Hamiltonian (12) toward these ends is to be preferred since in the method of the effective Hamiltonian the common set of parameters G_q^k can easily account for the effect of all configurations corresponding to one-electron excitations.

3.2. $Y_3Al_5O_{12}:Tm^{3+}$

A detailed experimental and theoretical study of the Tm^{3+} ion in $Y_3Al_5O_{12}$ (with symmetry D_2) was given in Ref. 30. The results of our calculations in the weak interconfigurational interaction approximation with the parameters of Ref. 30 are listed in column *a*) of Table II.[7] Here, as in Subsection 3.1., because of the special choice of the energy of the centroids the standard deviation differs insignificantly from that obtained in Ref. 30. The large value of the standard deviation 18.4 cm^{-1}, and the incorrect arrangement of the level corresponding to 252 cm^{-1} are proof of the inadequacy of the weak configurational interaction approximation.

In the intermediate configurational interaction approximation the description is improved: the correct arrangement of all levels is achieved and the standard deviation is reduced to 11.8 cm^{-1} (see Table II). For this system the ''quadratic crystal field'' also causes an insignificant additional splitting of the multiplets, around 3 cm^{-1}, i.e., to describe the spectra of the Ln^{3+} ion it is indeed possible to use the simplified variant (3) of the Hamiltonian (12).

An experimental determination and identification of levels of the Tm^{3+} ion encounters significant difficulties. It is probably for that reason that the authors of Ref. 31 proposed a new scheme of energy levels for this ion, which differs substantially from the scheme in Ref. 30. Application of the Hamiltonian (12) also improves the description of this new spectrum, lowering the standard deviation by 26%. In an earlier paper[32] they proposed another variant of the spectrum, but to use this spectrum to test the theory is not well motivated, because of its dubious realism.

TABLE II.

Multiplet	Experiment[30]		Calculated in the interconfigurational interaction approximation			
			a) weak (1)		*b*) intermediate (12)	
SLJ	Γ	E, cm^{-1}	Γ	E, cm^{-1}	Γ	E, cm^{-1}
$^3H_6$	2	0	2	−11.9	2	5.5
	1	27	1	15.2	1	23.1
	4	216	4	206.3	4	218.8
	3	241	3	215.7	3	241.9
	–	247	2	253.7[a]	2	247.6
	1	252	1	244.2[a]	1	262.8
	–	588	4	508.5	4	562.9
	–	610	3	601.3	3	615.8
	–	–	1	640.1	1	682.2
	–	690	2	674.8	2	687.1
	–	–	4	689.4	4	693.6
	–	730	3	741.9	3	724.5
	–	–	1	754.9	1	737.9
$^3F_4$	1	5556	1	5541.3	1	5537.9
	3	5736	3	5762.4	3	5753.6
	2	5832	2	5815.4	2	5814.6
	4	5901	4	5917.5	4	5910.2
	1	6041	1	6045.0	1	6035.7
	2	6108	2	6116.5	2	6097.3
	1	6170	1	6168.5	1	6182.8
	–	6224	4	6232.9	4	6223.4
	–	6233	3	6247.7	3	6251.1
$^3H_5$	4	8339	4	8340.2	4	8340.7
	3	8345	3	8349.6	3	8351.6
	3	8516	3	8504.0	3	8510.1
	1	8530	1	8513.4	1	8520.1
	–	–	2	8520.4	2	8531.0
	4	8556	4	8555.0	4	8557.3
	2	8711	2	8708.4	2	8725.0
	–	8773	1	8770.9	1	8770.8
	–	8800	3	8800.4	3	8807.9
	–	–	2	8869.0	2	8867.8
	–	8882	4	8880.8	4	8880.3

4. CALCULATION OF THE PARAMETERS G_q^k

The dimensionless parameters G_q^k prescribe the amplitude of the interconfigurational interaction, which previously was not taken into account in the description of the experimental data. Therefore it is of interest to estimate these parameters from microscopic arguments. It may be expected that the largest contribution to G_q^k comes from lower excited configurations of the type $4f^{N-1}5p^5$, the configuration with electron transfer from the ligand to the $4f$ shell, and the configurations $4f^{N-1}5l$ $(l=d,g)$.

The contribution of the excited configuration $4f^{N+1}5p^5$ can be estimated from formula (13), assuming that

TABLE II. (*continued*)

Multiplet	Experiment[30]		Calculated in the interconfigurational interaction approximation			
			a) weak (1)		*b*) intermediate (12)	
SLJ	Γ	E, cm^{-1}	Γ	E, cm^{-1}	Γ	E, cm^{-1}
$^3H_4$	1	12607	1	12610.2	1	12611.1
	2	12679	2	12674.2	2	12677.9
	4	12747	4	12745.0	4	12744.6
	–	–	2	12814.8	2	12801.7
	3	12824	3	12829.3	3	12833.8
	–	–	1	12951.5	1	12949.5
	4	13072	4	13066.7	4	13067.5
	3	13139	3	13124.1	3	13129.9
	–	13159	1	13155.8	1	13154.9
$^3F_3$	4	14659	4	14652.9	4	14654.5
	–	–	2	14658.0	2	14654.5
	–	–	2	14670.1	2	14667.9
	3	14679	3	14693.0	3	14693.8
	–	14705	4	14719.5	4	14714.3
	–	14720	3	14740.1	3	14743.6
	1	14741	1	14747.1	1	14745.5
$^3F_2$	3	15245	3	15244.5	3	15244.8
	4	15264	4	15258.7	4	15262.9
	–	–	1	15300.6	1	15305.6
	–	–	2	15430.3	2	15436.9
	1	15438	1	15438.5	1	15438.2
$^1G_4$	1	20805	1	20805.5	1	20804.4
	–	–	2	21181.7	2	21186.1
	3	21227	3	21214.9	3	21240.1
	4	21381	4	21376.9	4	21378.2
	1	21530	1	21502.6	1	21524.7
	2	21687	2	21671.8	2	21698.1
	1	21757	1	21756.5	1	21757.6
	–	–	4	21813.6	3	21839.7
	–	–	3	21853.6	4	21839.8
$^1D_2$	1	27868	1	27895.1	1	27885.4
	3	27877	3	27921.8	3	27906.7
	2	28023	2	28008.3	2	28014.4
	4	28044	4	28030.0	4	28037.4
	1	28075	1	28047.9	1	28057.6
$^1I_6$	1	34391	1	34384.4	1	34390.8
	4	34422	4	34428.6	4	34423.9
	3	34440	3	34450.9	3	34446.6
	2	34449	2	34454.4	2	34456.7
	4	34520	4	34526.5	4	34542.1
	–	–	3	34683.8	3	34709.7

TABLE II. (*continued*)

Multiplet	Experiment[30]		Calculated in the interconfigurational interaction approximation			
			a) weak (1)		*b*) intermediate (12)	
SLJ	Γ	E, cm^{-1}	Γ	E, cm^{-1}	Γ	E, cm^{-1}
	1	34748	1	34730.0	1	34742.1
	1	35033	1	35039.6	1	35033.2
	–	–	2	35039.8	2	35035.9
	–	–	2	35206.4	2	35237.0
	–	–	4	35230.1	4	35253.3
$^3P_0$	1	35372	1	35372.0	1	35372.0
$^1I_6$	–	–	3	35390.0	3	35455.8
	–	–	1	35401.9	1	35462.5
$^3P_1$	3	36234	3	36249.0	3	36234.0
	4	36391	4	36376.4	4	36405.9
	2	36418	2	36403.0	2	36418.0
$^3P_2$	3	37932	3	37948.9	3	37925.8
	4	38066	4	38015.0	4	38046.0
	1	38098	1	38101.8	1	38097.5
	2	38398	2	38414.3	2	38402.1
	1	38440	1	38423.1	1	38446.2
$^1S_0$	1	–	1	79604.0	1	79604.0
σ				18.4		11.8

Parameters B_q^k (in cm^{-1}) and G_q^k (in 10^{-4}, dimensionless)		
B_0^2	474	392
B_2^2	47.0	103
B_0^4	−213	−82.7
B_2^4	−1571	−1634
B_4^4	−824	−835
B_0^6	−984	−884
B_2^6	−310	−409
B_4^6	591	493
B_6^6	−193	−153
G_0^2		29.0
G_2^2		−11.9
G_0^4		−27.3
G_2^4		20.5
G_4^4		−2.3
G_0^6		−54.6
G_2^6		43.2
G_4^6		27.3
G_6^6		−20.5
		$F_q^k = G_q^k$

Remark.[a]—inverted levels.

$$B_q^k(5p) \approx \frac{\langle 5p|r^k|4f\rangle}{\langle 4f|r^k|4f\rangle} B_q^k(4f). \tag{16}$$

The necessary integrals (in atomic units)

$$\langle 4f|r^2|4f\rangle = 1.064, \quad \langle 5p|r^2|4f\rangle = 1.415,$$

$$\langle 4f|r^4|4f\rangle = 2.623, \quad \langle 5p|r^4|4f\rangle = 5.769 \tag{17}$$

were calculated in the $5p$ and $4f$ functions of the Pr^{3+} ion from Ref. 33. Next, using the value $\Delta(5p) = 60000\,\text{cm}^{-1}$ (Ref. 33), it is possible to obtain the following values, e.g., for the contributions to G_q^k:

$$10^4 G_0^4(5p) = -1.5, \quad 10^4 G_0^6(5p) = -0.9, \ldots \tag{18}$$

Their magnitudes are much smaller than the values obtained in Table I from the experimental data, i.e., admixture of the configuration $4f^{N+1}5p^5$ gives a negligibly small contribution to G_q^k.

In this situation it may be expected that processes involving electron transport from the ligand to the f shell of the rare-earth ion will be especially important. The value of the contributions to G_q^k from such processes can be estimated with the aid of expression (15) after the following transformation:

$$G_q^k(cov) = \sum_b J^k C_q^{k*}(\Theta_b, \Phi_b), \tag{19}$$

where

$$J^2=\frac{5}{28}[2(\lambda_{\sigma f}^2+\lambda_{sf}^2)+3\lambda_{\pi f}^2],$$

$$J^4=\frac{3}{14}[3(\lambda_{\sigma f}^2+\lambda_{sf}^2)+\lambda_{\pi f}^2],$$

$$J^6=\frac{13}{28}[2(\lambda_{\sigma f}^2+\lambda_{sf}^2)-3\lambda_{\pi f}^2]. \quad (20)$$

The covalency parameters λ can be found in Refs. 34 and 35:

$$\lambda_{sf}=0.02, \quad \lambda_{\sigma f}=-0.05, \quad \lambda_{\pi f}=0.04.$$

After summing over nearest neighbors of the Pr^{3+} ion, we obtain

$$10^4G_0^2(cov)=9.5, \quad 10^4G_0^4(cov)=-32.8,$$

$$10^4G_4^4(cov)=-30.8, \quad 10^4\,\mathrm{Im}\,G_4^4(cov)=-26.1,$$

$$10^4G_0^6(cov)=-1.5, \quad 10^4G_4^6(cov)=-5.1;$$

$$10^4\,\mathrm{Im}\,G_4^4(cov)=-6.7. \quad (21)$$

These values are found to be in satisfactory order-of-magnitude agreement with the experimental values in Table I. For these values of G_q^k the influence of the interconfigurational interaction on the multiplet splitting is substantial and the fact that it is taken into account in formula (12) improves the description of the energy spectrum of the Ln^{3+} ion in crystals.

However, there is one circumstance that indicates that besides charge-transfer processes, excited configurations of opposite parity can also play an important role. Indeed, according to Eqs. (20), the quantities J^2 and J^4 are formed by the sum of squares of covalency parameters while the quantity J^6 is formed by a difference. Therefore, J^6 should be significantly smaller in magnitude than J^2 and J^4 and, consequently, $G_q^6(cov)$ should be much smaller than $G_q^k(cov)$ for $k=2,4$. This conclusion correlates poorly with the parameters G_q^k in Tables I and II. This immediately confirms that the admixture of configurations of opposite parity can make a significant contribution to G_q^k.

Quite detailed information about the parameters of an odd crystal field and the excitation energies for the Y_2O_3 crystal is given in Ref. 36. Employing this information, we convinced ourselves that the excited configuration $4f^{N-1}5g$ gives a 5–10-times smaller contribution to G_q^k than $4f^{N-1}5d$ and the effects of admixture of the configuration $4f^{N-1}5g$ can be neglected. If we exclude covalent contributions (21) from G_q^k in Table I, then according to Eq. (13) a one-to-one correspondence will be observed between the parameters of the odd crystal field and the parameters obtained in this way $G_q^k(d)$. These values of $G_q^k(d)$ correspond to the following parameters of the odd crystal field:

$$B_2^3(d)=1248, \quad \mathrm{Im}\,B_2^3(d)=-2247,$$

$$B_2^5(d)=16610, \quad \mathrm{Im}\,B_2^5(d)=405. \quad (22)$$

They are of a similar order of magnitude as the parameters (in cm^{-1})

$$B_2^3(d)=611, \quad \mathrm{Im}\,B_2^3(d)=-620,$$

$$B_2^5(d)=-8334, \quad \mathrm{Im}\,B_2^5(d)=-185, \quad (23)$$

which we obtained from the quantities $R(k)$ calculated in Ref. 27. A similar correlation is observed between the parameters of an even crystal field calculated on the basis of microscopic models and obtained from the experimental data. From this point of view, parameters (22) are entirely realistic. Thus, the combination of Eqs. (12) and (13) makes it possible on the basis of an analysis of the Stark multiplet structure to obtain information about the parameters of an odd crystal field responsible for the admixture of excited configurations of opposite parity and, consequently, for the intensity characteristics of absorption and luminescence.

5. CONCLUDING REMARKS

Excited configurations influence high- and low-lying multiplets to a substantially different degree. This is the cause of the dependence of the crystal-field parameters and the intensity parameters on the energy of the multiplets. For Ln^{3+} ions the most acceptable approximation is the intermediate configurational interaction approximation, in which the parameters $\tilde{B}_q^k$ depend on the multiplet energy linearly. In the case of the actinides the more adequate approximation is the strong configurational interaction approximation with more-complicated-than-linear dependence of the parameters $\overline{B_q^k}$.

Microscopic estimates of the parameters prescribing the amplitude of the interconfigurational interaction are in satisfactory agreement with the experimental values, which is an indication of the realism of the proposed crystal-field Hamiltonians.

The new functional dependence of the crystal-field parameters makes it possible for the first time to determine the parameters of an odd crystal field from the results of an analysis of the structure of the Stark levels. These parameters can then be used to calculate the intensity characteristics of the absorption and luminescence of laser crystals.

This work was carried out with the partial financial support of the Russian Fund for Fundamental Research and the State Scientific Program ''Fundamental Spectroscopy,'' and also of CNRS/RFFI (Grant No. 97-02-16040). The authors note that their corporation in the Joint Open Laboratory ''Laser Crystals and Precision Laser Systems'' facilitated in a substantial way the completion of the studies reported here.

*)E-mail: kaminalex@public.mtu

[1] D. J. Newman, Adv. Phys. **20**, 197 (1971).
[2] M. V. Eremin and A. A. Kornienko, Phys. Status Solidi B **79**, 775 (1977).
[3] G. G. Siu and D. J. Newman, J. Phys. C: Solid State Phys. **16**, 7019 (1983).
[4] B. R. Judd, Phys. Rev. Lett. **39**, 242 (1977).
[5] B. R. Judd and H. Crosswhite, J. Opt. Soc. Am. B **1**, 255 (1984).
[6] M. F. Reid, J. Chem. Phys. **87**, 2875 (1987).
[7] K. Jankowski and L. Smentek-Mielczarek, Mol. Phys. **38**, 1445 (1979).
[8] B. G. Wybourne, *Spectroscopic Properties of Rare-Earths* (Wiley, New York, 1965).
[9] B. R. Judd, Phys. Rev. **127**, 750 (1962).

[10] G. S. Ofelt, J. Chem. Phys. **37**, 511 (1962).
[11] A. A. Kornienko and E. B. Dunina, JETP Lett. **59**, 412 (1994).
[12] A. A. Kornienko, A. A. Kaminskii, and E. B. Dunina, Phys. Status Solidi B **157**, 267 (1990).
[13] A. A. Kornienko, E. B. Dunina, and V. L. Yankevich, Zh. Prikl. Spektrosk. **63**, 1003 (1996).
[14] A. A. Kornienko, E. B. Dunina, V. L. Yankevich, Opt. Spektrosk. **81**, 951 (1996) [Opt. Spectrosc. **81**, 871 (1996)].
[15] P. Goldner and F. Auzel, J. Appl. Phys. **79**, 7972 (1996).
[16] M. A. Bunuel, R. Cases, M. A. Chamarro, and R. Alcalá, Phys. Chem. Glasses **33**, 16 (1992).
[17] K. Suzuki and R. Okamoto, Prog. Theor. Phys. **70**, 439 (1983).
[18] A. A. Kornienko, A. A. Kaminskii, and E. B. Dunina, Phys. Status Solidi B **157**, 261 (1990).
[19] A. A. Kaminskii, A. A. Kornienko, and M. I. Chertanov, Phys. Status Solidi B **134**, 717 (1986).
[20] A. A. Kornienko, A. A. Kaminskii, and E. B. Dunina, Phys. Status Solidi B **178**, 385 (1993).
[21] O. A. Anikeenok, M. V. Eremin, Fiz. Tverd. Tela (Leningrad) **23**, 706 (1981) [Sov. Phys. Solid State **23**, 401 (1981)].
[22] M. V. Eremin, Fiz. Tverd. Tela (Leningrad) **24**, 423 (1982) [Sov. Phys. Solid State **24**, 239 (1982)].
[23] R. W. Stevens, in *Proceedings of the Second International Conference on Crystal Field Effects in Metals and Alloys*, Zürich (1976), p. 1.
[24] A. Messiah, *Quantum Mechanics* (Wiley, New York, 1962).
[25] M. V. Eremin, Fiz. Tverd. Tela (Leningrad) **21**, 4634 (1979) [Sov. Phys. Solid State **21**, 2097 (1979)].
[26] B. R. Judd, J. Chem. Phys. **66**, 3163 (1977).
[27] L. Esterowitz, F. J. Bartoli, R. E. Allen *et al.*, Phys. Rev. B **19**, 6442 (1979).
[28] D. Garcia and M. Faucher, J. Chem. Phys. **90**, 5280 (1989).
[29] D. Garcia and M. Faucher, J. Chem. Phys. **91**, 7461 (1989).
[30] J. B. Gruber, M. E. Hills, R. M. Macfarlane *et al.*, Phys. Rev. B **40**, 9464 (1989).
[31] C. Tiseanu, A. Lupei, and V. Lupei, J. Phys.: Condens. Matter **7**, 8477 (1995).
[32] A. Lupei, C. Tiseanu, and V. Lupei, Phys. Rev. B **47**, 14084 (1993).
[33] M. Synek and W. Timmons, Phys. Rev. **185**, 38 (1969).
[34] O. A. Anikeenok, M. V. Eremin, V. P. Meiklyar *et al.*, J. Phys. C: Solid State Phys. **15**, L105 (1982).
[35] O. A. Anikeenok, M. V. Eremin, M. L. Falin *et al.*, J. Phys. C: Solid State Phys. **17**, 2813 (1984).
[36] W. F. Krupke, Phys. Rev. **145**, 325 (1966).

Translated by Paul F. Schippnick

Nonlinear dynamics of vortices pinned to unidirectional twins

V. A. Shklovskiĭ*) and A. K. Soroka

Khar'kov State University, 310077 Khar'kov, Ukraine

A. A. Soroka

National Scientific Center ''Khar'kov Physicotechnical Institute,'' Institute of Theoretical Physics, 310108 Khar'kov, Ukraine

(Submitted 14 May 1999)
Zh. Éksp. Teor. Fiz. **116**, 2103–2129 (December 1999)

The nonlinear resistive properties of superconductors in the mixed state in the presence of a system of unidirectional planar defects (twins) have been investigated theoretically within the framework of the two-dimensional stochastic model of anisotropic pinning based on the Fokker–Planck equations with a concrete form of the pinning potential. These equations allow one to obtain an exact analytical solution of the problem. Formulas are obtained for experimentally observable even and odd (relative to reversal of the direction of the external magnetic field) nonlinear longitudinal and transverse magnetoresistivities $\rho_{\parallel,\perp}^{\pm}(j,t,\alpha,\varepsilon)$ as functions of the transport current density j, temperature t, the angle α between the directions of the current and the twins, and the relative volume fraction ε occupied by the twins. In light of the great variety of types of nonlinear resistive dependences contained in these expressions for $\rho_{\parallel,\perp}^{\pm}$, the most characteristic of them are presented in the form of graphs with commentary. The desired nonlinear dependences $\rho_{\parallel,\perp}^{\pm}$ are linear combinations of the even and odd parts of the function $\nu(j,t,\alpha,\varepsilon)$, which has the sense of the probability of overcoming the potential barrier of the twins; this makes it possible to give a simple physical treatment of the nonlinear regimes. New scaling relations for the Hall conductivity are obtained and investigated which differ from the previously known relations for isotropic pinning. The interaction of vortex motion directed along the twins and the Hall effect is considered for Hall constants which are arbitrary in magnitude and sign, and it is shown that in the case of small Hall viscosity vortex motion directed along the twins has an effect on the odd magnetoresistivities $\rho_{\parallel}^{-}$ and $\rho_{\perp}^{-}$, whereas the reverse effect can be neglected. It is shown that pinning anisotropy is sufficient to manifest the new nonlinear (in the current) magnetoresistivities $\rho_{\perp}^{+}$ and $\rho_{\parallel}^{-}$. © *1999 American Institute of Physics.*
[S1063-7761(99)01812-0]

1. INTRODUCTION

Recently, an ever increasing number of experimental and theoretical works have appeared[1–19] investigating the influence of systems of unidirectional planar defects on the anisotropy of the resistive behavior of superconductors in the mixed state. The urgency of this problem is connected with two circumstances. First, in all of the currently known high-temperature superconductors (HTSC's) (e.g., based on Y and La) during growth of the crystals twins appear with their planes oriented parallel to the **c** axis. If such twins are unidirectional (i.e., the twinning planes are parallel) but the transport current **j** flows entirely in the plane of the crystalline layers, then, as experiment has shown,[1,4] the longitudinal and transverse magnetoresistivities of the sample in an external magnetic field $\mathbf{H}\|\mathbf{c}$ depend substantially on the angle α with which the vector **j** intersects the boundaries of the twins. Second, in layered anisotropic HTSC's the system of parallel ab planes itself can be considered as a set of unidirectional planar defects, which are a source of the so-called intrinsic[11] pinning for the vortices located parallel to the ab planes and subjected to the action of the Lorentz force directed along the **c** axis.

From the theoretical point of view, the two indicated cases differ only in the specific form of the pinning potential and, as the following analysis shows, the qualitative conclusions in both cases coincide. However, the difficulties of preparing samples with the required configuration of transverse and longitudinal contacts to realize various angles α in these two cases are substantially different. Thus, to prepare samples with twins from thin single crystals[1,4–6] or YBaCuO epitaxial films[19] is technically much simpler than in the case when it is necessary to pass a transport current at a certain angle to the ab planes of the film or crystal of YBaCuO (so far we know of only one attempt of this kind[20]). Therefore, in what follows the main illustrative material that we provide in this paper (in the form of graphs of evolutions of the current and temperature dependence of the longitudinal and transverse magnetoresistivity of a sample with variation of the size of the angle α (see Figs. 8–13 below) is only in regard to the case of pinning on twins.

From a more general point of view, the influence of the twin boundaries on the transport properties of HTSC's is a subject of increasingly focused attention.[1–19] One of the main reasons for such interest has to do with the reliably established fact[11] that the order parameter in twins in HTSC's is easily suppressed. As a consequence of this, an isolated twin boundary attracts vortices and pins them.[11] The corresponding pinning force acting on vortices directed along the **c** axis, by virtue of the geometry of the problem, is perpendicular to the twins and, as a consequence, its direction depends on the angle α between the vector **j** and the direction of the twins. Hence it also follows that the twins should have almost no effect on the dynamics of the vortices, whose velocity is parallel to their planes (in contrast to point defects, for which the direction of the pinning force is opposite the direction of motion of the vortex).

Recently there have appeared several works in which the influence of the twins on the vortex dynamics in a planar geometry was examined with the aid of numerical simulation.[12–14] Two of these studies[12,13] examined the interaction of moving vortices with an isolated twin boundary and investigated a number of interesting dynamical peculiarities of this interaction. It should be noted, however, that in a real transport experiment[1,4–6] we are usually studying the self-averaged dynamics of vortices, which is a consequence of the interaction of vortices with many twin boundaries distributed with some mean density between the measuring contacts. Such self-averaging apparently ''smears out'' some of the fine features of the interaction of the vortices with an isolated twin boundary that were detected in Refs. 12 and 13.

A different model was investigated in a recent paper.[14] In this paper, the equation of motion of a vortex takes account of the interaction of the vortices with each other, and also with point defects and planar defects in the presence of thermal fluctuations. However, the large number of independent physical parameters in the investigated problem greatly hinder the choice of their values, whose variation leads to nontrivial physical results. The latter can be summarized as saying that the vortex dynamics depend substantially on the orientation of the twin boundaries relative to the direction of the transport current and also on its magnitude.

The most specific manifestation of pinning anisotropy in twins are effects associated with directed motion of vortices along twins (guided motion, the G effect),[6] where there is a greater probability of the vortices moving probability parallel to the twins than overcoming the twin barriers (slipping, the S effect). For $\alpha \neq 0,\ \pi/2$ the G effect leads to the appearance of an even (relative to reversal of the direction of the external magnetic field) component of the transverse magnetoresistivity of the sample, $\rho_{\perp}^{+}$, which usually is significantly greater in magnitude than the familiar odd Hall component $\rho_{\perp}^{-}$.

Earlier, a number of experimental and some theoretical aspects of anisotropic pinning and guided motion of vortices in the flux-flow (FF) regime in samples of the Nb–Ta alloy subjected to cold rolling were discussed in detail by Niessen and Weijsenfeld.[21] Interest in these questions has grown substantially since the discovery of twins in single crystals based on Y and La. Besides the experimental works[4–9] we should also mention in this regard the recent theoretical work of

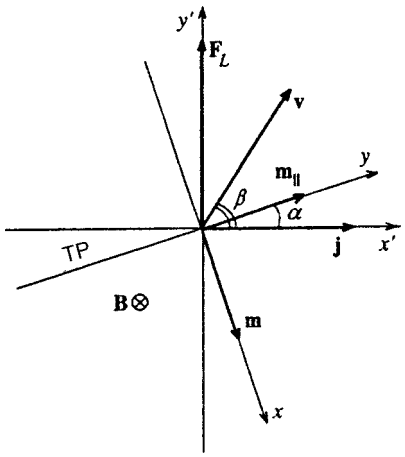

FIG. 1. System of coordinates xy associated with the twins (the anisotropy vector **m** points along the x axis) and the system of coordinates $x'y'$ associated with the direction of the current (the current density vector **j** points along the x' axis); α is the angle between the twin plane (TP) and the current density vector **j**, β is the angle between the velocity vector of the vortices **v** and the current density vector **j**; $\mathbf{F}_L$ is the Lorentz force.

Mawatari,[16] which discusses the dynamics of anisotropic pinning within the framework of a stochastic approach based on the Fokker–Planck equation. Although the general scheme of this approach also includes a derivation of the nonlinear current–voltage characteristics (CVC's) of samples with planar defects, Ref. 16 investigated only linear regimes of motion of vortices.

Note that the specifics of anisotropic pinning created by unidirectional twins do not reduce to just a representation of their role in the formation of guided vortex motion (as was done in Refs. 6, 12–14) but also includes, as will be shown below, the subtler question of the influence of the twins on the Hall effect. On the phenomenological level, this aspect of anisotropic pinning, as far as we know, was first discussed in Ref. 9 (see also Ref. 10). In Ref. 9, on the basis of general arguments about the number and symmetry of invariants for an anisotropic conductor in the case where all the currents lie in the xy plane perpendicular to the magnetic field directed along the z axis, the authors postulated a phenomenological expression for the linear Ohm's law in such a medium. A superconductor with uniaxial anisotropy caused, for example, by a system of unidirectional twins or a specially formed texture, according to the form of Ohm's law in Ref. 9 is characterized (for arbitrary values of the angle α) by only four constants: ρ_l, ρ_t, ρ_{Hl}, and ρ_{Ht}, which are components of the magnetoresistivity tensor (in the coordinate system with axes aligned with and perpendicular to the twins) and depend only on the magnitude of the magnetic field and the temperature. The physical meaning of these constants is simple: ρ_l and ρ_{Hl} are the ordinary longitudinal (even) and transverse Hall (odd) magnetoresistivities of the sample in the L geometry, where the current is perpendicular to the twins (or $\mathbf{j}\|\mathbf{m}$, where **m** is the unit vector in the direction perpendicular to the twins, Fig. 1), and ρ_t and ρ_{Ht} are the analogous quantities for the T geometry, where $\mathbf{j}\|\mathbf{m}_{\|}$, where $\mathbf{m}_{\|}$ is the unit vector pointing in the direction of the twins (Fig. 1). In fact, the main result of Refs. 9 and 10

consists in writing down formulas which allow one to express the longitudinal and transverse (relative to the direction of the transport current) magnetoresistivities of the sample $\rho_{\perp,\parallel}(\alpha)$ in terms of the constants ρ_l, ρ_t, ρ_{Hl}, and ρ_{Ht} for arbitrary values of the angle α between the vectors $\mathbf{j}$ and $\mathbf{m}_\parallel$. An analysis of Ohm's law obtained in this way shows that for $\alpha \neq 0$, $\pi/2$ the observed resistive response contains not only the ordinary longitudinal $\rho_\parallel^+(\alpha)$ and transverse $\rho_\perp^-(\alpha)$ magnetoresistivity [even $(+)$ and odd $(-)$] relative to reversal of the direction of the magnetic field, but also two new components induced by the pinning anisotropy: an even transverse component $\rho_\perp^+(\alpha)$ and an odd longitudinal component $\rho_\parallel^-(\alpha)$. If the physical origin of the component $\rho_\perp^+(\alpha)$ is related in an obvious way with the possibility of guided vortex motion along twins, then the appearance of an odd longitudinal contribution is associated with anisotropy of the Hall conductivity in a sample with twins, as a consequence of which in the L and T geometries $\rho_{Hl} \neq \rho_{Ht}$. Recently,[5] the magnetoresistivity $\rho_\parallel^-(H)$ was observed experimentally for the first time in a single crystal of $YBa_2Cu_3O_{7-\delta}$ with twins oriented at an angle $\alpha = 45°$ relative to the direction of the transport current, i.e., in the case where the quantity $\rho_\parallel^-(\alpha)$ is expected to be maximal.[9]

Later, in works of one of the authors[17,18] within the framework of a phenomenological approach, a derivation of Ohm's law (postulated earlier in Refs. 9 and 10) in terms of both linear (electronic) and nonlinear, vortex-velocity-dependent ''pinning'' viscosities was proposed. Taking into account the relative contribution of both isotropic (on point defects, Ref. 22) and anisotropic (on twins) pinning, in this approach it turned out to be possible to clarify the genesis of the four phenomenological constants ρ_l, ρ_t, ρ_{Hl}, and ρ_{Ht} introduced earlier in Refs. 9 and 10, mentioned above. It has been shown[18] that if isotropic pinning is neglected the phenomenologically introduced electronic and pinning viscosities can be ''reconstructed'' from CVC measurements in the L and T geometries, after which it is possible to predict the behavior of the quantities $\rho_{\parallel,\perp}^{\pm}(\alpha, j)$ for any α, $0 < \alpha < \pi/2$ (Ref. 18). In this approach it also turned out to be possible to interpret some general aspects of the anisotropic vortex dynamics in terms of physically simple quantities accessible to experimental measurement. In particular, the existence of scaling of the Hall conductivities was discussed, as well as some aspects of nonlinear regimes of guided vortex motion and a number of other results of a general nature.

Unfortunately, the phenomenological approach does not enable a direct theoretical calculation of the anisotropic nonlinear pinning viscosity, i.e., its dependence on current, temperature, and angle, starting from more detailed assumptions of vortex dynamics and the form of the pinning potential. Therefore, with the intention of performing such calculations in substantially nonlinear vortex-dynamics regimes we employed a stochastic pinning model (see, e.g., Refs. 23 and 24), a two-dimensional anisotropic variant of which has been proposed by Mawatari.[16] We augmented this model with an anisotropic Hall conductivity and specified the form of the pinning potential (in regard to the presence of unidirectional twins), which admits an analytical solution of all interesting effects. As a result, it became possible to derive and physical interpret quite simple formulas for the experimentally observable nonlinear longitudinal and transverse magnetoresistivities $\rho_{\parallel,\perp}^{\pm}(j,t,\alpha,\varepsilon)$ as functions of the transport current density j, temperature t, angle α, and relative volume fraction ε occupied by the twins. In the limiting case $\varepsilon = 1$ these results describe the effects of intrinsic pinning of an anisotropic layered HTSC. In light of the wide variety of types of resistive dependences contained in these expressions for $\rho_{\parallel,\perp}^{\pm}$, we tabulate the most characteristic of them in the form of graphs accompanied by descriptive commentary elucidating the physical meaning of the features of these graphs.

The organization of the article is as follows. The second section presents those general results in the stochastic model of anisotropic pinning which, for their exposition, do not require specification of the form of the pinning potential. Subsection 2.1. is introductory, it describes the Fokker–Planck method in the two-dimensional model of anisotropic pinning. Subsection 2.2. discusses the derivation of expressions for the nonlinear conductivity and resistivity tensors. Subsection 2.3. presents general results on the reconstruction of the conductivity tensor from CVC measurements and from scaling of the anisotropic Hall conductivity in the L and T geometries. The third section is fundamental. It substitutes a specific form of the pinning potential into the general formulas of the preceding section which enables in one of the limiting cases an analytical description of pinning on an isolated twin, and in the other, intrinsic pinning of the ab planes of the HTSC. Subsection 3.1. discusses the implemented pinning model and analyzes the behavior of the function $\nu(f,t,\varepsilon)$, which is the main nonlinear component of the model under discussion, where this component has the sense of the probability of overcoming the potential pinning barrier as a function of the external force f and the temperature t. Formulas are also obtained for $\rho_{\parallel,\perp}^{\pm}(j,t,\alpha,\varepsilon)$. Subsection 3.2. is dedicated to an analysis of the nonlinear G effect, and Subsection 3.3. discusses the interaction of the Hall effect and guided vortex motion in nonlinear regimes. Subsection 3.4. considers scaling relations for pinning on twins and discusses their stability with respect to small deviations of the angle α from its values adopted in the L and T geometries. Finally, the Conclusion discusses the results obtained and formulates conclusions.

2. GENERAL RESULTS IN THE STOCHASTIC MODEL OF ANISOTROPIC PINNING

2.1. The Fokker–Planck method in the anisotropic pinning model

Let us consider Mawatari's microscopic model,[16] generalized to the case in which the anisotropic Hall constant is taken into account. The Langevin equation for a vortex moving with velocity $\mathbf{v}$ in a magnetic field $\mathbf{B} = \mathbf{n}B$ ($B \equiv |\mathbf{B}|$, $\mathbf{n} = n\mathbf{z}$, $\mathbf{z}$ is the basis vector in the z direction, and $n = \pm 1$) has the form

$$\hat{\eta}\mathbf{v} + n\hat{\alpha}\mathbf{v} \times \mathbf{z} = \mathbf{F}_L + \mathbf{F}_p + \mathbf{F}_{th}, \tag{1}$$

where $\mathbf{F}_L = n(\Phi_0/c)\mathbf{j} \times \mathbf{z}$ is the Lorentz force (Φ_0 is the magnetic flux quantum, c is the speed of light, and $\mathbf{j}$ is the current density), $\mathbf{F}_p = -\nabla U_p$ is the pinning force (U_p is the

pinning potential), $\mathbf{F}_{th}$ is the thermal fluctuation force, $\hat{\eta}$ is the electronic viscosity tensor, and $\hat{\alpha}$ is the ''Hall'' tensor, describing effects associated with anisotropy of the Hall constant. If x and y are the coordinates along and transverse to the anisotropy axis (see Fig. 1), then in the xy representation the tensors $\hat{\eta}$ and $\hat{\alpha}$ are diagonal, and it is convenient to define η_0 and γ by means of the formulas

$$\eta_0=\sqrt{\eta_{xx}\eta_{yy}}, \quad \gamma=\sqrt{\eta_{xx}/\eta_{yy}},$$
$$\eta_{xx}=\gamma\eta_0, \quad \eta_{yy}=\eta_0/\gamma. \tag{2}$$

Here γ is the anisotropy parameter and η_0 is the averaged viscous friction coefficient. For $\alpha_{xx}=\alpha_{yy}=\alpha$ we regain the results of Ref. 16. The fluctuational force $\mathbf{F}_{th}(t)$ is represented by Gaussian white noise, whose stochastic properties are assigned by the relations

$$\langle F_{th,i}(t)\rangle=0, \quad \langle F_{th,i}(t)F_{th,j}(t')\rangle=2T\eta_{ij}\delta(t-t'), \tag{3}$$

where T is the temperature in energy units. Employing relations (3), we can reduce Eq. (1) to a system of Fokker–Planck equations:

$$\frac{\partial P}{\partial t}=-\nabla\mathbf{S}, \tag{4}$$

$$\hat{\eta}\mathbf{S}+n\hat{\alpha}\mathbf{S}\times\mathbf{z}=(\mathbf{F}_L+\mathbf{F}_p)P-T\nabla P, \tag{5}$$

where $P(\mathbf{r},t)$ is the probability density associated with finding the vortex at the point $\mathbf{r}=(x,y)$ at the time t, and $\mathbf{S}(\mathbf{r},t)\equiv P(\mathbf{r},t)\mathbf{v}(\mathbf{r},t)$ is the probability flux density of the vortex. Since the anisotropic pinning potential is assumed to depend only on the x coordinate and is assumed to be periodic [$U_p\equiv U_p(x)=U_p(x+a)$, where a is the period], the pinning force is always directed along the anisotropy axis x (with unit anisotropy vector $\mathbf{m}$, see Fig. 1) so that it has no component along the y axis [$F_{py}=-dU_p/dy=0$]. Thus, Eq. (5) in the stationary case for the functions $P=P(x)$ and $\mathbf{S}=(S_x(x),S_y(x))$ reduces to the equations

$$\gamma\eta_0 S_x+n\alpha_{xx}S_y\left(F_{Lx}-\frac{dU_p}{dx}\right)P-T\frac{dP}{dx}, \tag{6}$$

$$-n\alpha_{yy}S_x+\frac{\eta_0}{\gamma}S_y=F_{Ly}P. \tag{7}$$

Invoking the condition of stationarity for Eq. (4) and eliminating S_y from Eqs. (6) and (7), we obtain

$$T\frac{dP}{dx}+\left(-F+\frac{dU_p}{dx}\right)P=-\gamma\eta_0(1+\epsilon_x\epsilon_y)S_x, \tag{8}$$

where $\epsilon_x\equiv\alpha_{xx}/\eta_0$, $\epsilon_y\equiv\alpha_{yy}/\eta_0$, and $F\equiv F_{Lx}-n\gamma\epsilon_x F_{Ly}$. From the mathematical point of view, Eq. (8) is the Fokker–Planck equation of one-dimensional vortex dynamics.[23,24] Thus, the problem of two-dimensional vortex motion reduces to a one-dimensional problem, where a combination of x and y components of the Lorentz force enters as the external force:

$$F=n\frac{\Phi_0}{c}(j_y+n\gamma\epsilon_x j_x)=n\frac{\Phi_0}{c}(\cos\alpha+n\gamma\epsilon_x\sin\alpha)j. \tag{9}$$

Here α is the angle between the direction of the current and the vector $\mathbf{m}_\parallel$ directed perpendicular to the anisotropy axis (see Fig. 1). The solution of Eq. (8) for periodic boundary conditions $P(0)=P(a)$ and pinning potential of general form is

$$P(x)=\frac{\gamma\eta_0(1+\epsilon_x\epsilon_y)S_x}{T}\frac{f(a)f(x)}{f(a)-f(0)}\int_x^{x+a}\frac{d\xi}{f(\xi)}, \tag{10}$$

where $f(x)=\exp[(Fx-U_p(x))/T]$. Hence we obtain an expression for the x component of the vortex mean velocity:

$$\langle v_x\rangle=\frac{F\nu(F)}{\gamma\eta_0(1+\epsilon_x\epsilon_y)}, \tag{11}$$

where

$$\frac{1}{\nu(F)}\equiv\frac{F}{Ta(1-\exp(-Fa/T))}\int_0^a dx\int_0^a dx' \times\exp\left(-\frac{Fx}{T}\right)\exp\left[\frac{U_p(x+x')-U_p(x')}{T}\right]. \tag{12}$$

The dimensionless function $\nu(F)$ in the limit $F\to 0$ coincides with the analogous quantity introduced in Ref. 16. It has the physical meaning of the probability of the vortex overcoming the potential barrier, the characteristic value of which we denote as U_0. This can be seen by considering the limiting cases of high ($T\gg U_0$) and low ($T\ll U_0$) temperatures. In the case of high temperatures we have $\nu\approx 1$, and expression (11) corresponds to the flux-flow regime (FF regime). Indeed, in this case the influence of pinning can be neglected. In the case of low temperatures ν is a function of the current. For strong currents ($Fa\gg U_0$) the potential barrier disappears, $\nu\approx 1$, and the FF regime is realized. For weak currents ($Fa\ll U_0$) we have $\nu\sim\exp(-U_0/T)$, which corresponds to the regime of thermally activated flux flow[11] (the TAFF regime). The transition from the TAFF regime to the FF regime is associated with a lowering of the potential barrier with growth of the current.

2.2. The nonlinear conductivity and resistivity tensors

The electric field induced by a moving vortex system is

$$\mathbf{E}=\frac{1}{c}\mathbf{B}\times\langle\mathbf{v}\rangle=n\frac{B}{c}(-\langle v_y\rangle\mathbf{m}+\langle v_x\rangle\mathbf{m}_\parallel). \tag{13}$$

Since the mean velocity of a vortex is equal to

$$\langle\mathbf{v}\rangle=\frac{\int\int\mathbf{S}d^2\mathbf{r}}{\int\int P d^2\mathbf{r}}=\frac{\int_0^a\mathbf{S}(x)dx}{\int_0^a P(x)dx},$$

taking Eq. (13) into account, integration of Eq. (7) leads to a simple linear relation between the electric field components E_x, E_y and the transport current density j_x:

$$(1/\gamma)E_x+n\epsilon_y E_y=\rho_f j_x, \tag{14}$$

where $\rho_f\equiv\Phi_0 B/\eta_0 c^2$ is the average [see Eq. (2)] resistance to flux flow. It follows from Eq. (14) that the components of the conductivity tensor $\sigma_{xx}=(\gamma\rho_f)^{-1}$ and $\sigma_{xy}=n\epsilon_y/\rho_f$ obey linear scaling:

$$\sigma_{xy}/\sigma_{xx}=\rho_{yx}/\rho_{yy}=n\gamma\epsilon_y.$$

Below we will see that such simple scaling does not exist for the components σ_{yy} and σ_{yx} of the conductivity tensor.

From formulas (11), (13), and (14) we obtain the CVC in the xy coordinate system:

$$E_x=\gamma j_x-\frac{\epsilon_y\,\nu(f)f}{1+\epsilon_x\epsilon_y}, \tag{15}$$

$$E_y=\frac{n\,\nu(f)f}{\gamma(1+\epsilon_x\epsilon_y)}. \tag{16}$$

Here the dimensionless components of the electric field are measured in units of $E_0=BU_0/ca\eta_0$, and of the current, in units of $j_0=cU_0/\Phi_0 a$, and $f=Fa/U_0=nj(\cos\alpha+n\gamma\epsilon_x\sin\alpha)=nj_y+\gamma\epsilon_x j_x$. From expressions (15) and (16) we find the conductivity tensor (whose components are measured in units of $1/\rho_f$) for the nonlinear Ohm's law $\mathbf{j}=\hat{\sigma}(\mathbf{E})\mathbf{E}$:

$$\hat{\sigma}=\begin{pmatrix}\sigma_{xx} & \sigma_{xy}\\ \sigma_{yx} & \sigma_{yy}\end{pmatrix}=\begin{pmatrix}1/\gamma & n\epsilon_y\\ -n\epsilon_x & \sigma_{yy}(E_y)\end{pmatrix}, \tag{17}$$

where the only component depending on the electric field (or current), σ_{yy}, is given by

$$\sigma_{yy}(E_y)=\frac{n}{E_y}\tilde{\nu}^{-1}(n\gamma(1+\epsilon_x\epsilon_y)E_y)-\gamma\epsilon_x\epsilon_y, \tag{18}$$

$\tilde{\nu}^{-1}$ is the inverse function of $\tilde{\nu}(f)\equiv\nu(f)f$. From physical arguments it follows that the function $\nu(f)$ is monotonically increasing in f and, consequently, $\tilde{\nu}(f)$ is also monotonic and its inverse function $\tilde{\nu}^{-1}(f)$ is unique. From the definition (12) it can be shown that for a periodic potential possessing even parity, $U_p(-x)=U_p(x)$, the function $\nu(f)$ is even in f, i.e., $\nu(-f)=\nu(f)$. Correspondingly, $\tilde{\nu}(f)$ is odd in f. The resistivity tensor $\hat{\rho}$ (whose components are measured in units of ρ_f), which is the inverse tensor to $\hat{\sigma}$, has the form

$$\hat{\rho}=\begin{pmatrix}\rho_{xx} & \rho_{xy}\\ \rho_{yx} & \rho_{yy}\end{pmatrix}=\begin{pmatrix}\gamma[1-\epsilon_x\epsilon_y\nu(f)/(1+\epsilon_x\epsilon_y)] & -n\epsilon_y\nu(f)/(1+\epsilon_x\epsilon_y)\\ n\epsilon_x\nu(f)/(1+\epsilon_x\epsilon_y) & \nu(f)/[\gamma(1+\epsilon_x\epsilon_y)]\end{pmatrix}. \tag{19}$$

It is clear from Eq. (19) that all components of the tensor $\hat{\rho}$ (in contrast to the tensor $\hat{\sigma}$) are functions of the current density $\mathbf{j}$ and the angle α between the direction of the current and the vector $\mathbf{m}_\parallel$.

We introduce the L and T geometries in which $\mathbf{j}\parallel\mathbf{m}$ and $\mathbf{j}\perp\mathbf{m}$, respectively. If we neglect the Hall terms in Eq. (19), then in the L geometry vortex motion takes place along the pinning planes (the G effect), and in the T geometry—transverse to the pinning planes (the S effect). In the L geometry the critical current is equal to zero since the flux-flow regime (FF regime) is realized for guided vortex motion along pinning planes. In the T geometry, i.e., for vortex motion transverse to the pinning planes, a pronounced nonlinear regime is realized for $T\ll U_0$, the onset of which corresponds to the crossover point $j=j_{\rm cr}$, and for $T=0$ we have $j_{\rm cr}=j_c$, where j_c is the critical current.

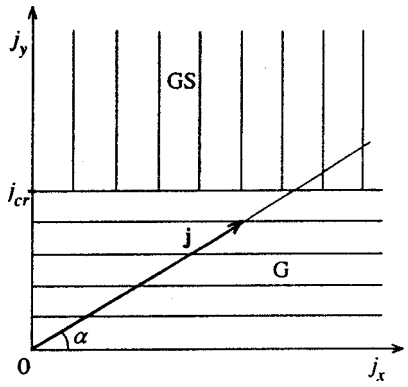

FIG. 2. Diagram of dynamic states of the vortex system neglecting the Hall effect in the (j_xj_y) plane; G is the region of motion of the vortices along the pinning plane (the G effect), GS is the region of motion of the vortices along and transverse to the pinning plane (the G and S effects); $j_{\rm cr}$ is the crossover current corresponding to a transition from the region G to the region GS when the y component of the current density is increased.

Let us consider a diagram of the dynamical states of a vortex system in the j_xj_y plane (Fig. 2). For arbitrary angle α the tip of the $\mathbf{j}$ vector can lie in two different regions which are different in their physical significance. As long as $j_y<j_{\rm cr}$ guided vortex motion takes place (the G region). For $j_y>j_{\rm cr}$ guided motion along the pinning planes is joined by motion transverse to the pinning planes (the GS region).

2.3. Reconstruction of the conductivity tensor from CVC data, and scaling of the Hall conductivity

The experimentally measured quantities are referenced to the coordinate system associated with the current (see Fig. 1). The longitudinal and transverse components of the electric field relative to the direction of the current, $E_\parallel$ and $E_\perp$, are related in a simple way to E_x and E_y:

$$E_\parallel=E_x\sin\alpha+E_y\cos\alpha, \tag{20}$$

$$E_\perp=-E_x\cos\alpha+E_y\sin\alpha. \tag{21}$$

The field $E_\parallel(j,\alpha)$ as a function of j for $\alpha=$const is monotonically increasing and reduces to $E_x(j)$ for $\alpha=\pi/2$ (the L geometry) and $E_y(j)$ for $\alpha=0$ (the T geometry). The field $E_\perp(j,\alpha)$ as a function of j for $\alpha=$const exhibits a pronounced nonlinearity and has an extremum associated with the G effect (see below).

We will show that between the components of the tensor $\hat{\sigma}$ and the CVC in the L and T geometries there exists a definite universal interrelationship. The general form of the formulas for reconstructing the tensor $\hat{\sigma}$ from CVC data in the L and T geometries for the anisotropic pinning model was formulated within the framework of the phenomenological approach in Ref. 18:

$$\sigma_{Ht}=-j[nE_\perp^L(j)+f_T(E_\parallel^L(j))]^{-1}, \tag{22}$$

$$\sigma_{Hl}=-j[nE_\perp^T(j)+f_L(E_\parallel^T(j))]^{-1}, \tag{23}$$

$$\sigma_l(x)=\frac{j_\parallel^L(x)}{x}\left[1+\frac{f_L^{-1}(x)}{f_T(x)}\right]^{-1}, \tag{24}$$

$$\sigma_t(x)=\frac{j_{\parallel}^{T}(x)}{x}\left[1+\frac{f_T^{-1}(x)}{f_L(x)}\right]^{-1}, \tag{25}$$

where

$$\sigma_{Ht}=-\alpha_{yy}\frac{c^2}{B\Phi_0}=-\frac{\epsilon_y}{\rho_f},\quad \sigma_{Hl}=-\alpha_{xx}\frac{c^2}{B\Phi_0}=-\frac{\epsilon_x}{\rho_f},$$

$$\sigma_l(x)=\frac{\sigma_{xx}(E_x)}{\rho_f},\quad \sigma_t(x)=\frac{\sigma_{yy}(E_y)}{\rho_f},$$

$E_{\parallel,\perp}^{L,T}(j)$ are the longitudinal and transverse components relative to the CVC current in the L and T geometries, and the functions f_L and f_T establish a relation between $E_{\parallel}$ and $E_{\perp}$ in the L and T geometries, respectively: $E_{\parallel}^{L,T}=f_{L,T}(nE_{\perp}^{L,T})$. Knowing the CVC's (15) and (16), it is possible within the framework of the microscopic model to find the analytical form of all functions entering into formulas (22)–(25):

$$E_{\perp}^{L}(j)=\frac{n\tilde{\nu}(\gamma\epsilon_x j)}{\gamma(1+\epsilon_x\epsilon_y)},\quad E_{\parallel}^{L}(j)\gamma j-\frac{\epsilon_y\tilde{\nu}(\gamma\epsilon_x j)}{1+\epsilon_x\epsilon_y},$$

$$E_{\perp}^{T}(j)=\frac{n\epsilon_y\tilde{\nu}(j)}{1+\epsilon_x\epsilon_y},\quad E_{\parallel}^{T}(j)=\frac{\tilde{\nu}(j)}{\gamma(1+\epsilon_x\epsilon_y)},$$

$$f_T(x)=\frac{1}{\gamma\epsilon_y}x,$$

$$f_L(x)=-\gamma\epsilon_y x+\frac{1}{\epsilon_x}\tilde{\nu}^{-1}[\gamma(1+\epsilon_x\epsilon_y)x].$$

Inspection shows that formulas (22)–(25) reconstructing the tensor $\hat{\sigma}$ within the given model are identities. This is because the diagonal component σ_{yy} of the conductivity tensor [see Eq. (18)] depends only on the corresponding component of the electric field E_y (the component σ_{xx} is a constant in the given model by virtue of the absence of pinning in the direction of the vector $\mathbf{m}_{\parallel}$). The fact that the diagonal components of the conductivity tensor depend only on the corresponding components of the field is an important postulate of the phenomenological model, on which the derivation of formulas (22)–(25) is based. The general form of these formulas is essential to the phenomenological approach. They make it possible from the results of CVC measurements in the L and T geometries to reconstruct the form of the conductivity tensor, in other words, to predict the CVC for an arbitrary angle α, and also to determine experimentally the Hall constants α_{xx} and α_{yy}, the electronic viscosities η_{xx} and η_{yy} forming respectively the tensors $\hat{\alpha}$ and $\hat{\eta}$, and the form of the function ν [see Eq. (17)]. Note that the function $f_T(x)$ in fact expresses linear scaling in the L geometry (for the components σ_{xy} and σ_{xx}) whereas in the T geometry (for the components σ_{yx} and σ_{yy}) such linear scaling is absent.

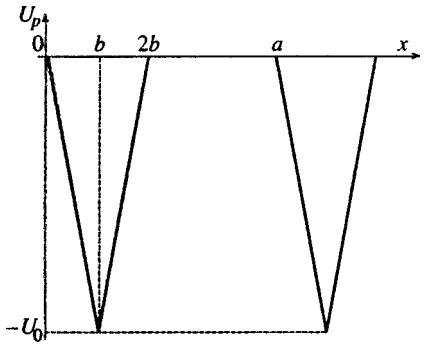

FIG. 3. Pinning potential of twins $U_p(x)$: a—potential period (distance between twins), $2b$—width of the twin potential well, U_0—depth of the twin potential well. Twin concentration is $\varepsilon=2b/a$. In all of the following relations $\varepsilon=0.001$.

3. ANISOTROPIC PINNING ON UNIDIRECTIONAL TWINS. ANALYSIS OF NONLINEAR REGIMES

3.1. Discussion of the model and analysis of the nonlinear behavior of the probability $\nu(j,t,\alpha,\varepsilon)$ of overcoming the twin barrier

The nonlinear properties of the resistivity tensor $\hat{\rho}$, as can be seen from formula (19), are completely determined by the behavior of the function $\nu(F,T)$, which has the physical sense of the probability of a vortex overcoming the potential barriers created by pinning centers. In turn, the function $\nu(F,T)$, according to formula (12), depends on the form of the pinning potential. In the present section we specify this potential for HTSC's of the type YBaCuO, in which the experimental realization of anisotropic pinning centers consists of twins. As is well known, a twin is a region of lowered value of the order parameter.[11] Therefore, it is energetically favorable for vortices to localize in the vicinity of a twin. In the present paper we analyze the resistive properties on the basis of a pinning potential of the form (Fig. 3)

$$U_p=\begin{cases}-F_p x, & 0\leqslant x\leqslant b,\\ F_p(x-2b), & b\leqslant x\leqslant 2b,\\ 0, & 2b\leqslant x\leqslant a,\end{cases} \tag{26}$$

where $F_p=U_0/b$ is the pinning force. Wells of width $2b$ in the potential (26) correspond to regions of twins, and zero pinning potential corresponds to the regions between the twins. As the parameter characterizing the twin concentration we use $\varepsilon=2b/a$ (more accurately, this is the volume fraction occupied by the twins).

Substituting the potential (26) into formula (12) for the probability function ν that a vortex will overcome the potential barrier of a twin gives the following expression:

$$\nu(f,t,\varepsilon)=\frac{2f(f^2-1)^2}{2f(f^2-1)(f^2-1+\varepsilon)-\varepsilon tG}, \tag{27}$$

where

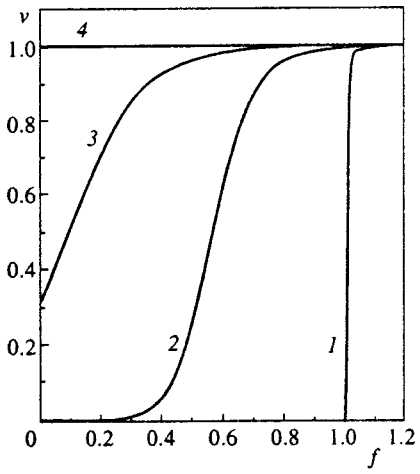

FIG. 4. The dependence $\nu(f)$ for the series of values of the parameter t_0: *1*—0, *2*—0.05, *3*—0.1, *4*—0.5.

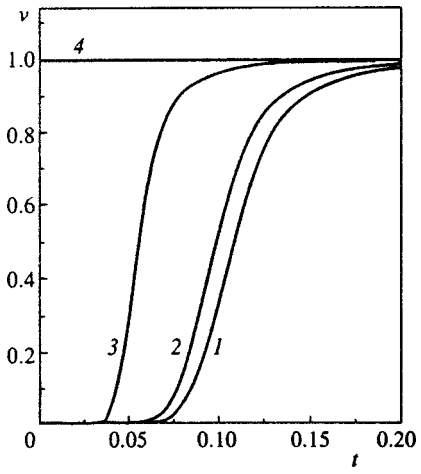

FIG. 5. The dependence $\nu(t)$ for the series of values of the parameter f_0: *1*—0.01, *2*—0.1, *3*—0.5, *4*—1.5.

$$G=\left[(3f^2+1)\cosh\left(\frac{f}{t\varepsilon}\right)+(f^2-1)\cosh\left(\frac{f(1-2\varepsilon)}{t\varepsilon}\right)\right.$$
$$-2f(f-1)\cosh\left(\frac{f(1-\varepsilon)}{t\varepsilon}-\frac{1}{t}\right)-2f(f+1)$$
$$\left.\times\cosh\left(\frac{f(1-\varepsilon)}{t\varepsilon}+\frac{1}{t}\right)\right]\Big/\sinh\left(\frac{f}{t\varepsilon}\right).$$

In formula (27) the effective external force F acting on the vortices in the direction perpendicular to the twins and causing them to move in this direction (the S effect) is characterized by the parameter f, which gives ratio of this force to the pinning force $F_p=U_0/b$ on the twins; the temperature is characterized by the parameter t, which gives the ratio of the energy of the thermal fluctuations of the vortices to the depth of the potential wells U_0 created by the twins. The influence of the external force F acting on the vortices is that it lowers the height of the potential barrier for vortices localized on the twins and, consequently, increases the probability of escape from them. Raising the temperature also increases the probability that a vortex will escape from a twin because of the increase in the energy of the thermal fluctuations of the twins. Thus, the pinning potential of the twins, leading as $F,T\to 0$ to localization of vortices, can be suppressed both by an external force and by an increase in the temperature.

The function $\nu_0(f)$ (see Fig. 4)

$$\nu_0(f)=\begin{cases}0, & 0\leqslant f\leqslant 1,\\ (f^2-1)/(f^2-1+\varepsilon), & f>1.\end{cases} \tag{28}$$

corresponds to the zero-temperature limit. In the zero-temperature limit, for $F<F_p$, the vortices are trapped in the potential wells of the twins and they cannot move, while for $F>F_p$ the potential barrier disappears and they begin to move. The value of the current at which $F=F_p$ corresponds to the critical current at zero temperature. Let us consider in turn the dependence of the probability function $\nu(f,t,\varepsilon)$ on each of the quantities f, t, and ε for the remaining quantities held fixed (denoted by the subscript ''0'').

The dependence $\nu(f)=\nu(f,t_0,\varepsilon_0)$ characterizing ν as a function of the external force acting on a vortex at constant temperature for constant twin concentration is monotonically increasing from the value $\nu(0,t_0,\varepsilon_0)$ to its limiting value of 1 as $f\to\infty$. By virtue of its even parity, the function $\nu(f)$ has a horizontal tangent at the point $f=0$. In the vicinity of infinity (for $f\gg 1$) the function $\nu(f)$ has a divergence $\nu(f)\approx 1-\varepsilon_0 f^{-2}$. The qualitative form of the dependence $\nu(f)$ is determined by the value of the dimensionless parameter t_0 characterizing the temperature. From the expansions of $\nu(0,t_0,\varepsilon_0)$ in a power series about t_0 in the vicinity of zero and infinity

$$\nu(0,t_0,\varepsilon_0)\approx\frac{\exp(-1/t_0)}{\varepsilon_0(1-\varepsilon_0)t_0}\left(1-\frac{\varepsilon_0 t_0}{1-\varepsilon_0}\right),\quad t_0\ll 1,\frac{1-\varepsilon_0}{\varepsilon_0}, \tag{29}$$

$$\nu(0,t_0,\varepsilon_0)\approx 1-\varepsilon_0\left(\frac{1}{3}-\frac{\varepsilon_0}{4}\right)t_0^{-2},\quad t_0\gg 1, \tag{30}$$

it is easy to understand the influence of the temperature on the qualitative form of $\nu(f)$. Specifically, at low temperatures ($T\ll U_0$) a nonlinear transition takes place from the TAFF regime of vortex motion perpendicular to the twins to the FF regime with growth of the external force, wherein the function $\nu(f)$ has a characteristic nonlinear shape (see Fig. 4). At high temperatures ($T\geqslant U_0$) the FF regime is realized over the entire range of variation of the external force.

The dependence $\nu(t)=\nu(f_0,t,\varepsilon_0)$ characterizing ν as a function of temperature for the external force and twin concentration fixed is also a monotonically increasing function, and its qualitative form is determined by the value of the parameter f_0, on which the quantity $\nu(f_0,0,\varepsilon)=\nu_0(f_0)$ depends [see formula (28)]. For $t\geqslant 1$ the expansion of $\nu(t)$ in a power series in t has the same form as the expansion of $\nu(0,t_0,\varepsilon_0)$ in t_0 in formula (30). Thus, the temperature dependence $\nu(t)$ (Fig. 5) depicts the nonlinear transition from the TAFF regime of vortex motion perpendicular to the twins to the FF regime as the temperature is raised (for $f_0\leqslant 1$) or a smoother transition to the FF regime from the dynamical state corresponding to $\nu_0(f_0)$ (for $f_0>1$). It also follows

from the above-said that the width of the transition from the TAFF to the FF regime, depending on t or f, decreases as ε is decreased [see also formula (28)].

The dependence $\nu(\varepsilon)=\nu(f_0,t_0,\varepsilon)$ characterizes the probability that a vortex will escape from the potential well of a twin as a function of the twin concentration for fixed values of the external force and temperature and is monotonically decreasing from the value $\nu(f_0,t_0,0)\equiv 1$ corresponding to the absence of twins. The probability density associated with finding a vortex on a twin grows with decrease of the temperature and the external force; therefore the smaller the parameters f_0 and t_0, the faster is the falloff of $\nu(\varepsilon)$. For $\varepsilon=1$ we regain the results of Ref. 24:

$$\nu(f,t,1)=(f^2-1)^2 \times\left[f^2(f^2-1)-2\,ft\,\frac{\cosh(f/t)-\cosh(1/t)}{\sinh(f/t)}\right]^{-1}. \tag{31}$$

The dynamics of a vortex system with the Hall effect taken into account depends substantially on the direction of the magnetic field. According to formula (9), an effective external force F perpendicular to the twins is created by a uniform transport current flowing through the sample; such a force contains two components—the Lorentz force $n(\Phi_0/c)j_y$ and the Magnus force $(\Phi_0/c)\gamma\epsilon_x j_x$, acting on the vortex along the x axis. Depending on the direction of the magnetic field assigned by the factor $n=\pm 1$, these two components can be identically or oppositely directed, and the resulting force will be different in these two cases. In what follows we express the current density in units of $j_c=cU_0/\Phi_0 b$ (in our study of the problem of a potential of general form we used the units $cU_0/\Phi_0 a$) so that for the dimensionless parameter f, characterizing the external force we obtain

$$f=nj\phi(\alpha),\quad \phi(\alpha)=\cos\alpha+n\gamma\epsilon_x\sin\alpha.$$

Thus, the value of the external force $F=F_p$ at which the height of the potential barrier vanishes corresponds to the dimensionless current $j=1/\phi(\alpha)$, equal to the critical current at $T=0$, and in the case $0<T\ll U_0$, to the crossover current $j_{\rm cr}$.

In the function $\nu(j\phi(\alpha),t)=\nu(f,t,\varepsilon_0)$ we separate out the even component $\nu^+=[\nu(n=1)+\nu(n=-1)]/2$ and the odd component $\nu^-=[\nu(n=1)-\nu(n=-1)]/2$ in the magnetic field. These components determine the observed resistive characteristics—the even and odd components (in the magnetic field) of the longitudinal and transverse resistivity $\rho^{\pm}_{\parallel,\perp}$. From the definition of $\nu^{\pm}$ it follows that their possible values for any values of j, t, α, ϵ_x, and ε, like the values of the function ν, always lie between zero and one. Note that only the Hall constant ϵ_x enters into the even and the odd components of $\nu(j\phi(\alpha),t)$ since the x component of the Magnus force depends on it, and this component contributes to the effective force F along the x axis, perpendicular to the twins. It is easy to see that the component ν^+ is even in ϵ_x, i.e., $\nu^+(-\epsilon_x)=\nu^+(\epsilon_x)$, and the component ν^- is odd, i.e., $\nu^-(-\epsilon_x)=-\nu^-(\epsilon_x)$. Therefore, for $\epsilon_x=0$ the even component $\nu^+\equiv\nu(j_y,t)$, and the odd component $\nu^-\equiv 0$. The qualitative behavior and the limits of the component $\nu^+(j)$ as $j\to 0,\infty$ coincide with the corresponding limits of $\nu(f)$. The component $\nu^-(j)$ tends to zero in the linear regimes (as $j\to 0,\infty$) and is nonzero in the region of nonlinearity of ν, forming a characteristic peak (Fig. 6). The limits of the components $\nu^+(t)$ and $\nu^-(t)$ as $t\to\infty$ are the same as for $j\to\infty$, and for $t\to 0$ they are equal to the corresponding components $\nu_0^{\pm}$ of the function $\nu_0[j\phi(\alpha)]$.

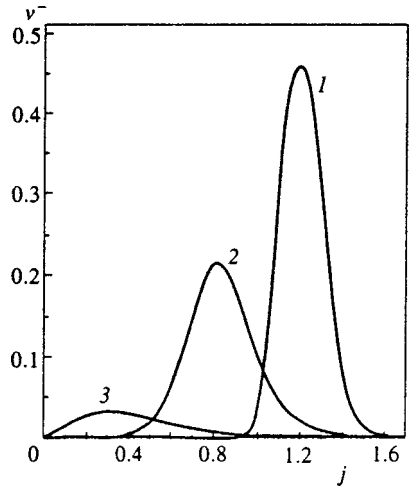

FIG. 6. The dependence $\nu^-(j)$ for the series of values of the parameter t_0: *1*—0.02, *2*—0.05, *3*—0.1, for $\alpha=45°$, $\gamma=1$, and $\epsilon_x=0.1$.

In the case of a small Hall effect ($\epsilon_x,\epsilon_y\ll 1$) the function $\nu(f)=\nu[j_y(1+n\gamma\epsilon_x\tan\alpha)]$ can be expanded in the parameter $n\gamma\epsilon_x\tan\alpha$, which is small in the range of angles where $\tan\alpha\ll 1/\gamma\epsilon_x$. In the linear approximation in the parameter $n\gamma\epsilon_x\tan\alpha$ the expressions for the even and odd components (in the magnetic field) of the function ν are respectively $\nu^+=\nu(j\cos\alpha)$ and $\nu^-=\nu'(f)|_{j\cos\alpha}\,j\gamma\epsilon_x\sin\alpha\sim\epsilon_x\ll 1$.

The behavior of the probability of overcoming the potential barrier of a twin, ν, and of its components ν^+ and ν^-, as functions of the external force (or current) and the temperature is connected with their influence on the height of the barrier. At zero temperature, the relative decrease in the height of the potential barrier caused by the external force (current) can be estimated as $\delta=\Delta U/U_0=1-|f|=1-j\phi(\alpha)$, so that vortex motion is impossible for $\delta>0$. The influence of temperature, with growth of which the probability of overcoming the barrier is increased, is such that for $t\leqslant 0.1\delta$ the TAFF regime of vortex motion transverse to the twins is realized, and for $t\gg 0.1\delta$ the FF regime is realized. The nature of the transition from the TAFF to the FF regime in the functional dependences $\nu(f)$ and $\nu(t)$ is substantially different. The dependence $\nu(f)$ shifts to the left with growth of t, and its steepness decreases (see Fig. 4). Thus, the higher the temperature, the smoother the transition from the TAFF to the FF regime, and the lower the values of the external force at which it occurs. The dependence $\nu(t)$ also shifts to the left with growth of f, but its steepness grows (see Fig. 5). Consequently, the lower the barrier δ, the steeper the transition from the TAFF to the FF regime and the lower the temperatures at which it occurs. The behavior of $\nu^+(j)$ and $\nu^+(t)$ is similar to the behavior of $\nu(f)$ and $\nu(t)$, respec-

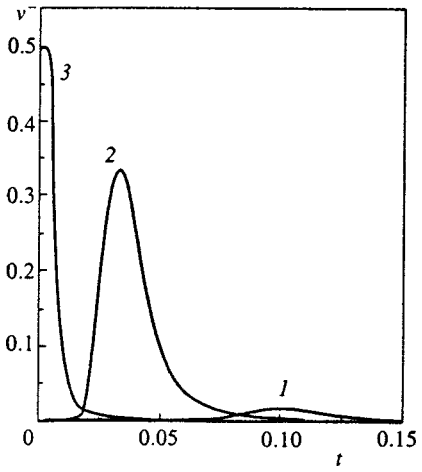

FIG. 7. The dependence $\nu^-(t)$ for the series of values of the parameter j_0: *1*—0.1, *2*—1.0, *3*—1.5, for $\alpha=45°$, $\gamma=1$, and $\epsilon_x=0.1$.

tively. The location of the peaks determined by the dependences $\nu^-(j)$ and $\nu^-(t)$ corresponds to the region of the transition in the current and the temperature from the TAFF to the FF regime, and their amplitude is inversely proportional to the width of the transition. Depending on $\nu^-(j)$, with increasing t the peak shifts to the left, and its amplitude decreases (see Fig. 6). Depending on $\nu^-(t)$, with increasing j the peak shifts to the left, but its amplitude increases (Fig. 7). These peculiarities of the behavior of the dependences $\nu^\pm(j)$ and $\nu^\pm(t)$ underlie the behavior of the experimentally observed quantities—the even and odd components (in the magnetic field) of the longitudinal and transverse resistivity $\rho^\pm_{\parallel,\perp}$.

We will obtain expressions from formulas (19)–(21)for the experimentally observed longitudinal $\rho_\parallel=E_\parallel/j$ and transverse $\rho_\perp=E_\perp/j$ resistivities (relative to the current) with the Hall effect taken into account. We separate out their even $\rho^+[\rho(n=1)+\rho(n=-1)]/2$ and odd $\rho^-=[\rho(n=1)-\rho(n=-1)]/2$ components relative to the magnetic field:

$$\rho^\pm_\parallel \frac{1}{1+\epsilon_x\epsilon_y}\left[(\epsilon_x-\epsilon_y)\nu^\mp \sin\alpha\cos\alpha+\left(\frac{1}{\gamma}\cos^2\alpha - \gamma\epsilon_x\epsilon_y\sin^2\alpha\right)\nu^\pm\right]+\frac{\gamma}{2}(1\pm1)\sin^2\alpha, \tag{32}$$

$$\rho^\pm_\perp=\frac{1}{1+\epsilon_x\epsilon_y}\left[(\epsilon_x\sin^2\alpha+\epsilon_y\cos^2\alpha)\nu^\mp+\left(\frac{1}{\gamma}+\gamma\epsilon_x\epsilon_y\right)\nu^\pm\sin\alpha\cos\alpha\right]-\frac{\gamma}{2}(1\pm1)\sin\alpha\cos\alpha, \tag{33}$$

where $\nu^\pm$ are the above-defined even and odd components relative to the magnetic field of the function $\nu(j\phi(\alpha),t)$. In formulas (32) and (33) the nonlinear and linear (nonzero only for $\rho^+_{\parallel,\perp}$) terms separate out in a natural way. The physical reason for the appearance of linear terms is that in the model under consideration for $\alpha\neq0$ there is always an FF regime of vortex motion along the twins.

3.2. The nonlinear G effect

Let us proceed now to a treatment of the vortex dynamics and the resistive properties associated them them, based on the anisotropic pinning potential introduced above [see formula (26)]. For simplicity we will neglect the usually small Hall effect, i.e., we take $\epsilon_x=\epsilon_y=0$. As a consequence, the nondiagonal components of the resistivity tensor vanish ($\rho_{xy}=\rho_{yx}\equiv0$). Neglecting the Hall effect, the formulas for the experimentally observed longitudinal $\rho_\parallel$ and transverse $\rho_\perp$ resistivities relative to the current can be represented formally in a manner analogous to the linear case[9]:

$$\rho_\parallel=\rho_{xx}\cos^2\alpha+\rho_{yy}\sin^2\alpha,$$
$$\rho_\perp=\sin\alpha\cos\alpha(\rho_{yy}-\rho_{xx}), \tag{34}$$

with one substantial difference, namely, that in the nonlinear case one of the components of the resistivity tensor depends on the current j and the angle α, whereas the second is constant:

$$\rho_{xx}=\gamma, \quad \rho_{yy}=(1/\gamma)\nu(j_y,t) \tag{35}$$

[here, as earlier, see formula (19), and the resistance is measured in units of $\rho_f=B\Phi_0/c^2\eta_0$]. Therefore, under certain conditions in the current and temperature dependences of the observed resistivities (34) a pronounced nonlinearity appears while in the dynamics of a vortex system as a consequence of anisotropy of the pinning viscosity a nonlinear G effect is observed in both the temperature and the current. As a consequence of the even parity of $\nu(f,t,\varepsilon)$ in f the quantities $\rho_\parallel$ and $\rho_\perp$ in formula (34) are even in the magnetic field, as they should be neglecting the Hall effect.

As is well known,[18] the specifics of anisotropic pinning consist in the noncoincidence of the directions of the external motive force $\mathbf{F}_L$ acting on the vortex, and its velocity $\mathbf{v}$ (for isotropic pinning $\mathbf{F}_L\|\mathbf{v}$ if we neglect the Hall effect). The anisotropy of the pinning viscosity along and transverse to the twin boundaries leads to the result that for those values of j,t,α for which the component of the vortex velocity perpendicular to the twins, v_x, is suppressed, a tendency is shown toward a substantial prevalence of guided vortex motion along the twins (the G effect) over motion transverse to the twins (the S effect). Note that the probability function ν describing the nonlinear properties of the observed resistive characteristics in the model under consideration is inversely proportional to the effective nonlinear viscosity introduced in Ref. 18, where the phenomenological approach was developed.

In the experiment, the function

$$\cot\beta=-\frac{\rho_\perp}{\rho_\parallel}=\frac{\gamma^2-\nu(j_y,t)}{\gamma^2\tan\alpha+\nu(j_y,t)\cot\alpha}, \tag{36}$$

is used to describe the G effect, where β is the angle between the vortex velocity vector $\mathbf{v}$ and the current density vector $\mathbf{j}$ (see Fig. 1). The G effect is expressed that much more strongly, the larger is the difference in directions of $\mathbf{F}_L$ and $\mathbf{v}$, i.e., the smaller is the angle β. Here it is possible that $\cot\beta\gg1$, i.e., $\rho_\perp\gg\rho_\parallel$. In the T and L geometries $\beta(\alpha=0)=\beta(\alpha=\pi/2)\equiv\pi/2$ since the Lorentz force is directed exactly perpendicular or parallel to the twins.

Let us consider the sensitivity of the magnitude of the angle β to small deviations of the angle α from the values 0 and $\pi/2$ corresponding to the T and L geometries, in the linear approximation, for which we will calculate the derivative $d\beta/d\alpha$ at $\alpha=0$ and $\alpha=\pi/2$:

$$\left.\frac{d\beta}{d\alpha}\right|_{\alpha=0}=1-\frac{\gamma^2}{\nu(j,t)},\quad \left.\frac{d\beta}{d\alpha}\right|_{\alpha=\pi/2}=1-\frac{\nu(0,t)}{\gamma^2}. \tag{37}$$

As can be seen from Eqs. (37), in the T geometry ($\alpha=0$) the derivative $d\beta/d\alpha$ depends on both temperature and the current. In the TAFF regime, where $\nu(j,t)\ll 1$, this derivative has very large values, consistent with the G effect, while in the FF regime (for $j\to\infty$ and $t\to\infty$) it has the limit $1-\gamma^2$, i.e., the dependence $\beta(\alpha)$ in this case is due to the anisotropy of the electronic viscosity. In the L geometry ($\alpha=\pi/2$) the derivative $d\beta/d\alpha$ depends only on the temperature since $j_y\equiv 0$. For $t=0$ we have $(d\beta/d\alpha)_{\alpha=\pi/2}=1$, and in the limit $t\to\infty$ we obtain $(d\beta/d\alpha)_{\alpha=\pi/2}=1-1/\gamma^2$.

Let us consider the current and temperature dependence of $\cot\beta(j,t)$ for fixed values of the angle $\alpha\neq 0,\ \pi/2$. The limits of $\cot\beta(j,t)$ as $j,t\to 0$ and as $j,t\to\infty$ are obtained by substituting the corresponding limits of the function $\nu(j\cos\alpha,t)$ into formula (36). For the limit of $\cot\beta(j,t)$ as $j\to 0$ we obtain $\lim_{j\to 0}\cot\beta\approx\cot\alpha$ correct to within $\max[\cot^2\alpha\nu(0,t)/\gamma^2,\nu(0,t)/\gamma^2]$. In the temperature region corresponding to the TAFF regime, we have $\beta\approx\alpha$ and, consequently, at low currents the G effect arises. At large currents ($j_y\gg 1$), where for vortex motion transverse to the twins the FF regime is set up, the pinning viscosity becomes isotropic so that for $\gamma=1$ we have $\mathbf{v}\|\mathbf{F}_L$ for arbitrary values of the angle α. In the temperature dependence of $\cot\beta(j,t)$ the G effect is most strongly pronounced for currents $j_y<1$, where $\beta(t=0)=\alpha$.

3.3. The Hall effect and the G effect in nonlinear regimes

In this subsection we consider peculiarities of the resistive characteristics in the investigated model due to the Hall effect. Experimentally, two types of measurements of the observed resistive characteristics are possible in a prescribed geometry defined by a fixed value of the angle α: CVC measurements and resistive measurements, which investigate the dependence of the observed resistivities on the current density at a fixed temperature $\rho^{\pm}_{\|,\perp}(j)$ and on the temperature for fixed current density $\rho^{\pm}_{\|,\perp}(t)$. The form of these dependences is governed by a geometrical factor—the angle α between the directions of the current density vector $\mathbf{j}$ and the twin plane. There are two different forms of the dependence of $\rho^{\pm}_{\|,\perp}$ on the angle α [see formulas (32) and (33)]. The first of these is the ''tensor'' dependence, also present in the linear regimes (TAFF and FF), which is external to the function ν. The second is through the dependence of the function ν on its argument $f=nj\phi(\alpha)$, which in the region of the transition from the TAFF to the FF regime is substantially nonlinear [see Eq. (27)].

First recall that in the absence of the Hall effect ($\epsilon_x=\epsilon_y\equiv 0$ there exist only even resistivities (in the magnetic field) $\rho^{+}_{\|,\perp}$—the odd resistivities $\rho^{-}_{\|,\perp}$ are zero [see formulas (32) and (33)]. The presence of nonzero values of ϵ_x,ϵ_y leads not only to the appearance of a Hall contribution to the observed resistivities on account of the even component ν^+ of the function ν, but also to the appearance of the odd component ν^-, which has a maximum in the region of the nonlinear transition from the TAFF to the FF regime and is essentially equal to zero outside this transitional region (see Figs. 6 and 7). As a consequence, ''crossover'' effects arise: contributions from ν^- to effects due to ν^+, and vice versa: contributions from ν^+ to effects due to ν^-. Thus, in the even resistivities $\rho^{+}_{\|,\perp}$, in addition to the main contribution created by the G effect and described by ν^+, there is present a Hall contribution arising due to ν^-. The expressions for the odd resistivities $\rho^{-}_{\|,\perp}$ contain, in addition to the Hall terms arising due to ν^+, terms due to ν^- [see formulas (32) and (33)].

Let us analyze the resistive dependences $\rho^{\pm}_{\|,\perp}(j)$ and $\rho^{\pm}_{\|,\perp}(t)$ with allowance for the small Hall effect. We first consider the simplest case of isotropy of the electronic viscosity $\gamma=1$ and a small isotropic Hall effect ($\epsilon_x=\epsilon_y=\epsilon\ll 1$). In this case, the expressions for $\rho^{\pm}_{\|,\perp}$, out to terms of order $\epsilon_x\epsilon_y\ll 1$, have the form

$$\rho^{+}_{\|}\approx\nu^{+}\cos^2\alpha+\sin^2\alpha,\quad \rho^{+}_{\perp}\approx(\nu^{+}-1)\sin\alpha\cos\alpha, \tag{38}$$

$$\rho^{-}_{\|}\approx\nu^{-}\cos^2\alpha,\quad \rho^{-}_{\perp}\approx\epsilon\nu^{+}+\nu^{-}\sin\alpha\cos\alpha. \tag{39}$$

in the case of a small Hall effect ($\epsilon\ll 1$) the expressions for the even and odd components of $\nu(f)$ in the linear approximation in the parameter $\epsilon\tan\alpha\ll 1$ (see Subsection 3.1.) are equal respectively to $\nu^+=\nu(j\cos\alpha)$ and $\nu^-=\nu'(f)|_{j\cos\alpha}\,j\gamma\epsilon\sin\alpha\sim\epsilon\ll 1$, i.e., the functions $\nu^{\pm}$ are related to ν in a simple way.

Expressions (38) and (39) with the condition $\epsilon\tan\alpha\ll 1$ taken into account lead to a new nonlinear scaling relation for the Hall conductivity having the following form:

$$\epsilon=\frac{(\rho^{-}_{\perp}-\rho^{-}_{\|}\tan\alpha)\cos^2\alpha}{(\rho^{+}_{\|}-\sin^2\alpha)}. \tag{40}$$

Thus, the isotropic Hall conductivity $\epsilon\ll 1$ is uniquely related to three experimentally observable nonlinear resistivities: $\rho^{+}_{\|}$, $\rho^{-}_{\|}$, and $\rho^{-}_{\perp}$ while the scaling relation depends on the angle α. This relation differs substantially from the power-law scaling law obtaining in the isotropic case.[22] In the particular case $\alpha=0$ we regain the results of Ref. 16, specifically $\epsilon=\rho^{-}_{\perp}/\rho^{+}_{\|}\equiv\rho_{\perp}/\rho_{\|}$, i.e., a linear relationship between $\rho_{\perp}$ and $\rho_{\|}$.

The nature of the behavior of the current and temperature dependence of $\rho^{\pm}_{\|,\perp}$ is completely determined by the behavior of the dependences $\nu^{\pm}(j)$ and $\nu^{\pm}(t)$. The linear limit in formulas (38) and (39) is realized in that region of currents and temperatures where $\nu^+=\text{const}$ and $\nu^-=0$ while the region of nonlinearity of the current and temperature dependence of $\rho^{\pm}_{\|,\perp}$ corresponds to those current and temperature intervals where the dependences $\nu^{\pm}(j)$ and $\nu^{\pm}(t)$ are nonlinear. Let us turn our attention now to the fact that the nonlinearity in the temperature dependences $\rho^{\pm}_{\|,\perp}(t)$ can be observed even at large currents $j>1$ in the case when $j\phi(\alpha)<1$, where this this latter relation depends on the magnitude of the angle α [for $\phi(\alpha)<1/j$ we have $f<1$ and $\nu(t\to 0)=0$] and determines the limiting values of $\nu^{\pm}(t)$ as

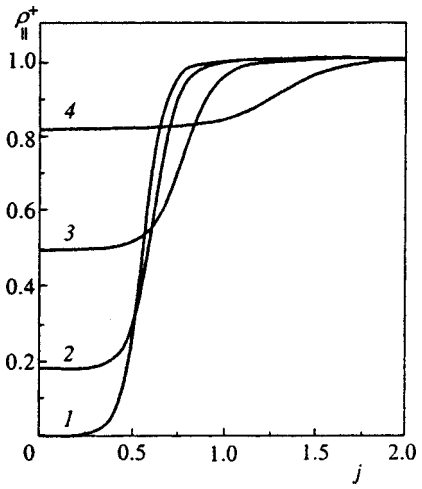

FIG. 8. The dependence $\rho_{\parallel}^{+}(j)$ for the series of values of the angle α: *1*—0°, *2*—25°, *3*—45°, *4*—65°, for $t_0=0.05$, $\gamma=1$, and $\epsilon_x=\epsilon_y=0.02$.

$t\to 0$. Thus, the linearity or nonlinearity of the dependences $\rho_{\parallel,\perp}^{\pm}(t)$ at currents larger than unity depends on the magnitude of the angle α.

In the even longitudinal resistivity $\rho_{\parallel}^{+}$ and the even transverse resistivity $\rho_{\perp}^{+}$ for a small Hall effect, terms proportional to $\epsilon\ll 1$ are absent, and only linear terms and contributions describing the nonlinear G effect are present. The qualitative form of the current dependence $\rho_{\parallel,\perp}^{+}(j)$ and the temperature dependence $\rho_{\parallel,\perp}^{+}(t)$ is determined only by the even component ν^{+} of the function ν.

The limiting values of the qualitatively similar dependences $\rho_{\parallel}^{+}(j)$ and $\rho_{\parallel}^{+}(t)$ corresponding to the TAFF regime of vortex motion transverse to the twins are determined by guided vortex motion along the twins and grow with increasing magnitude of the angle α since in this case the component of the Lorentz force along the twins increases. In the FF regime, as the pinning viscosity becomes isotropic the contribution to the dependences $\rho_{\parallel}^{+}(j)$ and $\rho_{\parallel}^{+}(t)$ due to vortex motion transverse to the twins becomes substantial, and the limiting values of these dependences are equal to unity (Fig. 8).

The main contribution to the even transverse resistivity $\rho_{\perp}^{+}$ is proportional to the factor $\sin\alpha\cos\alpha$; therefore, the angle most favorable for its observation is $\alpha=\pi/4$. The current dependence $\rho_{\perp}^{+}(j)$ and the temperature dependence $\rho_{\perp}^{+}(t)$ have their maximum absolute values in the TAFF regime of vortex motion transverse to the twins [see Eqs. (38)] (the same value is approached if the angle is replaced by its complement in the limit $j\to 0$ and $t\to 0$) and go to zero with the onset of the FF regime as a consequence of isotropization of the pinning viscosity (Fig. 9). In the case of isotropic electronic viscosity (the case under consideration) the resistivity $\rho_{\perp}^{+}$ can serve as a measure of the anisotropy of the pinning viscosity since it is determined by the difference of the pinning viscosities transverse to and along the twins [see also Eqs. (38)].

As was noted above, the odd longitudinal $\rho_{\parallel}^{-}$ and even transverse $\rho_{\perp}^{-}$ resistivities arise thanks to the Hall effect, and therefore their characteristic scale is proportional to $\epsilon\ll 1$ [see Eqs. (39)].

Only the odd component ν^{-} of the function ν contributes to the current dependence $\rho_{\parallel}^{-}(j)$ and temperature dependence $\rho_{\parallel}^{-}(t)$ of the odd longitudinal resistivity in the case of the isotropic Hall effect; therefore, their qualitative form is determined completely by the behavior of ν^{-} as a function of the current and temperature. A characteristic peak appears in the dependences $\rho_{\parallel}^{-}(j)$ and $\rho_{\parallel}^{-}(t)$ in the region of nonlinearity of ν^{-} as a function of the current and temperature while in the TAFF and FF regimes of vortex motion transverse to the twins they vanish (Fig. 10).

In the current dependence $\rho_{\perp}^{-}(j)$ and temperature dependence $\rho_{\perp}^{-}(t)$ of the odd transverse resistivity there are contributions both from the even ν^{+}, and from the odd ν^{-} component of the function ν, whose relative magnitudes are determined by the angle α, the Hall constant ϵ, the parameter t_0 [for the dependence $\rho_{\perp}^{-}(j)$] and the current density j_0 [for the dependence $\rho_{\perp}^{-}(t)$]. Thus, the magnitude of the contributions of $\epsilon\nu^{+}$ and ν^{-} can be of the same order of magnitude in the current dependence $\rho_{\perp}^{-}(j)$ at low temperatures ($t\ll 1$) and in the temperature dependence $\rho_{\perp}^{-}(t)$ at large currents [$j>1/\phi(\alpha)$] (Fig. 11) whereas in the temperature dependence $\rho_{\perp}^{-}(t)$ for weak currents ($j\ll 1$) the contribution of $\epsilon\nu^{+}$ dominates substantially. The limiting values of the current dependence $\rho_{\perp}^{-}(j)$ as $j\to 0,\infty$ are determined by the corresponding limits of ν^{+} while the contribution of the odd

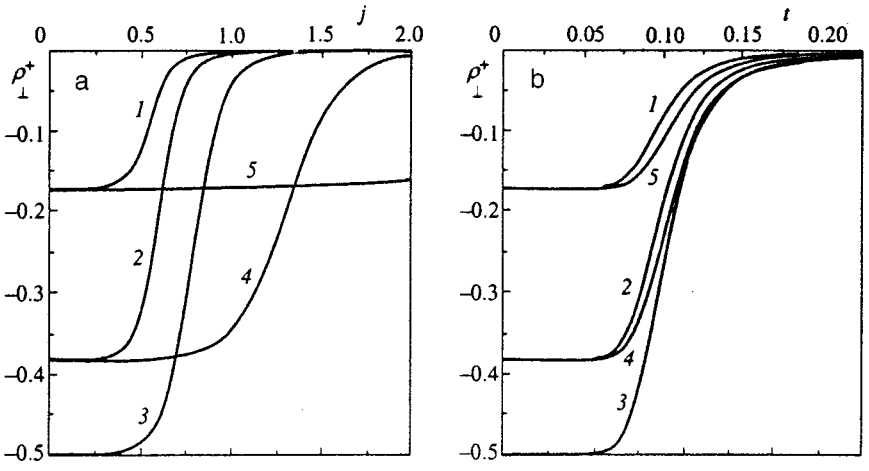

FIG. 9. Dependence of $\rho_{\perp}^{+}$ on j (a) for $t_0=0.05$ and on t (b) for $j_0=0.1$ for the series of values of the angle α: *1*—10°, *2*—25°, *3*—45°, *4*—65°, *5*—80°, for $\gamma=1$ and $\epsilon_x=\epsilon_y=0.02$.

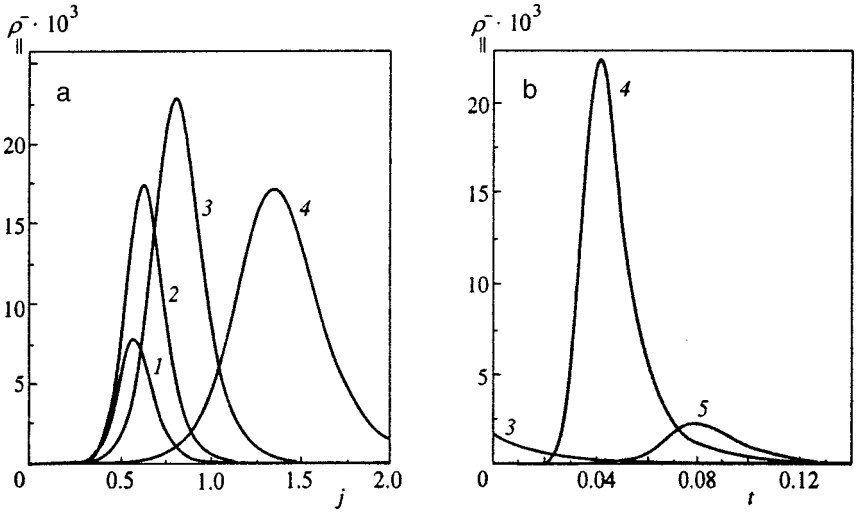

FIG. 10. Dependence of $\rho_{\parallel}^{-}$ on j (a) for $t_0=0.05$ and on t (b) for $j_0=1.5$ for the series of values of the angle α: *1*—10°, *2*—25°, *3*—45°, *4*—65°, *5*—80° for $\gamma=1$, and $\epsilon_x=\epsilon_y=0.02$.

component ν^- becomes considerable in its nonlinearity region where it forms its characteristic peak. In the temperature dependence $\rho_{\perp}^{-}(t)$ in the limit as $t\to 0$ a substantial contribution from the odd component ν^- is also possible if different regimes of vortex motion transverse to the twins are realized for opposite directions of the magnetic field.

Let us turn now to a discussion of the more general formulas with $\gamma\neq 1$ and $\epsilon_x\neq\epsilon_y$, including the case of different signs. The presence of anisotropy of the electronic viscosities leads to a change in the limiting values of the observable resistivities $\rho_{\parallel,\perp}^{\pm}$ in those cases when they depend on the anisotropy parameter γ (Fig. 12). Whereas for $\gamma=1$ complete isotropization of the system takes place in the FF regime of vortex motion transverse to the twins, for $\gamma\neq 1$ anisotropy remains in this regime, due to anisotropy of the electronic viscosity. The anisotropy of the Hall constants has a substantial effect on the odd resistivities $\rho_{\parallel,\perp}^{-}$ created by the Hall effect, and no effect on the even resistivities $\rho_{\parallel,\perp}^{+}$, which are essentially independent of the small Hall effect. The physical reason for the influence of anisotropy of the Hall constants on the behavior of the dependences $\rho_{\parallel,\perp}^{-}(j)$ and $\rho_{\parallel,\perp}^{-}(t)$ is that the relative values of the Hall constants determine the angle between the vortex velocity and the Magnus force acting on them and, consequently, they determine the direction of the vortex velocity itself (in the anisotropic case, $\epsilon_x\neq\epsilon_y$, this angle deviates from the value $\pi/2$, which corresponds to the isotropic case, $\epsilon_x=\epsilon_y$). Anisotropy of the Hall constants leads to the appearance of a new contribution to the odd longitudinal resistivity $\rho_{\parallel}^{-}$, which is absent in the isotropic Hall effect. In contrast to the isotropic case, for anisotropy of the Hall constants it is possible for the current and temperature dependences $\rho_{\parallel,\perp}^{-}$ to change sign (Fig. 13).

To conclude this discussion, note that the case $|\epsilon_x|,|\epsilon_y|\sim 1$ is of practical interest for pure superconductors[25] by virtue of the fact that the Lorentz and Magnus forces acting on a vortex have the same order of magnitude, but this point requires separate study due to the large variety of possible effects.

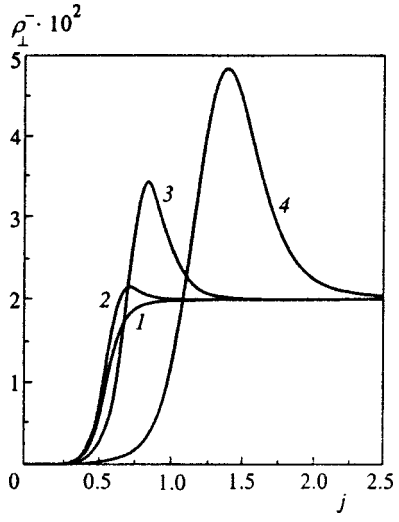

FIG. 11. The dependence $\rho_{\perp}^{-}(j)$ for the series of values of the angle α: *1*—0°, *2*—25°, *3*—45°, *4*—65°, for $t_0=0.05$, $\gamma=1$, and $\epsilon_x=\epsilon_y=0.02$.

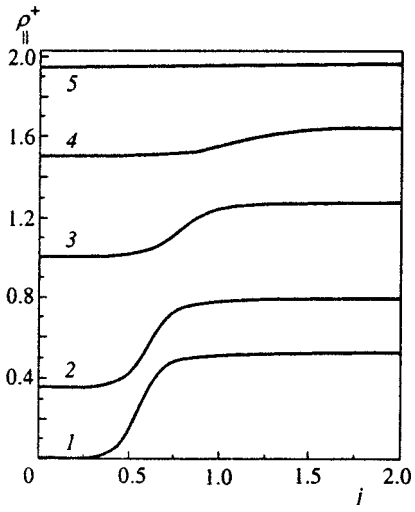

FIG. 12. The dependence $\rho_{\parallel}^{+}(j)$ for the series of values of the angle α: *1*—0°, *2*—25°, *3*—45°, *4*—60°, *5*—80° for $t_0=0.05$, $\gamma=2$, and $\epsilon_x=\epsilon_y=0.02$.

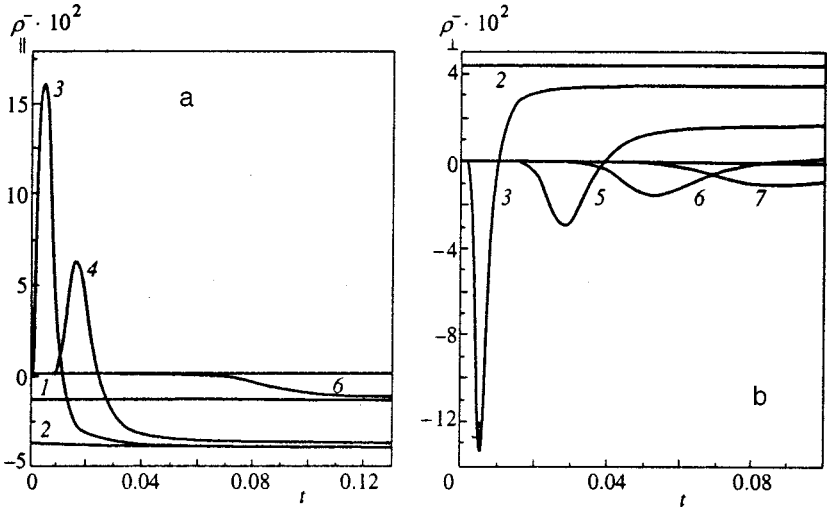

FIG. 13. Dependence of $\rho_\parallel^-$ (a, $\epsilon_x=0.02$) and $\rho_\perp^-$ (b, $\epsilon_x=-0.01$) on t in the case of anisotropy of the Hall effect for the series of values of the angle α: *1*—10°, *2*—45°, *3*—50°, *4*—55°, *5*—60°, *6*—80°, *7*—85°, for $j_0=1.5$, $\gamma=1$, and $\epsilon_y=0.1$.

3.4. Scaling relations and their stability

In order to give concrete form to the above-obtained scaling relations [see Eqs. (22)–(25)], let us consider the observed resistivities in the T and L geometries, where the current is directed exactly parallel ($\alpha=\pi/2$) or perpendicular ($\alpha=0$) to the twins. It follows from formulas (32) and (33) that in these limiting cases $\rho_\parallel^-=\rho_\perp^+\equiv 0$, and we obtain for $\rho_\parallel^+$ and $\rho_\perp^-$

$$\rho_{\parallel,T}^+=\frac{\nu(j,t)}{\gamma(1+\epsilon_x\epsilon_y)},$$

$$\rho_{\perp,T}^-=\frac{\epsilon_y\nu(j,t)}{(1+\epsilon_x\epsilon_y)}\ (\alpha=0,\ \ T-\text{geometry}), \tag{41}$$

$$\rho_{\parallel,L}^+=\gamma\left(1-\frac{\epsilon_x\epsilon_y\nu(\epsilon_x\gamma j,t)}{(1+\epsilon_x\epsilon_y)}\right),$$

$$\rho_{\perp,L}^-=\frac{\epsilon_x\nu(\epsilon_x\gamma j,t)}{(1+\epsilon_x\epsilon_y)}\ (\alpha=\pi/2,\ \ L-\text{geometry}). \tag{42}$$

Let us discuss the case of small Hall effects ($\epsilon_x,\epsilon_y\ll 1$). The longitudinal $\rho_\parallel^-$ and transverse $\rho_\perp^+$ resistivity for a superconductor with uniaxial pinning anisotropy vanish in the T and L geometries. The longitudinal even $\rho_\parallel^+$ and transverse odd $\rho_\perp^-$ resistivity in the T geometry are due to vortex motion transverse to the twins, described by the function $\nu(j,t)$. In the limit $j,t\to\infty$ to within terms proportional to $\epsilon_x\epsilon_y\ll 1$ we have $\rho_\parallel^+=1/\gamma$ and $\rho_\perp^-=\epsilon_y$. The main contribution to the resistivity $\rho_\parallel^+$ in the L geometry, which is equal to γ with the same accuracy, is due to guided vortex motion along the twins, for which pinning is absent. The magnitude of the resistivity $\rho_\perp^-$ in the L geometry is described by the function $\nu(\epsilon_x\gamma j,t)$. Since its first argument, the x component of the Magnus force, $\epsilon_x\gamma j$, is vanishingly small for a small Hall effect for realistically achievable currents $j\ll 1/\epsilon_x\gamma$ and the velocity component v_x is suppressed, the resistivity $\rho_\perp^-$ dependence mainly only on the temperature. For $t\ll 1$ the resistivity $\rho_\perp^-$ is so small it cannot be measured ($\rho_\perp^-\equiv 0$ in the limit $t=0$ since $\epsilon_x\gamma j<1$), and for $t\sim 1$ it approaches the value of the Hall constant, ϵ_x (to within terms proportional to $\epsilon_x\epsilon_y\ll 1$).

Formulas (41) and (42) express simple scaling relations between the observable resistivities $\rho_\parallel^+$ and $\rho_\perp^-$ in the T and L geometries on the one hand and the constants $\epsilon_x,\epsilon_y,\gamma$ on the other. The values of the latter and the form of the function $\nu(x)$ can be reconstructed, as can be seen from formulas (41) and (42), from measurements of $\rho_\parallel^+$ and $\rho_\perp^-$ in the T and L geometries. Therefore, it makes sense to consider the question of the stability of the measurements in these geometries since the preparation of the samples can lead to small deviations $\Delta\alpha$ from the values $\alpha=0,\pi/2$. Here it should also be borne in mind that besides the resistivities $\rho_\parallel^+$, and $\rho_\perp^-$ assigned by formulas (41) and (42), in the presence of an angle deviation, $\Delta\alpha$, the resistivities $\rho_\parallel^-$ and $\rho_\perp^+$, not present in the T and L geometries, also appear. The expansions of $\rho_{\parallel,\perp}^{\pm}$ in α about $\alpha=0$ (in the T geometry) and in $\Delta\alpha=\pi/2-\alpha$ about $\alpha=\pi/2$ (in the L geometry) out to the first nonvanishing terms neglecting terms of the order of $\epsilon_x\epsilon_y\ll 1$ have the form

$$\rho_{\parallel,T}^-\approx[(\epsilon_x-\epsilon_y)\nu(j,t)+(j\epsilon_x/\gamma)\nu'(j,t)]\alpha,$$

$$\rho_{\perp,T}^+\approx[(1/\gamma)\nu(j,t)-\gamma]\alpha,$$

$$\rho_{\parallel,L}^-\approx(\epsilon_x-\epsilon_y)\nu(\epsilon_x\gamma j,t)\Delta\alpha,$$

$$\rho_{\perp,L}^+\approx[(1/\gamma)\nu(\epsilon_x\gamma j,t)+j\epsilon_x\nu'(\epsilon_x\gamma j,t)-\gamma]\Delta\alpha,$$

$$\rho_{\parallel,T}^+\approx\frac{1}{\gamma}\nu(j,t)+\left[\gamma-\frac{1}{\gamma}\nu(j,t)-\frac{j}{2\gamma}\nu'(j,t)\right]\alpha^2,$$

$$\rho_{\perp,T}^-\approx\epsilon_y\nu(j,t)+[(\epsilon_x-\epsilon_y)\nu(j,t)+j(\epsilon_x-\epsilon_y/2)\nu'(j,t)]\alpha^2,$$

$$\rho_{\parallel,L}^+\approx\gamma+[j(\epsilon_x-\epsilon_y)\nu'(\epsilon_x\gamma j,t)+(1/\gamma)\nu(\epsilon_x\gamma j,t)-\gamma]\Delta\alpha^2,$$

$$\rho_{\perp,L}^-\approx\epsilon_x\nu(\epsilon_x\gamma j,t)+\left[(\epsilon_y-\epsilon_x)\nu(\epsilon_x\gamma j,t)+\frac{j}{\gamma}\nu'(\epsilon_x\gamma j,t)+\frac{j^2\epsilon_x}{2}\nu''(\epsilon_x\gamma j,t)\right]\Delta\alpha^2,$$

where the prime denotes the derivative $\nu'(f,t)\equiv\partial\nu(f,t)/\partial f$. The resistivities $\rho_\parallel^-$ and $\rho_\perp^+$, equal to zero in

the T and L geometries, vary linearly with α and $\Delta\alpha$ for small deviations from the T and L geometries. The more unstable of them (in both the L and the T geometry) is the even transverse resistivity $\rho_{\perp}^{+}$ in the TAFF regime of vortex motion transverse to the twins, where $\rho_{\perp,T}^{+} \approx -\gamma\alpha$ in the T geometry and $\rho_{\perp,L}^{+} \approx -\gamma\Delta\alpha$ in the L geometry. The resistivities $\rho_{\parallel}^{+}$ and $\rho_{\perp}^{-}$ vary quadratically in α and $\Delta\alpha$ from their values in the T and L geometries respectively. The relative deviation of the resistivity for a small deviation from the T and L geometries is negligibly small for $\rho_{\perp}^{-}$, but for $\rho_{\parallel}^{+}$ it is of the order $\Delta\rho_{\parallel,T}^{+}/\rho_{\parallel,T}^{+} \sim \alpha^2/\nu(j,t)$ in the T geometry and $\Delta\rho_{\parallel,L}^{+}/\rho_{\parallel,L}^{+} \sim \Delta\alpha^2$ in the L geometry. Thus, $\rho_{\parallel,T}^{+}$ is the most unstable in the TAFF regime of vortex motion transverse to the twins, where $\nu(j,t) \ll 1$. The physical reason for this instability is the rapid variation of the angle β from $\pi/2$ in the T geometry, where $v_y = 0$, to the angle $\alpha \ll 1$ corresponding to the G regime with $v_y \gg v_x$. The stability of the L geometry is physically clear from the fact that for $\Delta\alpha \ll 1$ the angle β varies hardly at all, i.e., the direction of the velocity vector $\mathbf{v}$ varies only slightly (in contrast to the case of the T geometry).

As was stated above, in an actual experiment small deviations $\Delta\alpha$ of the angle α from the values $\alpha = 0, \pi/2$ corresponding to the T and L geometries are always present. Utilizing experimental measurements of $\rho_{\perp,\parallel}^{\pm}$, these deviations can be found using the following scheme. First, neglecting small quadratic contributions in α and $\Delta\alpha = \pi/2 - \alpha$ to the resistivities $\rho_{\parallel}^{+}$ and $\rho_{\perp}^{-}$ (in the region where they are stable), it is possible to solve the scaling problem using formulas (41) and (42), i.e., the reconstruct the values of the constants $\epsilon_x, \epsilon_y, \gamma$ and the function $\nu(x)$. Knowing this, from the formulas for the resistivities $\rho_{\parallel}^{-}$ and $\rho_{\perp}^{+}$, which vanish in the T and L geometries and are linear in α and $\Delta\alpha = \pi/2 - \alpha$ for small deviations from the T and L geometries it is possible to find the corresponding values of α and $\Delta\alpha$. The self-consistency of this scheme is checked by calculating the quadratic corrections in α and $\Delta\alpha$, which should be small relative to the main contribution in the T and L geometries.

4. CONCLUSION

In the present work we have theoretically examined the strongly nonlinear resistive behavior of the two-dimensional vortex system of a superconductor with unidirectional twins as a function of the transport current density j, the temperature t, and the angle α between the directions of the current and the twins. The nonlinear (in j) resistive behavior of the anisotropic vortex ensemble can be caused by factors of both an electronic and a ''pinning'' origin. Digressing here from the possible nonlinearity of the electronic viscosity of superconductors (see, however, Refs. 26 and 27), in this work we take as the one source of nonlinearity the presence of intrinsically anisotropic pinning on twins. It is physically obvious that such pinning at low enough temperatures leads to anisotropy of the vortex dynamics since it is much easier for them to move along the twins (the G effect in the FF regime, which is linear in the current) than in the perpendicular direction, where it is necessary for them to overcome the pinning potential barrier on the twins, which also is a source of resistive nonlinearity. If under variation of one of the ''external'' parameters j, t, α the intensity of manifestation of the indicated nonlinearity is weakened, then (neglecting the usually small Hall terms) this weakening will lead to an ''effective isotropization'' of the vortex dynamics, i.e., to a convergence (and in the limit of the absence of nonlinearity and electronic anisotropy, to coincidence) of the directions of the mean velocity vector of the vortices and the Lorentz force.

It is physically clear that the current, temperature, and angle α have a qualitatively different effect on the weakening of pinning on twins and the corresponding transition from anisotropic vortex dynamics to isotropic. With growth of j the Lorentz force F_L grows and the height of the potential barrier $\delta(j)$ decreases, so that for $j \geqslant j_{cr}$ (where j_{cr} is the crossover current of the indicated transition, whose width grows with growth of t) this barrier essentially disappears. The quantity j_{cr} depends on α by virtue of the fact that the probability of overcoming the barrier δ is governed not by the magnitude of the force F_L, but only by its transverse component $F_L \cos\alpha$, so that $j_{cr}(\alpha) = j_{cr}(0)/\cos\alpha$ grows with growth of α. Since an increase in the temperature t always increases the probability of overcoming the pinning barrier $\delta(j)$, the transition to isotropization of the vortex dynamics is that much steeper in t, the smaller is $\delta(j)$.

In order to theoretically analyze the above-described physical picture of a nonlinear anisotropic resistive response, Sections 2 and 3 employed a comparatively simple, but at the same time quite realistic, planar model of stochastic pinning. It allows one to distinctly separate the pinning viscosity from the electronic (drift and Hall) viscosity and to reduce the calculations to the evaluation of analytical formulas (32) and (33), which have a simply physical interpretation. A distinguishing feature of this model is the possibility, within the framework of a unified approach, to describe consistently the nonlinear transition from the anisotropic dynamics of a vortex system (for currents $j \ll j_{cr}(\alpha)$ at relatively low temperatures) to isotropic behavior (for currents $j > j_{cr}(\alpha)$ at relatively high temperatures). In the model under consideration this approach corresponds (for $t > 0$) to a substantially nonlinear crossover from the linear low-temperature TAFF regime to the ohmic FF regime of vortex motion. Physically, this stochastic model provides the most adequate description of resistive behavior of a strongly pinned vortex liquid[28] in the immediate vicinity of its melting temperature $T_m(H)$, which depends on the magnetic field strength H (Ref. 11). Note that most resistive experiments on pinning on twins are carried out in this temperature range[4–9] since, on the one hand, here it is already possible in fact to neglect isotropic pinning on point defects and, on the other, the amplitude of the resistive response is still large enough that one does not have to use squid picovoltmeters.

Proceeding now to a brief description of the main theoretical results, we note here that an analytical representation of the nonlinear resistive response of the investigated system in terms only of elementary functions was possible thanks to the use of a simple but physically realistic model of anisotropic pinning on twins (see Sec. 3 and Fig. 3). The exact solution obtained made it possible for the first time to consistently analyze not only the qualitatively clear dynamics of

the nonlinear G effect, but also the nontrivial question of the interaction of guided vortex motion along twins and the Hall effect. The most important result (and unexpected from the point of view of ''linear'' intuition) in our opinion is the conclusion that the appearance of novel (specific to anisotropic pinning in the linear[9,10] variant) $\rho_{\perp}^{+}$- and $\rho_{\parallel}^{-}$-effects does not require the existence of linear anisotropy of an electronic origin, i.e., anisotropy of the electronic viscosity tensor $\hat{\eta}$ and the Hall conductivity tensor $\hat{\alpha}$ (see Sec. 2). The nonlinear formulas (32) and (33) in agreement with physical intuition (now already nonlinear) clearly demonstrate that the most natural and ''sufficient'' reason for the relatively large novel $\rho_{\perp}^{+}$- and $\rho_{\parallel}^{-}$-effects is the anisotropy of pinning on twins. At comparatively low temperatures and weak currents $[j \ll j_{cr}(\alpha) \equiv j(0)/\phi(\alpha)$, see Eq. (29)] it leads to the realization of a quite intense (over a wide interval of angles around $\alpha = \pi/4$) guided vortex motion along the twins in the TAFF regime, i.e., to the appearance of $\rho_{\perp}^{+}$-effects, and at currents $j \approx j_{cr}(\alpha)$, to the appearance of characteristic maxima in the curves of the Hall components of the resistivity tensor, $\rho_{\perp}^{-}$ and $\rho_{\parallel}^{-}$ [see Subsections 3.1. and 3.3. and Figs. 10, 11, and 13a].

A completely novel result of the present work is also contained in formulas (32) and (33). It is a quantitative description of the interaction of the G effect and the Hall effect, which is valid for all values of the Hall constants regardless of magnitude or sign. Formally, this interaction arises as a result of the fact that in the case of anisotropic pinning on twins the motive force F [see Eq. (9)], which determines the probability of overcoming the potential barrier (and therewith also determines the magnitude of the component of the vortex velocity perpendicular to the twins), is the sum of two forces. The first of these is the transverse component of the Lorentz force, $F_L \cos\alpha$, and the other is the transverse component of the Magnus force, $F_M^{\perp} = n\epsilon_x \gamma F_L \sin\alpha$ [see Eq. (9)], which is proportional to the longitudinal (relative to the twins) component of the velocity of guided vortex motion. This second force $F_M^{\perp}$, which changes its sign (relative to the sign of F_L) upon reversal of the sign of the external magnetic field, is the reason for the appearance of new, Hall-like in their origin, ν^{-}-terms in the formulas for the resistive responses (32) and (33). The key point in the physical interpretation of these formulas is our treatment of the function $\nu(f,t,\varepsilon)$ as the probability of overcoming the potential barrier of the twins, from which follows an understanding of the evolution of the functions associated with it, $\nu^{\pm}$ (see Subsection 3.1.), as functions of the magnitude of the current density j, temperature t, and angle α. Note that this treatment is not a unique property of the stochastic model of anisotropic pinning considered in this work, but can also be consistently realized within the framework of the nonlinear phenomenological approach[17,18] under much broader assumptions, including, in particular, an account of the intervortex interaction.

If, as is usually the case in experiment,[2,3,5] the Hall constants are sufficiently small $(\epsilon_x, \epsilon_y \ll 1)$, then formulas (32) and (33) simplify substantially since under these conditions $\nu^{-}(f_0) \sim \epsilon_x (d\nu/df_0) \ll 1$ (see Subsection 3.1.). If, in addition, $\gamma = 1$ and $\epsilon_x = \epsilon_y = \epsilon \ll 1$, then the interaction of the G effect and the Hall effect becomes ''unilateral'': formulas (38) for $\rho_{\parallel}^{+}$ and $\rho_{\perp}^{+}$ (and consequently for $\cot\beta$, see Subsection 3.2.) cease to depend on the Hall effects whereas in formulas (39) for $\rho_{\parallel}^{-}$ and $\rho_{\perp}^{-}$ the ν^{+} functions characteristic of the G effect are preserved.

In conclusion, it should also be noted that the curves of the temperature dependences $\rho_{\parallel,\perp}^{\pm}(t)$ presented in this work should not be understood literally since the entire list of parameters of the problem (components of the tensors $\hat{\eta}$ and $\hat{\alpha}$, the pinning potential U_0) can also depend implicitly on the temperature, which enters into the formulas for the resistivities in terms of $\nu(f,t,\varepsilon)$. Therefore, these curves should be understood only as an illustration of possible effects of temperature depinning of vortices on twins, valid under the condition that the indicated implicitly t-dependent parameters vary weakly in those temperature intervals that correspond to singularities of the functions $\nu^{\pm}(t)$. Nevertheless, we would still like to note that in a number of cases for the model under consideration [see Fig. 13b] the temperature dependence of the Hall magnetoresistivity $\rho_{\perp}^{-}$ for $j_0 = \text{const}$ is strongly reminiscent of the ''anomalous'' Hall behavior in HTSC's, which has been actively discussed in a number of experimental and theoretical works.[2,3,29]

*)E-mail: Valerij.A.Shklovskij@univer.kharkov.ua

[1] S. Fleshler, W.-K. Kwok, U. Welp *et al.*, Phys. Rev. B **47**, 14 448 (1993).
[2] T. R. Chien, T. W. Jing, N. P. Ong, and Z. Z. Wang, Phys. Rev. Lett. **66**, 3075 (1991).
[3] T. R. Chien, D. A. Brawner, Z. Z. Wang, and N. P. Ong, Phys. Rev. B **43**, 6242 (1991).
[4] A. V. Bondarenko, M. A. Obolenskii, R. V. Vovk, *et al.*, in *Proceedings of the Seventh IWCC on Superconductors*, Alpbach, Austria, edited by H. W. Weber (World Scientific, Singapore, 1994), p. 177.
[5] A. A. Prodan, V. A. Shklovskij, and V. V. Chabanenko *et al.*, Physica C **302**, 271 (1998).
[6] V. V. Chabanenko, A. A. Prodan, V. A. Shklovskij *et al.*, Physica C **314**, 133 (1999).
[7] H. Ghamlouch and M. Aubin, Physica C **269**, 163 (1996).
[8] H. Ghamlouch and M. Aubin, R. Gagnon, and L. Taillefer, Physica C **275**, 141 (1997).
[9] É. B. Sonin, A. L. Kholkin, Fiz. Tverd. Tela (Leningrad) **34**, 1147 (1992) [Sov. Phys. Solid State **34**, 610 (1992)].
[10] E. B. Sonin, Phys. Rev. B **48**, 10487 (1993).
[11] G. Blatter, M. V. Feigelman, V. B. Geshkenbein *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
[12] G. W. Crabtree, G. K. Leaf, H. G. Kaper *et al.*, Physica C **263**, 401 (1996).
[13] J. Groth, C. Reichhardt, C. J. Olson *et al.*, Phys. Rev. Lett. **77**, 3625 (1996).
[14] B. U. Zhu, J. Dong, D. Y. Xing, and Z. D. Wang, Phys. Rev. B **57**, 5075 (1998).
[15] E. Yu. Klimenko, S. V. Shavkin, P. V. Volkov, Zh. Éksp. Teor. Fiz. **112**, 1055 (1997) [JETP **85**, 573 (1997)].
[16] Y. Mawatari, Phys. Rev. B **56**, 3433 (1997).
[17] V. A. Shklovskiĭ, Fiz. Nizk. Temp. **23**, 1134 (1997) [Low Temp. Phys. **23**, 853 (1997)].
[18] V. A. Shklovskiĭ, Fiz. Nizk. Temp. **25**, 153 (1999) [Low Temp. Phys. **25**, 109 (1999)].
[19] C. Villard, G. Koren, D. Cohen, and E. Polturac, Phys. Rev. Lett. **77**, 3913 (1996).
[20] P. Berghuis, E. DiBartolomeo, G. A. Vagner, and G. E. Evetts, Phys. Rev. Lett. **79**, 2332 (1997).
[21] A. K. Niessen and C. H. Weijsenfeld, J. Appl. Phys. **40**, 384 (1969).

[22] V. M. Vinokur, V. B. Geshkenbein, M. V. Feigel'man, and G. Blatter, Phys. Rev. Lett. **71**, 1242 (1993).
[23] B. Shen and J. Dong, Phys. Rev. B **44**, 10 206 (1991).
[24] O. V. Usatenko and V. A. Shklovskij, J. Phys. A **27**, 5043 (1994).
[25] J. M. Harris, Y. F. Yan, O. K. S. Tsui *et al.*, Phys. Rev. Lett. **73**, 1711 (1994).
[26] A. I. Larkin and Yu. N. Ovchinnikov, Zh. Éksp. Teor. Fiz. **68**, 1915 (1975) [Sov. Phys. JETP **41**, 960 (1975)].
[27] A. I. Larkin and Yu. N. Ovchinnikov, in *Nonequilibrium Superconductivity*, edited by D. Langenberg and A. I. Larkin (North-Holland, Amsterdam, 1986), Ch. 11.
[28] V. M. Vinokur, V. B. Geshkenbeĭn, A. I. Larkin, and M. Feĭgel'man, Zh. Éksp. Teor. Fiz. **100**, 1104 (1991) [Sov. Phys. JETP **73**, 610 (1991)].
[29] S. J. Hagen, A. W. Smith, M. Rajeswari *et al.*, Phys. Rev. B **47**, 1064 (1993).

Translated by Paul F. Schippnick

Negative persistent photoconductivity in GaAs(δ-Sn) structures

V. A. Kul'bachinskiĭ,*) V. G. Kytin, R. A. Lunin, A. V. Golikov, A. V. Demin, A. S. Bugaev, and A. P. Senichkin

M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

A. De Visser and R. T. F. Van Schaijk

Van der Waals–Zeeman Institute, University of Amsterdam, 1018XE, Amsterdam, the Netherlands
(Submitted 21 May 1999)
Zh. Éksp. Teor. Fiz. **116**, 2130–2139 (December 1999)

The effect of illumination with various wavelengths λ (770 nm$<\lambda<$1120 nm) on the conductivity of GaAs structures with tin δ-doping of the vicinal faces was investigated in the temperature range 4.2–300 K. Negative persistent photoconductivity was found in strongly doped samples. It was shown on the basis of the results of investigations of the Hall and Shubnikov–de Haas effects that the negative photoconductivity is due to a large decrease in the electron mobility with increasing electron density. The decrease of electron mobility is explained by ionization of DX centers, which destroys the spatial correlation in the distribution of positively charged donors and negatively charged DX centers. © *1999 American Institute of Physics.* [S1063-7761(99)01912-5]

1. INTRODUCTION

Delta-doped semiconductors, where the impurity atoms are located in a layer one or several atomic monolayers thick, are now an object of intense experimental and theoretical investigations.[1] The charges of the dopants in the δ layer create a potential well, as a result of which a structure with two-dimensional electrons is formed. In δ-doped structures with a high impurity density, electrons fill many size-quantization subbands. The behavior of two-dimensional electrons in such systems in electric and magnetic fields is much more complicated than in ordinary low-dimensional structures with a single filled subband. Intersubband electron scattering is important, and electron mobilities in each subband are different.

Interest in the study of δ-doped semiconductors is justified not only from the scientific standpoint but also by the possibility of practical applications of such materials. Delta-doping is an example of an extremely narrow doping profile which gives high current-carrier densities. Even though high dopant concentrations are important in nanoelectronics, the mechanisms limiting the maximum achievable free-electron density at high doping levels are still not completely understood.

It should be noted that ordinarily silicon is used to produce *n*-type δ-layers in gallium arsenide, and it is important to investigate and compare the electronic properties of δ-layers with different dopants, for example, tin. As a donor impurity, tin is less amphoteric than silicon,[2] and the use of tin should make it possible to obtain higher densities of two-dimensional electrons in a δ-layer.

Quasi-one-dimensional and one-dimensional electronic systems, produced on the basis of two-dimensional systems, are now being investigated increasingly more actively.[3] Quantization of the conductivity as a function of the width of the conducting channel, quantum oscillations of a new type in a magnetic field, and other fundamental effects are observed in such systems. Ordinarily, submicron electron lithography is used to limit the lateral size in two-dimensional systems.[3] To obtain systems with quasi-one-dimensional electronic channels several tens and less nanometers in size, a promising method is to grow structures on the vicinal surface of gallium arsenide,[4–10] i.e., on a surface tilted from the basal plane [for example, (001)] by a small angle, as a result of which it becomes stepped.

An important experimental fact is that, usually, persistent positive photoconductivity is observed in structures with δ-doped layers at low temperatures, i.e., under illumination the conductivity of the structures increases and remains for a long time.[11] Various models of persistent photoconductivity exist. One model involves the photoionization of deep levels, called DX centers. It is believed that a DX center is a negatively charged localized state, which traps two free electrons.[12–15] In a different model, the separation of photoionized electron–hole pairs, so that the electrons remain in the δ layer while the holes escape into the interior volume, is taken into account. In this case, a logarithmic decay of the persistent photoconductivity is expected.[16,17] Conductivity anisotropy[4–6,18] and positive photoconductivity and its quenching by a strong electric field[19,20] are observed in GaAs structures with tin δ-doping of the vicinal face.

In the present paper we report the results of an investigation of negative persistent photoconductivity, which we observed in GaAs structures with tin δ-doping of the vicinal faces, in a wide range of photon energies in the temperature range 4.2 K$<T<$300 K. Mechanisms of negative and positive persistent photoconductivity in δ-doped GaAs structures are discussed.

TABLE I. Resistivity ρ, Hall density n_H, and Hall mobility μ_H of electrons and the sum Σn_{SdH} of the electron densities in all subbands, which is determined from the Shubnikov–de Haas effect, at temperature $T=4.2$ K.

Sample No.	Form of illumination	ρ, Ω	n_H, 10^{12} cm^{-2}	μ_H, cm^2/(V·s)	Σn_{SdH}, 10^{12} cm^{-2}
	In dark	202	31.5	981	26.2
1	$\lambda=791$ nm	198	31.6	1000	26.3
	$\lambda>850$ nm	232	30.4	886	27.9
	In dark	374	25.8	648	25.9
2	$\lambda=791$ nm	367	24.9	683	26.0
	$\lambda>850$ nm	417	26.0	576	29.6
	In dark	1330	8.03	586	8.28
3	$\lambda=791$ nm	1173	8.62	618	8.39
	$\lambda>850$ nm	1235	8.81	574	8.38

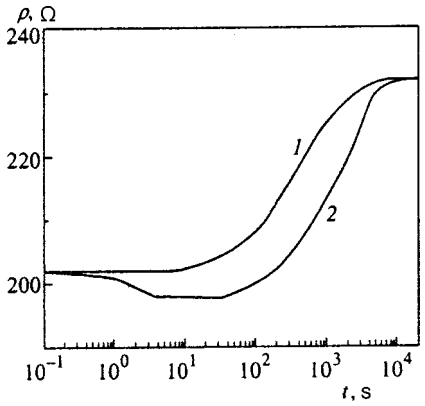

FIG. 1. Resistance of sample 1 versus the irradiation time for light with wavelength $\lambda\geqslant1120$ nm (*1*) and $\lambda=791$ nm (*2*) and illumination intensity 10 μW/cm^2.

2. SAMPLES AND MEASUREMENT PROCEDURE

The experimental structures were grown by molecular-beam epitaxy. A 0.45 μm thick undoped GaAs buffer layer was grown on a semi-insulating GaAs{Cr} substrate, disoriented by 3° from the (001) plane toward to the (110) plane. A system of steps formed on the vicinal face of the crystal. A 1 ML high step is 5.3 nm wide. Next, growth was stopped and a definite quantity of tin was precipitated onto the surface. After the tin was deposited, a 40 nm thick gallium arsenide layer was grown at low epitaxy temperatures $\approx$450 °C, which should make it possible to preserve a nonuniform distribution of tin. Then a 20 nm thick GaAs layer, doped with silicon to 2×10^{18} cm^{-3}, to fill the surface states was grown.

Samples in the form of double Hall bridges with the conducting channel oriented along the [110] direction and along the [−110] direction were prepared for resistivity and Hall effect measurements. In all samples the resistivity in the [110] direction was less than in the [−110] direction. At the same time, the influence of illumination of the conductivity was qualitatively independent of the direction of the conducting channel, so that in the present paper we report the results obtained for samples with the conducting channel oriented along the [110] direction. To investigate the photoconductivity the samples were illuminated with an incandescent lamp through a monochromator, which extracted radiation with wavelength ranging from 770 nm to 1120 nm with spectral linewidth 2.6 nm, and through various filters. Certain parameters of the experimental samples at $T=4.2$ K are presented in Table I.

3. EXPERIMENTAL RESULTS

The photoconductivity of the experimental samples depends on the wavelength of the incident light. In addition, the irradiation intensity was found to be important for observing negative photoconductivity with reasonable durations of the experiment. The typical dependence of the resistivity of sample 1 on the illumination time is displayed in Fig. 1. As one can see in this figure, when the sample is irradiated with $\lambda=791$ nm light (curve *2*), at first positive photoconductivity is observed and then, under further illumination, it changes to negative. For $\lambda\geqslant1120$ nm (curve *1*) only negative photoconductivity is observed. The dependence of the change in the resistivity as a function of the wavelength of the incident light for various radiation intensities for sample 1 is presented in Fig. 2. The same dependence is observed for sample 2. The resistivity of these samples, after cooling in the dark to temperature 4.2 K and illumination with monochromatic light with wavelength less than 835 nm and intensity less than 20–70 nW/cm^2, decreases (positive photoconductivity) and saturates in $\approx$30 min (points *1*, Fig. 2). Under further illumination of the same samples with radiation with wavelength 786 nm$<\lambda<$796 nm and intensity $I\approx10$ μW/cm^2, the resistivity at first decreases to a minimum (point *3* in Fig. 2), and then starts to grow, reaching at saturation a value (point *4* in Fig. 2) greater than the value in the dark (negative photoconductivity). Under continuous irradiation with light with wavelength greater than 835 nm and intensity $I\approx20$ nW/cm^2 (this intensity of light is obtained by illuminating through a monochromator) the resistivity of the samples remains unchanged for

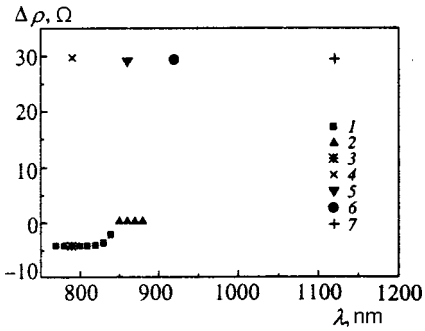

FIG. 2. Resistivity change $\Delta\rho$ of sample 1, measured from the dark resistivity, versus the wavelength λ of the incident radiation: *1*—illumination through a monochromator with intensity $I\approx70$ nW/cm^2 for 1 h; *2*—illumination through a monochromator with intensity $I\approx20$ nW/cm^2 for 1 h; *3*—illumination through a filter 786 nm $<\lambda<$796 nm with $I\approx10$ μW/cm^2 (minimum value of the resistivity); *4*—illumination through the same filter and with the same intensity $I\approx10$ μW/cm^2 to saturation; *5*—$\lambda>850$ nm filter to saturation, $I\approx60$ μW/cm^2; *6*—920 nm $<\lambda<$930 nm filter to saturation, $I\approx10$ μW/cm^2; *7*—$\lambda\geqslant1120$ nm filter, $I\approx60$ μW/cm^2 to saturation.

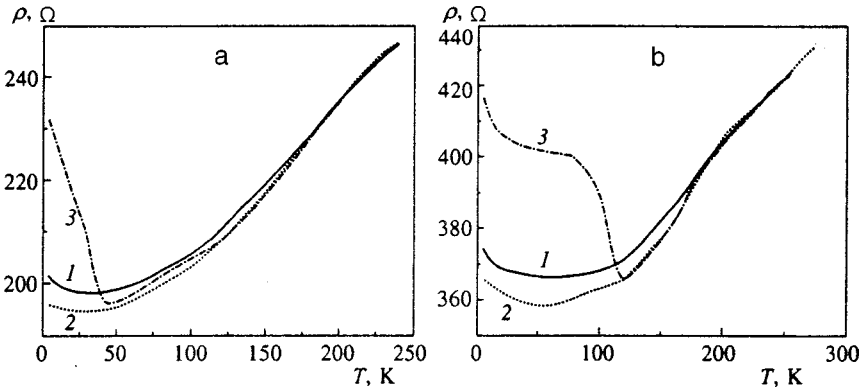

FIG. 3. Temperature dependences of the resistivity of samples 1 (a) and 2 (b) in the dark (curve *1*) and after illumination at $T=4.2$ K with $\lambda=791$ nm light (curve *2*) (the illumination was switched off after the minimum resistivity was reached) and $\lambda>1120$ nm light (curve *3*) (the illumination was switched off after resistivity saturation was reached).

at least 5 hs (points *2* in Fig. 2). When the samples 1 and 2 are illuminated with radiation with wavelength greater than 850 nm and intensity $I\approx 10-60$ μW/cm^2, the resistivity increases from the dark value and reaches at saturation (points *5*, *6*, and *7* in Fig. 2) the same value as for illumination with intense light with wavelength less than 835 nm. In this case an initial decrease of the resistivity is not observed. In the negative photoconductivity regime, the resistivity remains unchanged for at least 5 h after the illumination is switched off at temperature 4.2 K, i.e., negative persistent photoconductivity is observed. The critical photon energy corresponding to wavelength 835 nm is approximately 35 meV less than the GaAs band gap. This corresponds to the energy required to transfer electrons from shallow acceptors into the conduction band.[21]

The temperature dependences of the resistivity ρ of heavily doped samples 1 and 2, measured in the dark and after various forms of illumination at $T=4.2$ K and heating at a rate of 3 K/min, are presented in Fig. 3. When sample 1 is heated, after illumination at $T=4.2$ K (up to saturation of the resistivity) through a filter transmitting light with wavelength greater than 1120 nm, as the temperature increases, the resistivity decreases and crosses the dark curve $\rho(T)$ at $T\approx 40$ K (Fig. 3a). For sample 2 the heating curve, after illumination with ''long-wavelength'' radiation (we shall call radiation with energy less than the band gap in gallium arsenide long-wavelength radiation) crosses the dark temperature dependence at $T\approx 120$ K. After crossing the dark dependence $\rho(T)$, the resistivity curve after illumination with ''long-wavelength'' radiation lies somewhat above the resistivity curve after illumination with ''short-wavelength'' radiation (radiation with photon energy greater than the band gap in gallium arsenide), switched on at the moment when the resistivity reaches its minimum value. In all samples persistent photoconductivity exists up to temperatures ≈ 180 K.

For the less heavily doped sample 3 the resistivity decreases after illumination with ''short-wavelength'' and ''long-wavelength'' radiation, but the values of the resistivity at which saturation occurs are different for these two forms of illumination—the resistivity decreases more strongly for illumination with ''short-wavelength'' radiation. After illumination at temperature 4.2 K is switched off, the resistivity of this sample slowly returns to the dark value in several hours. The temperature dependences of the resistivity of sample 3, which were measured in the dark, and with heating after irradiation at $T=4.2$ K with light with various wavelengths are presented in Fig. 4.

Besides the temperature dependences of the resistivity, in the present work we investigated the Shubnikov–de Haas effect to determine the electron density in the size-quantization subbands. The magnetoresistivities for samples 1 and 2, respectively, at $T=4.2$ K measured in the dark (curves *1*) and after various forms of illumination (curves *2* and *3*) are displayed in Figs. 5(a) and 6(a), and the Fourier spectra corresponding to the oscillations are displayed in Figs. 5(b) and 6(b). The Shubnikov–de Haas effect showed that for positive persistent photoconductivity the frequencies (proportional to the two-dimensional electron densities in the size-quantization subbands), observed in the Fourier spectrum of the magnetoresistivity oscillations, remain practically unchanged (curves *2* in Figs. 5 and 6), while for negative persistent photoconductivity the frequencies increase (curves *3* in Figs. 5 and 6) compared with the dark values (curves *1* in Figs. 5 and 6). The quantum electron mobilities[22] in the size-quantization subbands in the presence of positive persistent photoconductivity increase by 10–20% (primarily in the upper subbands) compared with the dark values, and in the presence of negative persistent photoconductivity they decrease strongly (by a factor of 1.5) in the lower subbands. This change of the quantum electron mobilities affects the height and width of the peaks in the Fourier spectrum: For illumination with ''long-wavelength'' radia-

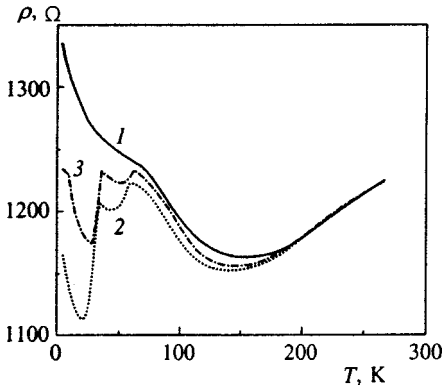

FIG. 4. Temperature dependences of the resistivity of sample 3 in the dark (curve *1*) and after illumination at $T=4.2$ K with $\lambda=791$ nm light (curve *2*) and $\lambda=850$ nm light (curve *3*). The illumination was switched off after resistivity saturation was reached.

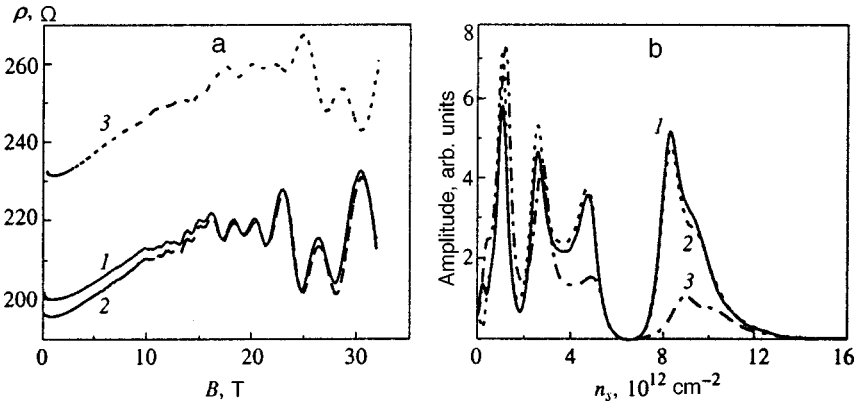

FIG. 5. (a)—Magnetoresistivity oscillations of sample 1 at $T=4.2$ K in the dark (*1*) and after irradiation with $\lambda=791$ nm light (*2*) (the illumination was switched off after the minimum resistivity was reached) and $\lambda>850$ nm light (*3*) (the illumination was switched off after resistivity saturation was reached). (b)—Amplitude of the Fourier transform of the Shubnikov–de Haas oscillations versus the density of two-dimensional electrons for sample 1 in the dark and after a corresponding illumination.

tion the height of the peaks corresponding to the lower subbands is much smaller and the width is much greater than in the dark [Figs. 5(b) and 6(b)]. The Hall effect investigations confirmed that the change in the resistivity of the samples after illumination is determined primarily by the change in the electronic Hall mobilities (see Table I). For illumination with ''short-wavelength'' radiation the Hall mobility averaged over all subbands increases, and the ''long-wavelength'' mobility decreases compared with the values in the dark.

4. DISCUSSION

The long-time character of the negative photoconductivity observed in the heavily doped samples 1 and 2 gives a basis for inferring that this effect is associated with filling and emptying of DX centers. The increase in the total electron density, determined from the Shubnikov–de Haas effect, under illumination is 6.5% for sample 1 and 14% for sample 2 (see Table I). This appears to be due to the ionization of filled DX centers.[12] It is believed that a DX center is a negatively charged localized state, trapping two free electrons.[13–15] The Coulomb interaction between the positively charged shallow donors and DX centers gives rise to a correlation in the spatial distribution of charged impurity atoms and decreases scattering of electrons by them.[23–27] Ionization of DX centers by the light destroys the correlation and this should decrease the electron mobility[23,28] and lead to the above-described negative persistent photoconductivity effect. It should also be noted that, possibly, the filled tin DX centers in GaAs are still neutral (they contain one electron each)[29] and weakly scatter electrons, while under illumination the tin atoms become positively charged, and the observed decrease of the electron mobility is simply due to increase in the number of scattering centers.

The positive persistent photoconductivity can be explained as follows. Under illumination with ''short-wavelength'' radiation, electron–hole pairs are produced, and they are separated by a weak electric field which exists at equilibrium in the *i*-GaAs buffer layer between the substrate and the δ-layer. In the process, the electrons slide down into the δ-layer, and the holes neutralize the charged acceptors, which are present in small quantities in *i*-GaAs, or slide down into the substrate. The characteristic acceptor density in *i*-GaAs is 4×10^{14} cm^{-3}, which for buffer layer thickness $d=0.45$ μm corresponds to a two-dimensional density 1.8×10^{10} cm^{-2}. As a result of the spatial separation of the electrons and holes in the buffer layer, an additional electric field arises and completely compensates the initial field, the bands are rectified,[11] and the electron–hole pairs no longer separate. The additional charge-carrier density required for such a nonequilibrium situation to arise is approximately

$$\Delta n_s=\frac{\varepsilon_0\varepsilon}{ed}\Delta V,$$

where $\Delta V=0.75$ V is the potential corresponding to a deep chromium level in the substrate[30] and $d=0.45$ μm is the buffer layer thickness in the experimental structures. The

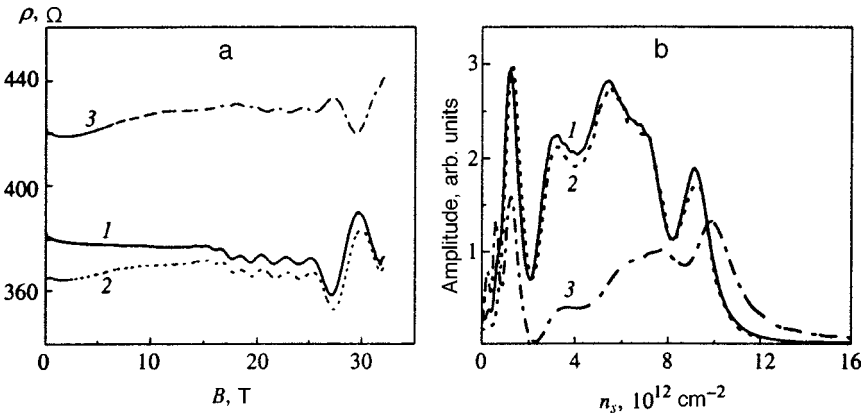

FIG. 6. (a)—Magnetoresistivity oscillations of sample 2 at $T=4.2$ K in the dark (*1*) and after irradiation with $\lambda=791$ nm light (*2*) (the illumination was switched off after the minimum resistivity was reached) and $\lambda>850$ nm light (*3*) (the illumination was switched off after resistivity saturation was reached). (b)—Fourier spectrum of the Shubnikov–de Haas oscillations for sample 2 in the dark and after corresponding illumination.

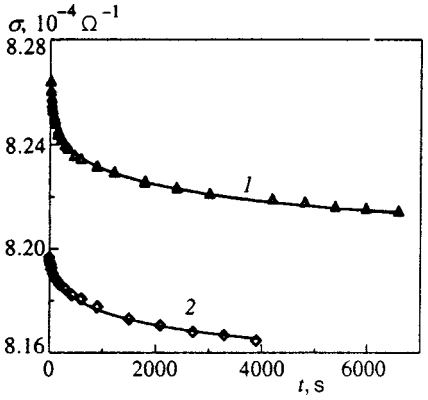

FIG. 7. Temporal relaxation of the positive photoconductivity of sample 3 after illumination at $T=77$ K with $\lambda=791$ nm light (*1*) and $\lambda>1120$ nm light (*2*. Solid lines—fit of the formula (1).

value of Δn_s obtained in this manner is 1.2×10^{11} cm^{-2}, which corresponds to a small increase of the Shubnikov electron densities in the samples (for example, Δn_s for sample 3 is 1.1×10^{11} cm^{-2}). The rectification of the conduction band bottom increases the effective width of the quantum well. At the same time, the wave functions of the electrons are concentrated farther away from the charged donors of the δ-layer. This decreases the electron scattering and increases the electron mobility. This should affect the electron mobility most strongly precisely in the upper subbands, since it is the wave functions of electrons in the upper subbands that are most sensitive to such an increase in the width of the potential well.[22]

Under ''long-wavelength'' illumination, together with ionization of DX centers, the electrons are excited from deep levels of chromium in the substrate and slide down into the δ-layer. At the same time, neutralization of the acceptors in the *i*-GaAs buffer layer by holes, just as with ''short-wavelength'' illumination, does not occur. The neutralization of charged acceptors under illumination with ''short-wavelength'' radiation results in an additional increase of the electron mobilities in the upper subbands,[22,31] and for this reason in the present case the resistivity of the samples decreases more strongly than for illumination with ''long-wavelength'' radiation.

The fact that positive persistent conductivity arises because of the spatial separation of the photogenerated electrons and holes is confirmed by measurements of the temporal relaxation of the photoconductivity. The decrease of the conductivity with time after illumination is switched off is described well by a logarithmic time dependence, which is characteristic for spatial charge separation.[16,17] Figure 7 displays the relaxation of the positive persistent photoconductivity in sample 3 and a fit of the function

$$\sigma(0)-\sigma(t)=A\ln(1+t/\tau), \tag{1}$$

obtained in Ref. 16 and valid for the initial time interval, to this photoconductivity. The relaxation parameter τ for sample 3 illuminated with ''short-wavelength'' radiation at $T=77$ K is 19 s, while for illumination with ''long-wavelength'' radiation it is 68 s. At $T=4.2$ K the relaxation parameter increases and is 23 s for ''short-wavelength'' radiation and several tens of minutes for ''long-wavelength'' radiation. Neutralization of charged acceptors under illumination with ''short-wavelength'' radiation results in faster relaxation of the positive persistent photoconductivity than for illumination with ''long-wavelength'' radiation, because of the recombination of the electrons in the δ-layer and the close-lying acceptors. In heavily doped samples 1 and 2, in the regime of positive persistent photoconductivity, the relaxation times of the photoconductivity are close to the relaxation times in sample 3.

5. CONCLUSIONS

In summary, we have investigated for the first time negative persistent conductivity in GaAs structures with tin δ-doped vicinal faces. This effect is observed only in samples with a high level of doping. An increase in resistivity is accompanied by an increase in the electron density and a substantial decrease of electron mobility, which is what determines the negative sign of the photoconductivity. The increase in electron density is a consequence of the ionization with deep metastable levels—DX centers. The decrease of the mobility could be due to breakdown of the spatial correlation in the arrangement of positively charged donors and negatively charged DX centers as well as to an increase in the density of positively charged scattering centers, if the DX centers were neutral before ionization.

This work was supported by the Russian Fund for Fundamental Research (Grant 97-02-17396) and the Dutch organization N.W.O.

*)E-mail: kulb@mig.phys.msu.su

[1] *Delta-Doping of Semiconductors*, edited by E. F. Shubert (Cambridge University Press, New York, 1996).
[2] Kin Man Yu, H. P. Lee, and S. Wang, Appl. Phys. Lett. **56**, 1784 (1990).
[3] C. G. Smith, Rep. Prog. Phys. **59**, 235 (1996).
[4] A. D. Visser, V. I. Kadushkin, V. A. Kul'bachinskiĭ *et al.*, JETP Lett. **59**, 363 (1994).
[5] V. I. Kadushkin, V. A. Kulbachinskii, A. P. Senichkin *et al.*, Phys. Low-Dimen. Struct. **1**, 53 (1994).
[6] V. A. Kulbachinskii, N. B. Brandt, V. G. Kytin *et al.*, Phys. Low-Dimens. Semicond. Struct. **12**, 237 (1995).
[7] M. Ramsteiner, J. Wagner, D. Behr, and R. Hey, Appl. Phys. Lett. **64**, 490 (1994).
[8] L. Daweritz, K.-J. Friedland, J. Behrend, and p. Schutzendube, Phys. Status Solidi A **146**, 277 (1994).
[9] L. Daweritz, H. Kostial, R. Hey *et al.*, J. Cryst. Growth **150**, 214 (1995).
[10] B. Etienne, F. Lelarge, Z. Z. Wang, and F. Laruelle, Appl. Surf. Sci. **113–114**, 66 (1997).
[11] V. V. Valyaev, V. L. Gurtovoĭ, D. Yu. Ivanov *et al.*, Zh. Éksp. Teor. Fiz. **113**, 693 (1998) [JETP **86**, 383 (1998)].
[12] D. K. Maude, J. C. Portal, L. Dmowski *et al.*, Phys. Rev. Lett. **59**, 815 (1987).
[13] P. Gibart, D. L. Williamson, J. Moser, and P. Basmaji, Phys. Rev. Lett. **65**, 1144 (1990).
[14] M. Baj and L. H. Dmowski, J. Phys. Chem. Solids **56**, 589 (1995).
[15] M. Hayne, A. Usher, J. J. Harris, and C. T. Foxon, Surf. Sci. **361/362**, 574 (1996).
[16] H. J. Queisser and D. E. Theodorou, Phys. Rev. B **33**, 4027 (1986).
[17] J. Chen, C. H. Yang, and R. A. Wilson, Appl. Phys. Lett. **60**, 2113 (1992).
[18] R. T. F. van Schaijk, A. de Visser, V. A. Kulbachinskii *et al.*, Physica B **256–258**, 243 (1998).

[19] V. A. Kul'bachinskiĭ, R. A. Lunin, E. V. Bogdanov *et al.*, JETP Lett. **63**, 336 (1996).
[20] V. A. Kulbachinskii, R. A. Lunin, E. V. Bogdanov *et al.*, Physica B **229**, 262 (1997).
[21] D. J. Ashen, P. J. Dean, D. T. J. Hurle *et al.*, J. Phys. Chem. Solids **36**, 1041 (1975).
[22] P. M. Koenraad, in *Delta-Doping of Semiconductors*, edited by E. F. Shubert (Cambridge University Press, New York, 1996), Chap. 17, p. 407.
[23] Z. Wilamowski, J. Kossut, T. Suski *et al.*, Semicond. Sci. Technol. **6**, B34 (1991).
[24] D. K. Maude, L. Eaves, and J. C. Portal, Appl. Phys. Lett. **60**, 1993 (1992).
[25] T. Suski, P. Wisniewski, I. Gorczyca *et al.*, Phys. Rev. B **50**, 2723 (1994).
[26] R. Shikler, M. Heiblum, and V. Umansky, Phys. Rev. B **55**, 15427 (1997).
[27] J. M. Shi, P. M. Koenraad, A. F. V. van de Stadt *et al.*, Phys. Rev. B **55**, 13093 (1997).
[28] E. Buks, M. Heiblum, and H. Shtrikman, Phys. Rev. B **49**, 14790 (1994).
[29] A. K. Saxena, Solid-State Electron. **25**, 127 (1982).
[30] A. Chandra, C. E. Wood, D. W. Woodard, and L. F. Eastman, Solid-State Electron. **22**, 645 (1979).
[31] V. P. Evtikhiev, P. S. Kop'ev, M. Yu. Nadtochiĭ, and V. M. Ustinov, Fiz. Tekh. Poluprovodn. **23**, 845 (1989) [Sov. Phys. Semicond. **23**, 530 (1989)].

Translated by M. E. Alferieff

Current transport along the [001] axis of YBCO in low-temperature superconductor—normal metal—high-temperature superconductor heterostructures

F. V. Komissinskiĭ

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia; M. V. Lomonosov Moscow State University, 119899 Moscow, Russia

G. A. Ovsyannikov*)

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia

N. A. Tulina and V. V. Ryazanov

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

(Submitted 27 May 1998; resubmitted 14 July 1999)
Zh. Éksp. Teor. Fiz. **116**, 2140–2149 (December 1999)

The electrophysical properties of heterojunctions several microns in size, obtained by successive deposition of the metal-oxide high-temperature superconductor $YBa_2Cu_3O_x$, a normal metal Au, and the low-temperature superconductor Nb, were studied experimentally. Current flows in the [001] direction of the epitaxial $YBa_2Cu_3O_x$ film. It is shown, by comparing the experimental data with existing theoretical calculations, that for the experimentally realizable transmittances ($D=10^{-5}-10^{-6}$) of the $YBa_2Cu_3O_x$—normal metal boundary the critical current of the entire heterostructure is low (of the order of the fluctuation current) because of a sharp change in the amplitude of the potential of the superconducting carriers at this boundary. The current–voltage characteristics of the heterostructure studied correspond to tunnel junctions consisting of a superconductor with $d_{x^2-y^2}$ type symmetry of the superconducting wave function and a normal metal. © *1999 American Institute of Physics.* [S1063-7761(99)02012-0]

1. INTRODUCTION

Currently many properties of HTSCs are being estimated using a *d*-type wave function for the superconducting carriers. Specifically, this model explains the magnetic field dependence of the critical current in bimetallic two-junction SQUIDs consisting of $YBa_2Cu_3O_x$ (YBCO) and Pd[1] and the spontaneous excitation of magnetic flux quanta in HTSC structures with three bicrystalline boundaries.[2] At the same time, experiments on electron tunneling in the *c* direction in HTSCs give contradictory results. On the one hand, in HTSC—low-temperature superconductor (*s*-type superconducting wave function) junctions there is no critical current for junctions in the *c* direction,[3–5] which agrees well with the theory of junctions consisting of superconductors with a *d*-type wave function for the superconducting carriers and an *s*-superconductor. On the other hand, an appreciable critical current, whose amplitude varies nonmonotonically as a function of the magnetic and microwave fields nonmonotonically as predicted for junctions with *s*-superconductors, has been observed in a number of experiments.[6–8] To explain the experiments of Refs. 6–8, it has been conjectured that in yttrium-group HTSC materials a mixture of superconducting *s*- and *d*-type carriers arises because of the orthorhombic nature of these materials, and diffuse scattering near the boundary or twinning of HTSC films results in a larger contribution from the *s* component.[9,10] We note that an estimate of the parameters of the Pb/(Au,Ag)/YBCO structures investigated in Refs. 6–8 gives transmittances $D=10^{-7}-10^{-9}$, averaged over the directions of the moments, for the barriers of the HTSC—normal metal barriers with quite large junction areas, $S=0.1-1$ mm^2.

In the present paper we report the results of an experimental investigation of current flow in *s*-superconductor—normal metal—HTSC heterojunctions, fabricated by successive deposition of YBCO, a normal metal (ordinarily Au), and Nb, with much smaller areas ($S\approx 8\times 8\ \mu$m^2) and higher transmittance ($10^{-5}-10^{-6}$) of the YBCO—normal metal boundary. The experimental data are analyzed from two standpoints: on the basis of the isotropic theory of *s* superconductivity and from the standpoint of the modern theory, which assumes a *d*-type wave function in the superconductor YBCO film.

2. EXPERIMENTAL PROCEDURE AND EXPERIMENTAL SAMPLES

The junctions were prepared by using the sequence of operations shown in Fig. 1. First, the epitaxial YBCO films were grown either by laser ablation or using cathodic sputtering in a diode configuration with dc current and high oxygen pressure. During YBCO film growth, a temperature 700–800 °C was maintained and the pure oxygen pressure was 0.3–1 mbar for laser ablation and 3 mbar for cathodic sputtering. Neodymium gallate with (110) orientation or the *r* plane of sapphire with a CeO_2 buffer layer was used as the

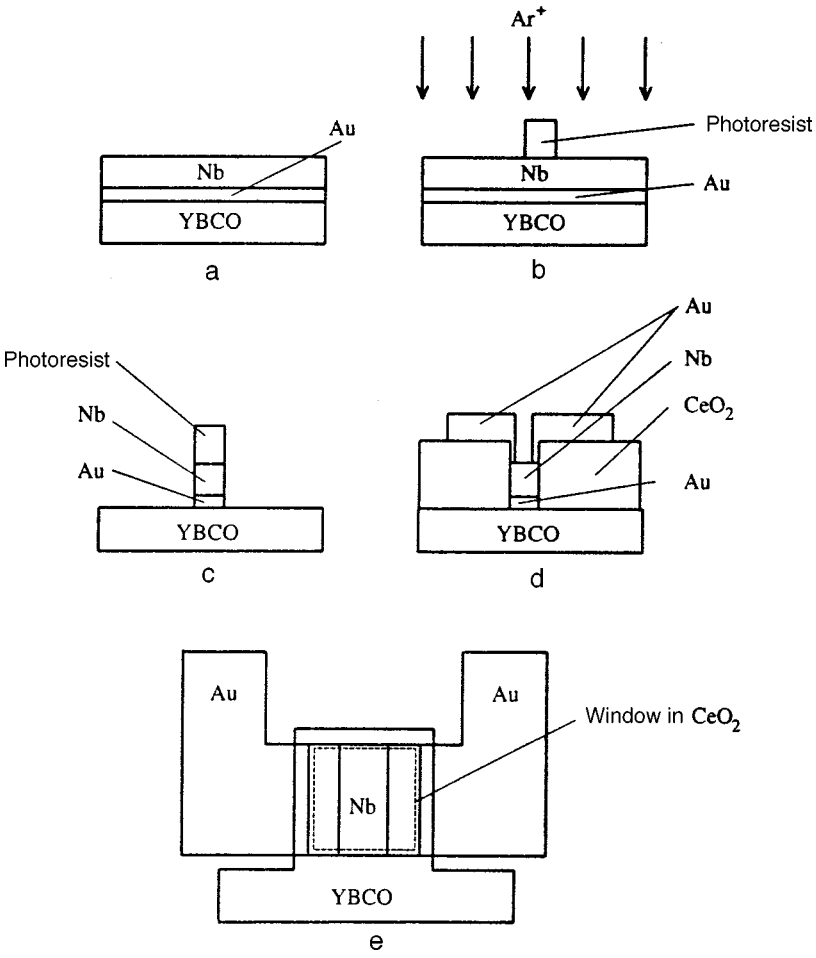

FIG. 1. Sequence for the preparation of HTSC–normal metal–superconductor heterostructures: a) deposition of a trilayer heterostructure Nb/Au/YBCO; b) formation of a region of the heterojunction using ionic etching; c) deposition of the insulator CeO_2 to prevent contacts with YBCO in the basal plane; d) fabrication of an Au electric layout; e) top view of the fabricated heterostructures.

substrate. Epitaxial YBCO films, 100–150 nm thick, with *c* orientation and the following superconducting parameters, measured by the resistive method, were obtained: 1) the critical temperature at which the resistance of the film deposited on a $5\times5\ \mathrm{mm}^2$ substrate is zero, $T_{cf}=84-89\ \mathrm{K}$; 2) the width of the superconducting transition (determined at the levels 0.9 and 0.1 times the resistance of the film at the onset of the transition into the superconducting state), $\Delta T_c=0.5-1\ \mathrm{K}$; 3) the ratio of the resistances at temperatures 300 K and 100 K, $\rho_{300\,\mathrm{K}}/\rho_{100\,\mathrm{K}}\approx2.8$. The number of 0.3–1 μm in diameter particles on the surface of the YBCO film, which are caused by the formation of different phases of YBCO as well as Y, Ba, and Cu oxides, was $\sim10^6\ \mathrm{cm}^{-2}$. Evidence of the high quality of the YBCO films fabricated is the small width of the (005) x-ray peak of YBCO, FWHM(005) $\approx0.2°$, for $\theta/2\theta$ scanning with 0.15 μm film thickness.

A thin, 20 nm thick, layer of normal metal (Au, Ag, Pt) was deposited at 100 °C immediately after the YBCO film, using either laser ablation or high-frequency cathodic sputtering (Fig. 1a). Next, a 100–150 nm thick Nb layer was deposited on a water-cooled substrate by a magnetron cathodic sputtering. The critical temperature of the superconducting transition in Nb films was 9.1–9.2 K. Niobium is used as the low-temperature superconductor because it does not enter into a solid-phase chemical reaction with Au. We note that in the experiments of Refs. 4–7, where Pb is used, a superconducting alloy of Au and Pb can form.

In the trilayer heterostructure obtained, photolithography and ion and plasma-chemical etchings were used to form regions of heterojunctions which during photolithography were fixed on sections with the minimum number of particles on the surface of the YBCO films (Fig. 1b). To prevent electrical contact in the basal (a–b) plane of the YBCO film, the lateral region of the junction was insulated with a CuO_2 layer with a central window with the dimensions $S=8\times8\ \mu\mathrm{m}^2$ (Fig. 1c). At the final stage explosion lithography was used to form junction areas and Au wiring in the form of two stripes, which enable separated input of current and voltage to the top electrode Nb (Figs. 1d, e). The geometry used for the gold contacts (see Fig. 1) makes it possible to investigate the electrophysical properties of Nb/Au/YBCO structures for the YBCO film in the superconducting state. More than 30 Nb/normal metal/YBCO samples, where Au, Ag, and Pt were used as the normal metal, were prepared. In the present paper the results of investigations performed on nine Nb/Au/YBCO samples, in which the variance of the characteristic resistances R_NS (R_N is the differential resistance, measured for $V>20\ \mathrm{mV}$) of the boundaries at liquid-helium temperature did not exceed a factor of 4 (see Table I).

3. EXPERIMENTAL RESULTS

The dependences of the resistances R of the heterojunctions on the temperature T and 4 μm wide the test bridges,

TABLE I. Electrophysical parameters of superconductor structures measured at $T=4.2$ K.

Sample	$R_d(0)$, Ω	R_N, Ω	R_NS, $10^{-6}\,\Omega\cdot\mathrm{cm}^2$	$R_d(0)/R_N$	$\overline{D}$, 10^{-6}
*P*9*J*2	12.2	7.0	4.5	1.7	4.8
*P*9*J*3	9.8	6.0	3.8	1.6	5.6
*P*10*J*2	10.5	5.9	3.8	1.8	5.6
*P*10*J*3	10.6	5.9	3.8	1.80	5.6
*P*11*J*2	4.9	4.2	2.7	1.2	7.9
*P*11*J*3	5.2	3.6	2.3	1.4	9.3
*P*12*J*2	2.4	2.0	1.3	1.2	16.7
*P*13*J*2	7.2	3.5	2.2	2.1	9.5
*P*13*J*3	7.5	6.6	4.2	1.1	5.1

consisting of YBCO films placed on the same substrate, with 1–5 μA bias currents and their current-voltage characteristics (IVCs) in the temperature 4.2–300 K were measured. Figure 2 shows the temperature dependences for one of the substrates. At temperatures $T>T_{cf}$ metallic behavior of $R(T)$ is observed, i.e., the resistance decreases with temperature, as is characteristic for a *c*-oriented YBCO film with current flow in the basal plane of YBCO. As a rule, T_{cf} of the bridges and heterojunctions was less than the critical temperature of YBCO films, measured immediately after the trilayer heterostructure was prepared. The degradation of the superconducting properties of the film is evidently due to a decrease in the amount of oxygen during ionic etching. The inset in Fig. 2 shows the function $R(T)$ for a heterojunction at temperatures $T<T_{cf}$, demonstrating that the resistance of the heterojunction increases as temperature decreases. The value of $R(T)$ at temperatures $T<T_{cf}$ depends on the current. This attests to a nonlinear current dependence of the differential resistance R_d of the heterojunction.

A family of curves of R_d versus the voltage V at various temperatures is shown in Fig. 3. It is evident that $R_d(0)$ increases as T decreases. This growth is reflected in an increase of the resistance $R(T)$ (Fig. 2). The nonlinearity observed in the IVC in the temperature range $72\,\mathrm{K}<T<84\,\mathrm{K}$ is due to the destruction of the superconductivity of the YBCO film. The function $R_d(V)$ increases. This is due to the systematic destruction of superconductivity in sections of the YBCO electrode as the current I increases. We note that the

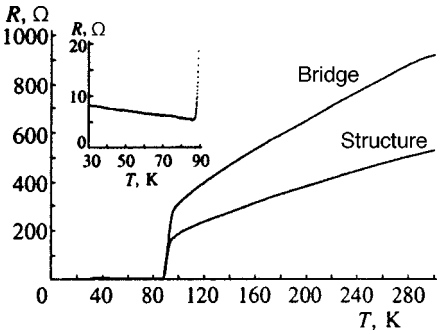

FIG. 2. Temperature dependence of the resistance of the heterostructure and of a 4 μm wide bridge arranged on the same substrate. Inset: $R(T)$, on an enlarged scale, of the heterojunction at temperatures $T<T_c$, where the resistance of the YBCO film is zero.

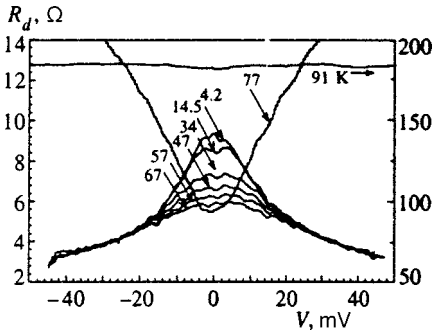

FIG. 3. Family of curves of the differential resistance of the heterojunction at various temperatures versus voltage. The scale along the resistance axis for $R_d(V)$ at $T=91$ K is shown on the right-hand side.

junction resistance at $T\approx T_{cf}$ is somewhat higher than the asymptotic resistance R_N, measured for $V>20$ mV and $T\ll T_{cf}$.

The results of the measurements of the electrophysical parameters of several samples, prepared by the same method are presented in Table I. The resistance R_NS of the boundary at $T=4.2$ K makes it possible to estimate the average (over the direction of the momentum of the quasiparticles) boundary transmittance, which we shall employ below,[11] as

$$\overline{D}=\frac{2\pi^2\hbar^3}{e^2p_F^2}\frac{1}{R_NS}=\frac{2\rho^{\mathrm{YBCO}}l^{\mathrm{YBCO}}}{3R_NS},$$

where p_F is the smallest value of the Fermi momentum for YBCO or Au.[11] The values of the transmittance of the boundaries of the fabricated structures for $\rho^{\mathrm{YBCO}}l^{\mathrm{YBCO}}\approx 3.2\times10^{-11}\,\Omega\cdot\mathrm{cm}^2$ (Ref. 4) are also presented in Table I.

Test samples with bilayer heterostructures Au/YBCO, Nb/YBCO, and Au/Nb, fabricated using a technology with the same conditions as for the formation of the experimental Nb/Au/YBCO heterostructures, were also investigated. The resistances R_NS of these boundaries measured at liquid-nitrogen temperature are $R_NS(\mathrm{Au/YBCO})\sim10^{-8}\,\Omega\cdot\mathrm{cm}^2$, $R_NS(\mathrm{Au/Nb})\sim10^{-12}\,\Omega\cdot\mathrm{cm}^2$, and $R_NS(\mathrm{Nb/YBCO})\sim10^{-4}\,\Omega\cdot\mathrm{cm}^2$. Here the series resistance of the YBCO film for $T_{cf}<77$ K was taken into account. Comparing these quantities with the data presented in Table I, it is evident that the resistance of the Au/Nb boundary can be neglected, and the resistance of the Au/YBCO boundary, which increases when Nd is deposited on top of Au, probably, because of the interaction of Nb with YBCO, makes the main contribution to the resistance of the experimental heterojunctions. The resistance of a direct Nb/YBCO contact is very large. Most likely, the increase in the contact resistance is due to the displacement of oxygen out of the YBCO film into the Nb, which has good gettering characteristics, deposited on top. We note that the oxygen mobility in the $a-b$ planes of YBCO is much higher than in the *c* direction.

Figure 4 shows the surface of a bilayer Au/YBCO heterostructure, measured with an atomic-force microscope. It is evident that its surface consists of Au granules separated by ~1 μm. The subsequently deposited Nb film covers the surface of the Au granules, where a good electric contact with the YBCO film is created, and forms a direct contact with

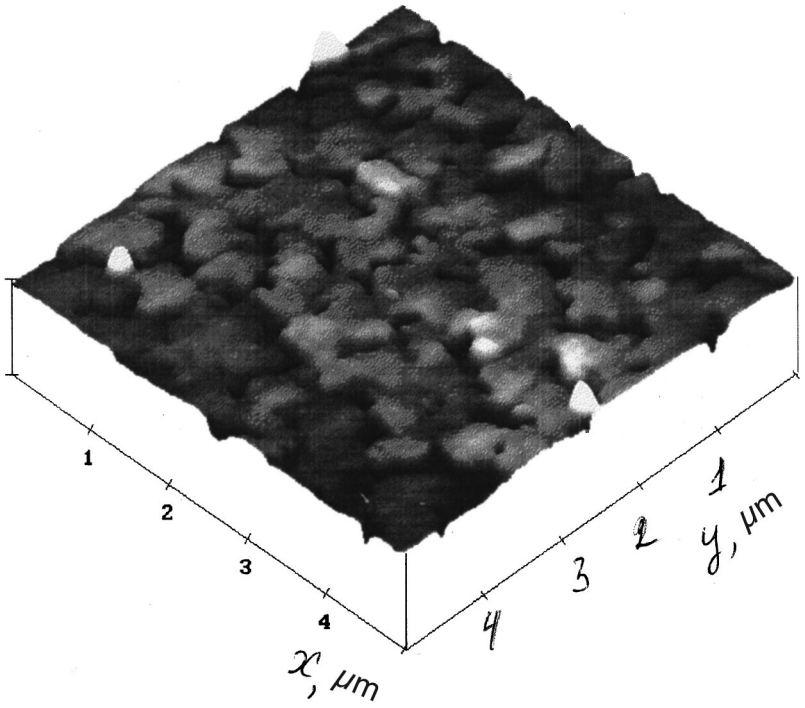

FIG. 4. Three-dimensional image of the surface of a bilayer heterostructure Au/YBCO. The image was obtained with an atomic-force microscope.

YBCO, where, as a result of a decrease in the amount of oxygen in the basal planes, the contact resistance is much higher. This could be the reason why the resistance of the trilayer heterostructure Nb/Au/YBCO is higher than that of a bilayer Au/YBCO heterostructure.

4. DISCUSSION OF THE EXPERIMENTAL RESULTS

The experimental trilayer heterostructure can be represented as Nb/Au/YBCO granules connected in parallel and sections of direct contact of Nb and YBCO via pores in the Au film. Since the characteristic resistance of the Nb/YBCO boundary is several orders of magnitude greater than $R_N S$ of the trilayer heterostructure Nb/Au/YBCO, and the surface area of the granules and pores, according to our estimates, differ severalfold (see Fig. 4), current flows mainly through the boundary of Nb/Au/YBCO granules. A trilayer Nb/Au/YBCO heterostructure can be described by the model shown in Fig. 5; a 100–150 nm superconducting YBCO electrode (S_d) with critical superconducting transition temperature $T_{cf}=87\,\mathrm{K}$; a 1–3 nm YBCO (S_d') layer with an oxygen deficit and therefore disrupted superconducting properties; a 10–20 nm thick layer of normal metal (Au); a 100–150 nm thick superconducting Nb (S_s) electrode with $T_c=9.2\,\mathrm{K}$. A similar model has been proposed in Ref. 4 to estimate the electrophysical parameters of the system Pb/Au/YBCO.

First, we shall estimate the change in the superconducting order parameter in Nb as a result of the contact with Au. Since the measured value of the boundary resistance is quite small, it can be assumed that the superconducting Green's function, characterizing the amplitude of the interaction potential Φ of the superconducting carriers and its derivative with respect to the coordinate x are continuous at the boundary. Using the calculations of Refs. 12 and 13, we find that the superconducting order parameter Δ_1 of Nb at the Nb/Au boundary is somewhat less than its equilibrium value Δ_{Nb} in the interior volume of the film and is $\Delta_1/e\approx 560\,\mu\mathrm{V}$. For estimates, the following values of the electrophysical parameters of Nb and Au were used at $T=4.2\,\mathrm{K}$: $\rho^{\mathrm{Nb}}l^{\mathrm{Nb}}=4\times10^{-12}\,\Omega\cdot\mathrm{cm}^2$, $\xi^{\mathrm{Nb}}=0.73\times10^{-6}\,\mathrm{cm}$, $v_F^{\mathrm{Nb}}=3\times10^7\,\mathrm{cm/s}$, $T_{c0}^{\mathrm{Nb}}=9.2\,\mathrm{K}$ and $\rho^{\mathrm{Au}}l^{\mathrm{Au}}=8\times10^{-12}\,\Omega\cdot\mathrm{cm}^2$, $\xi^{\mathrm{Au}}=10^{-6}\,\mathrm{cm}$, and $v_F^{\mathrm{Au}}=1.4\times10^8\,\mathrm{cm/s}$, where $v_F^{\mathrm{Nb,Au}}$ is the Fermi velocity and $l^{\mathrm{Nb,Au}}$ is the mean-free path length in Nb and Au, respectively.

Let us estimate the change in the order parameter at the YBCO/Au boundary. We assume that as a result of the interaction of YBCO and Nb, a superconducting surface layer S_d' of the order of 3 nm thick with critical temperature less than 4 K is formed.[4] Assuming that the coherence length of S_d' differs negligibly from $\xi_{c-\mathrm{YBCO}}$ and is $\xi_{S_d'}=5\times10^{-8}\,\mathrm{cm}$ and that the resistivity increases by an order of magnitude[4]—from $\rho_{c-\mathrm{YBCO}}=10^{-4}\,\Omega\cdot\mathrm{cm}^2$ to $\rho_{S_d'}=1\times10^{-3}\,\Omega\cdot\mathrm{cm}^2$, we obtain that at the Au/YBCO boundary

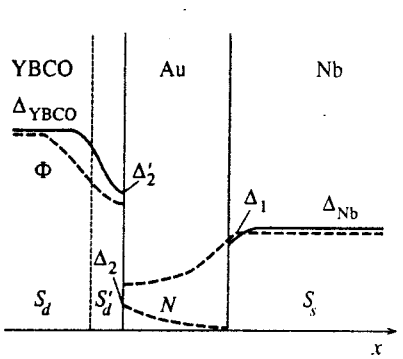

FIG. 5. Schematic diagram of the distribution of the order parameter (solid line) and the amplitude of the pair potential (dashed lines) in a direction perpendicular to the surface of an Nb/Au/YBCO heterostructure.

the order parameter decreases on the YBCO side decreases by a factor of approximately 100, $\Delta_2'/e \approx 140\ \mu$V. A potential barrier with low transmittance, $D \sim 10^{-6}$, is present at the Au/YBCO boundary. This barrier decreases Δ_2 by another factor of D, $\Delta_2 = \Delta_2' D$. Here we used the theoretical estimates, which are strictly applicable for superconductors with *s*-type pairing. However, as the calculations of Refs. 10 and 14 show, the character of the change in the order parameter at the boundary of a *d* superconductor with a normal metal or insulator does not differ much from a junction with an *s* superconductor for orientations of the normal to the *d* superconductor along the principal crystallographic axes.

As a result, we can estimate the amplitude of the superconducting current through the entire structure by using the model of a superconductor—normal metal—superconductor ($S_d'NS$) junction, on the boundaries of whose weak section the values of the order parameters are known: $\Delta_2/e \approx 0.004\ \mu$V and $\Delta_1/e \approx 560\ \mu$V. In what follows, we shall employ the theory developed for $S-N-S$ junctions. The thickness of the N layer is of the order of the coherence length, so that the change in the superconducting order parameter in the interlayer can be neglected. As a result, the product of the critical current I_c by R_N at low temperature is $I_c R_N \approx \sqrt{(\Delta_1 \Delta_2)}/e = 0.09\ \mu$V. Taking account of the resistance of the heterojunctions ($R_N = 10\ \Omega$), we obtain that the critical current of the structure $I_c \approx 0.009\ \mu$A is less than the fluctuation current $I_f = 1\ \mu$A of the measuring system and does not affect the experiment even if YBCO contains a mixture of *d* and *s* components of the superconducting order parameter and the number of *s* components is greater than the number of *d* components. For pure *d* pairing, the superconducting current for flow along the *c* direction in YBCO must be zero because of the type of symmetry of the superconducting order parameter. To estimate the critical current we assumed that the large width of the potential barrier (several coherence lengths) prevents direct tunneling of the superconducting current through the barrier. We note that we have considered quite strong suppression of the order parameter at the YBCO boundary because of degradation of the superconducting parameters of the HTSC film. However, even in the absence of suppression of the order parameter in the surface layer of YBCO, $\Delta_1/e = 14$ mV, the critical current of the heterojunctions Nb/Au/YBCO will once again be comparable to the fluctuation current because of the decrease in the order parameter on the low-transmittance barrier.

The finite critical current, observed in a number of works,[6–8] in Pb(Au,Ag)/YBCO heterostructures with a much larger value of $R_N S$ and large junction areas could be due to the fact that treatment of the YBCO electrode with a solution of bromine and alcohol, as was done in these works, opens up the basal planes of YBCO, the transmittance of whose boundaries with a normal metal or ordinary superconductors is three orders of magnitude higher than in the *c* direction ($R_{ab} S_{ab} \ll R_c S_c$). Ultimately, the superconducting current flows along the contacts to the basal plane of YBCO, and the normal resistance is determined by parallel connection of the resistances of the boundaries along the *c* and in the basal plane. In our case current flow is impeded in the direction of the basal plane along the Nb/YBCO junctions, most likely because of substantial displacement of oxygen out of YBCO into Nb.

It is important that lead can react with Au, forming a superconducting alloy. Then, the Pb/Au/YBCO structure contains a superconductor instead of a layer of normal metal. This is confirmed by the appearance of gap features of lead in the IVCs[6–8] at sufficiently low temperatures ($T = 1.2$ K).

A new explanation of the experimental data on the flow of a superconducting current through low-temperature superconductor—HTSC junctions was proposed recently. It has been shown theoretically[15] that a strong spin-orbit interaction, which is observed in Pb/Ag structures, intensifies superconducting current flow through a barrier. Replacement of Pb by an Al- or Nb-type superconductor decreases the spin-orbit interaction, and the superconducting current decreases as a result.

Let us discuss the dependences $R_d(V)$ for heterojunctions as a function of temperature in the range 4.2–100 K (Fig. 3). For $T \ll T_c$ the IVC as a whole corresponds to heterojunctions of the type superconductor—insulator—normal metal ($S-I-N$): There is a location where R_d increases at low voltages. However, the feature on $R_d(V)$ that is due to the gap in YBCO is not observed in the experiment. This corresponds to a junction with a superconductor with gapless superconductivity, including with *d*-type superconductivity.[14,16,17] According to the calculations performed in Ref. 14, the feature at $eV \approx \Delta$ in the density of states of a *d* superconductor gives a logarithmic dependence $R_d \propto \ln(T)$, $\ln(|eV| - |\Delta|)$, subjected to strong temperature broadening just as for a gapless *s* superconductor. We note that for *s* superconductors with a gap a power-law divergence is observed $R_d \propto T^{-1/2}\ ((eV)^2 - \Delta^2)^{-1/2}$. The features in the form of changes in $R_d(V)$ at voltages $V < 2$ mV due to the niobium gap have virtually no effect in our experiment, and we did not study them in detail.

For *s*-type symmetry of the order parameter in a superconductor at low temperatures, $kT \ll \Delta$, the number of excited quasiparticles decreases exponentially with temperature. Therefore the resistance increases proportionally $R_d(0) \propto (-\Delta/T)$.[17] In a superconductor with *d*-type pairing, the presence of nodes with a zero order parameter makes it possible to excite a quasiparticle even at very low temperature, $T \ll \Delta$. As a result, $R_d(0)$ grows more slowly as temperature decreases.[14] As one can see in the inset in Fig. 2, nearly linear growth of $R_d(0)$ with decreasing T is observed in the experiment. The dependence $R_d(V)$ is quadratic as $V \to 0$, which agrees qualitatively with calculations for a *d* superconductor.[14]

One of the most surprising features of superconductors with *d*-type pairing is the appearance of two types of bound states, which, as a rule, are not observed in *s*-superconductors.[17] Surface states with low energies at the boundary of the *d* superconductor with an insulator are due to the change in sign of the order parameter at the Fermi surface for quasiparticles reflected from the boundary.[16,17] The superconducting parameter for a *d*-type superconducting wave function changes sign with a 90° circuit around the *c* axis. Since the direction of the momentum of a quasiparticle

changes on mirror reflection from a boundary, bound states arise at zero energies because of Andreev reflection. This leads to the appearance of a dip in $R_d(V)$ at small V, as is observed experimentally for a transport current in the [110] direction in a YBCO film (see Refs. 6–8, 18, and 19). In our case, the contribution of such quasiparticles is small because the normal to the boundary is oriented along one of the principal crystallographic directions in YBCO. For mirror-reflected quasiparticles, there is no Andreev reflection because the phases of the order parameter are the same for incident and reflected quasiparticles.

An additional mechanism was recently predicted theoretically for the appearance of bound states due to the suppression of the order parameter of a d-superconductor for orientations of the normal with respect to the boundary different from the principal crystallographic axes or for diffuse reflection at a boundary with an insulator.[17] These states are observed at energies different from zero, and estimates in Ref. 17 show that they are more stable with respect to the quality of the boundary. The appearance of bound states should be observed in the dependences $R_d(V)$ as a decrease of R_d for eV_r of the order of the gap in the d-superconductor, and in addition the ratio eV_r/Δ depends on the angle between the normal and the crystallographic axes of the d-superconductor. The condition for the existence of bound states with nonzero energy is suppression of the order parameter near the boundary. This occurs in our experiment because of the degradation of the superconducting properties of the surface. Indeed, in all samples we observe features at $V_r = 15\,\mathrm{mV}$, where V_r is virtually temperature-independent.

5. CONCLUSIONS

In the present work, heterojunctions with dimensions of several microns, obtained by successive deposition of YBCO, Au, and Nb, with transport current flowing in YBCO along the c axis were fabricated and studied experimentally. The transmittances of the heterostructures, as estimated from the resistance of the junctions, are two orders of magnitude greater than the existing experimental data, and the areas of the heterojunctions are much smaller. The IVCs of the heterojunctions with resistances differing from one another by a factor of 4 were investigated. Estimates based on the proximity effect showed that the absence of a critical current in heterojunctions is probably due to a decrease in the amplitude of the potential of the superconducting carriers at the Au/YBCO boundary. The curves of the differential resistance of the heterojunctions versus the voltage are similar to the case of $S-I-N$ junctions with a gapless superconductor, specifically, the absence of a YBCO gap feature could also correspond to d-type superconductivity, specifically, to the presence of nodes of the order parameter as the direction of the momentum of the quasiparticles changes by 45°. The dependence of $R_d(0)$ on T also corresponds to a d-type superconductor.

We thank Yu. S. Barash, D. A. Golubev, A. V. Zaĭtsev, Z. G. Ivanov, and M. Yu. Kupriyanov for a helpful discussion of the experimental results, and D. Ertz, P. B. Mozhaev, and T. Henning for assisting in the experiment.

This work was supported in part by the program ''Current Problems of Condensed-State Physics'' (subsection ''Superconductivity''), the Russian Fund for Fundamental Research, and the INTAS program of the European Union.

*)E-mail: gena@lab235.cplire.ru

[1] D. A. Wollman, D. J. Van Harlingen, W. C. Lee *et al.*, Phys. Rev. Lett. **71**, 2134 (1993).
[2] C. C. Tsuei, J. R. Kirtley, C. C. Chi *et al.*, Phys. Rev. Lett. **73**, 593 (1994).
[3] H. Akoh, C. Camerlingo, and S. Takada, Appl. Phys. Lett. **56**, 1487 (1990).
[4] J. Yoshida, T. Hashimoto, S. Inoue *et al.*, Jpn. J. Appl. Phys., Part 1 **31**, 1771 (1992).
[5] J. Lesueur, L. H. Greene, W. L. Feldmann *et al.*, Physica C **191**, 325 (1992).
[6] A. G. Sun, A. Truscott, A. S. Katz *et al.*, Phys. Rev. B **54**, 6734 (1996).
[7] A. S. Katz, A. G. Sun, R. C. Dynes *et al.*, Appl. Phys. Lett. **66**, 105 (1995).
[8] J. Lesueur, M. Aprili, A. Goulon *et al.*, Phys. Rev. B **55**, 3398 (1997).
[9] J. R. Kirtley, K. A. Moler, and D. J. Scarlapino, E-print archive cond-mat/9703067 (1997).
[10] L. J. Buchholtz, M. Palumbo, D. Rainer, and J. A. Sauls, J. Low Temp. Phys. **101**, 1099 (1995).
[11] A. V. Zaĭtsev, Zh. Éksp. Teor. Fiz. **86**, 1742 (1984) [Sov. Phys. JETP **59**, 1015 (1984)].
[12] M. Yu. Kupryanov and K. K. Likharev, IEEE Trans. Magn. **27**, 2400 (1991).
[13] G. Deutscher, Physica C **185–189**, 216 (1991).
[14] Yu. S. Barash, A. V. Galaktionov, and A. D. Zaikin, Phys. Rev. B **52**, 665 (1995).
[15] G. S. Lee, Physica C **292**, 171 (1997).
[16] Y. Tanaka and S. Kashiwaya, Phys. Rev. B **53**, 11957 (1996).
[17] Yu. S. Barash, A. A. Svidzinsky, and H. Burkhardt, Phys. Rev. B **55**, 15282 (1997).
[18] F. V. Komissinskiĭ, G. A. Ovsyannikov, N. A. Tulina, and V. V. Ryazanov, in *Abstracts of Reports at the 31st Conference on Low-Temperature Physics (LT-31)*, Moscow, 1998, p. 236.
[19] P. V. Komissinski, G. A. Ovsyannikov, N. A. Tulina, and V. V. Ryazanov, E-print archive cond-mat/9903065 (1999).

Translated by M. E. Alferieff

Effective conductivity of three-phase nondissipative randomly-inhomogeneous media

V. E. Arkhincheev*)

Buryat Science Center, Siberian Branch of the Russian Academy of Sciences, 670047 Ulan-Udé, Russia
(Submitted 10 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 2150–1253 (December 1999)

An inhomogeneous medium consisting of a random mixture of three nondissipative (Hall) phases is studied. An exact expression is obtained for the effective conductivity of such a medium with arbitrary concentrations of the phases and the additional condition that the concentrations of two phases are the same. © *1999 American Institute of Physics.* [S1063-7761(99)02112-5]

1. As is well known, randomly inhomogeneous media have been studied mostly in the two-dimensional case. This is due to the additional symmetry of the two-dimensional equations for the constant current and Ohm's law

$$\operatorname{div}\mathbf{j}=0, \quad \operatorname{curl}\mathbf{e}=0, \tag{1}$$

$$\mathbf{j}=\tilde{\sigma}\mathbf{e}, \tag{2}$$

and their invariance relative to rotational transformations:

$$\mathbf{j}=a\mathbf{j}'+ib\mathbf{e}', \quad \mathbf{e}=c\mathbf{e}'+id\mathbf{j}'. \tag{3}$$

Here the vectors $\mathbf{j}$ and $\mathbf{e}$ are the electric current and field, $\tilde{\sigma}=\sigma/(1+i\beta)$ is the conductivity tensor of the medium in a magnetic field, σ is the conductivity, $\beta=\mu B$ is the Hall factor, μ is the particle mobility, and the coefficients a,b,c, and d are real. The imaginary unity i describes rotation by $\pi/2$ in the complex plane. The magnetic field is directed perpendicular to the plane. Using the transformations (3), we obtain an expression for the conductivity tensor of the primed system:

$$\tilde{\sigma}'=\frac{b+ia\tilde{\sigma}}{d\tilde{\sigma}+ia}. \tag{4}$$

A similar expression is also obtained for the effective conductivity tensor of the primed system. For two-phase randomly-inhomogeneous media and equal concentrations, an expression for the effective conductivity tensor follows from Eq. (4).[1] We note that this expression is obtained for equal concentrations of the phases (at one concentration point).

In the present paper the effective conductivity of three-phase nondissipative (Hall) randomly-inhomogeneous media is calculated and a qualitative description of the results obtained is given. The possibility of solving the problem exactly is based on the representation of the Dykhne transformations in the form of linear-fractional conformal transformations of the conduction plane of the initial medium onto the conduction plane of the transformed, primed system.[2] According to the general theory of conformal transformations, they are given by three points in the plane and their images. This is why it is possible to transform a three-phase two-dimensional medium into a similar medium. However, according to Ref. 1, this possibility is realized only for one value of the conductivity of the third phase: $\sigma=\sqrt{\sigma_1\sigma_2}$. We shall show that this limitation becomes immaterial for current flow under quantum Hall effect conditions ($\sigma_{xx}=0$, $\sigma_{xy}=\text{const}$) or, in other words, for current flow along nondissipative phases. Thus, a three-phase Hall medium can be transformed for any values of the conductivities and arbitrary concentrations of the phases provided that the concentrations of two phases are the same, but they can be arbitrary in magnitude, just as the concentration of the third phase (the sum of the three concentrations is 1). This makes it possible to solve the problem of the effective conductivity of three-phase randomly inhomogeneous media in the entire concentration interval.

2. To obtain an exact expression for the effective conductivity of nondissipative three-phase randomly-inhomogeneous media, we return once again to the expression relating the conductivity of the primed and initial systems [Eq. (4)]. We shall consider this as a conformal mapping of one plane into another:[2]

$$W=L(z)=\frac{b+iaZ}{dZ+ic}. \tag{5}$$

Let us rewrite the expression (4) in the form

$$d\sigma'\sigma-b=i(a\sigma-c\sigma). \tag{6}$$

Let us interchange the phases 1 and 2: $\sigma_1'=\sigma_2$, $\sigma_2'=\sigma_1$. This is possible if the concentrations X_i of the phases are the same:

$$X_1=X_2. \tag{7}$$

The condition (7) is important and will be used below to obtain the macroscopic equivalence of the initial and primed systems. Then, from Eq. (6) we obtain two relations between the coefficients a,b,c, and d:

$$a=-c, \quad d\sigma_1\sigma_2-b=ia(\sigma_1-\sigma_2). \tag{8}$$

If the conductivities are real, which is the case considered in Ref. 1, the coefficient a in the relation (5) must be set equal to zero. Otherwise, we obtain complex expressions for the coefficients a,b,c, and d which are real by definition. In this case, there arises the above-indicated restriction on the conductivity of the third phase, because of the fact that there are not enough parameters in the problem. However, if the conductivities of the phases are assumed to be purely imaginary

— capacitive or inductive, $\sigma = i\omega L$, $\sigma = (i\omega C)^{-1}$, or purely Hall, $\sigma = i\sigma_{xy}$ (in the quantum Hall effect regime $\sigma_{xx} = 0$, $\sigma_{xy} =$ const), and the coefficient a can be assumed to be nonzero, the coefficients b and d will also be real. Then the condition for the conductivity of the third phase will be simply another condition for determining all coefficients of the transformations (3):

$$d\sigma_3\sigma_3' - b = ia(\sigma_3 + \sigma_3'). \tag{9}$$

Let $\sigma_3' = \sigma_3$. Then the primed system with the additional condition (6) is macroscopically equivalent to the initial system: $\sigma_e' = \sigma_e$. From Eq. (6) we also obtain an expression for the effective conductivity of a three-phase randomly-inhomogeneous medium in the entire concentration range:

$$\sigma_e = i\left\{\frac{a}{d} \pm \left[\left(\frac{b}{c} + \frac{a}{d}\right)\frac{a}{d}\right]^{1/2}\right\}, \tag{10}$$

where the coefficients $a, b, c,$ and d were determined above by the conditions (8) and (9).

3. Let us discuss the result obtained. In the derivation of Eq. (10), no restrictions on the concentrations of the phases, except for the condition (6), were used. Therefore the result obtained is valid for any concentrations of the phases, specifically, when the concentrations of two phases are the same and arbitrary in magnitude (the sum of the concentrations of all phases is 1).

Let us consider the limiting cases following from Eq. (10) and elucidating the meaning of this equation. For $\sigma_3 = 0$, according to Eq. (8), the coefficient b vanishes and we obtain for the effective conductivity the two solutions

$$\sigma_e^{(1)} = 0, \quad \sigma_e^{(2)} = 2i\frac{a}{d} = i\sigma_{xy}^{(e)} = i\frac{2\sigma_{xy}^{(1)}\sigma_{xy}^{(2)}}{\sigma_{xy}^{(1)} + \sigma_{xy}^{(2)}}. \tag{11}$$

Let us clarify the results obtained. For concentrations of the first and second conducting phases such that $X_1 + X_2 \leqslant 1/2$, it is impossible to form an infinite cluster — an ensemble of conducting paths going to infinity, so that the effective conductivity of the medium is zero. This corresponds to the first zero solution. For concentrations of the first and second phases greater than the critical value, the effective conductivity changes abruptly and remains constant as the concentration of the dielectric phase decreases. This means that as the concentration of the dielectric varies, the electric field in the conducting Hall phases increases in a manner so that the total current in the system remains constant and therefore the effective Hall conductivity also remains constant.

For $\sigma_3 = \infty$ the coefficient d vanishes: $d = 0$. The effective conductivity also assumes two values:

$$\sigma_e^{(1)} = \infty, \quad \sigma_e^{(2)} = i\frac{b}{2c} = i\sigma_{xy}^{(e)} = i(\sigma_{xy}^{(1)} + \sigma_{xy}^{(2)})/2. \tag{12}$$

The meaning of the results is also obvious. If the sum $X_1 + X_2$ of the concentrations is less than the critical value, the effective conductivity of the medium is determined by the superconducting phase, shunting the metallic phases. If the sum of the concentrations is greater than the critical value, the effective conductivity is determined by the resistance of the metallic phases. We also note that in the cases considered above, the effective conductivity reaches its limiting values: the minimum and maximum possible values.

The situation is similar if the conductivity of the third phase is finite. If the sum of the concentrations of the first and second phases is below a threshold value, where it is impossible to construct an infinite cluster without the third phase, the effective conductivity of the medium has one value. As the percolation threshold with respect to the common concentration of the first and second phases is crossed, the effective conductivity changes abruptly and the conductivity of the third phase likewise participates in determining the total effective conductivity by a different, completely determined, way. We underscore once again that the effective conductivity is independent of the concentrations of the phases in a wide range of concentrations.

For a random mixture of Hall and metallic phases, it has been shown in Ref. 3 that in a wide range of concentrations, specifically, as long as flow along the Hall phase occurs, the effective characteristics of such a system are constant and equal to the corresponding values for a Hall medium. This corresponds to current flow in the system described with the minimum (zero) heating. Unfortunately, similar arguments are not applicable for the system studied, where all phases are nondissipative.

This work was supported by the Russian Fund for Fundamental Research (Grant 99-02-17355).

*)E-mail: varkhin@bsc.buryatia.ru

[1] A. M. Dykhne, Zh. Éksp. Teor. Fiz. **59**, 110, 641 (1970) [Sov. Phys. JETP **32**, 63, 348 (1971)].
[2] V. E. Arkhincheev, JETP Lett. **67**, 1004 (1998).
[3] V. E. Arkhincheev and E. G. Batiyev, Solid State Commun. **12**, 1059 (1989).

Translated by M. E. Alferieff

Dynamics of a Peierls system in a light field

A. L. Semenov*)

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia
(Submitted 23 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 2154–2175 (December 1999)

Equations describing the temporal dynamics of the order parameter $\xi(t)$ of a metal–semiconductor phase transition and the density $n(t)$ of electron–hole pairs in a Peierls system in a light field are obtained on the basis of the Lagrange equation for the phonon mode and the Liouville equation for the density matrix of the electronic subsystem. The equations obtained are analyzed for a stationary state (with adiabatically slow variation of the light intensity I) and for a transient process near the initial and final states of dynamic equilibrium (with the light field switched on abruptly). It is shown that for adiabatically slow growth of the intensity I up to a certain critical value I_c the band gap of the electronic spectrum decreases but the semiconductor phase of the Peierls system remains stable. For $I>I_c$ the stationary semiconductor state ($\xi\neq 0$) becomes unstable. When the light is switched on abruptly, the deviation of the system parameters from the initial values is described by an exponential law with a characteristic reciprocal of the rise time of the process linearly dependent on the irradiation intensity I. As a new position of equilibrium is approached, three qualitatively different regimes of behavior of the order parameter ξ and density n are possible. For low intensities I ($I<I_1$) a purely relaxational aperiodic process occurs. For intermediate intensities I ($I_1<I<I_c$) damped oscillations of ξ and n are observed near a new stationary semiconductor state with a smaller band gap. For $I>I_c$ the stationary semiconductor state with $\xi\neq 0$ is absent. The experimental data on the irradiation of a vanadium dioxide film with a powerful laser pulse is interpreted on the basis of the theory developed.
[S1063-7761(99)02212-X]

1. INTRODUCTION

It is well known that as temperature T decreases below a certain critical value T_0, a chain of equidistant atoms, each of which contains a single external electron, undergoes a reversible thermodynamically equilibrium phase transition characterized by a change in the crystal structure (pairwise convergence of atoms in the chain) and the formation of a band gap in the electronic spectrum at the Fermi level.[1] This transition, for which, in addition, a uniform deformation of the atomic chain occurs,[2] is called a Peierls transition, and the system in which the transition occurs is called a Peierls system.

A thermodynamically equilibrium metal–semiconductor (or semiconductor–metal) phase transition in a Peierls system can also be initiated by pressure (uniaxial or hydrostatic),[3,4] breakdown of the ideality of the crystal lattice (specifically, as a result of doping with substitution impurities),[1,4–8] a constant electric field,[9] adsorption of molecules from the gas phase, and so on.[4,10,11] Near the critical point T_0 of the thermodynamically equilibrium metal–semiconductor phase transition, a one-dimensional spatial, periodic, heterophase structure of alternating metallic and semiconductor phases forms in a film consisting of a Peierls material on a substrate.[12]

There is great interest in theoretical and experimental investigation of a thermodynamically nonequilibrium phase transition in a low-temperature semiconductor phase of a Peierls system, each atom of which contains n ($0<|n-1|<1$) external electrons, in a constant electric field directed along the principal axis of the quasi-one-dimensional crystal.[13–18] This transition, manifested as a sharp increase of the electrical conductivity when the intensity of the electric field is greater than a certain threshold value, is due to the appearance of a charge-density wave and an initially static Fröhlich phonon mode, associated with the charge-density wave and formed by the displacement of atoms accompanying the Peierls metal–semiconductor phase transition, moving along the atomic chain.

Different types of thermodynamically nonequilibrium, photostimulated instabilities and phase transitions in a Peierls system have been studied in Refs. 19–25. Specifically, it was shown[19–22] that when a Peierls semiconductor is irradiated with light with frequency greater than the band gap E_g of the electronic spectrum, the density n of nonequilibrium electron–hole pairs increases and, in consequence, the band gap decreases smoothly. When the density n reaches a certain critical value n_c, the band gap E_g abruptly vanishes (photoinduced semiconductor–metal phase transition).

On account of the characteristic features of the electronic spectrum of a Peierls semiconductor (the presence of van Hove singular points,[23] sharp[24] or smeared[25] band-gap edges) in a light field with a specially selected central frequency of the optical spectrum, sharp photoinduced transitions from one semiconductor state into another are also possible.[23–25] Optical bistability with increasing absorption without a reso-

nator is observed near the critical point of these transitions in a Peierls system.[26]

The temporal dynamics of the development of a photoinduced phase transition, consisting of a change in the structure of the crystal lattice and the band gap of the electronic spectrum of the semiconductor when the semiconductor is irradiated with light, has been investigated for various materials in an entire series of works.[27–34] It has been shown that for high nonequilibrium carrier densities temporal[27–29] and spatial[30] periodic oscillations of the parameters of the system can arise. Depending on the specific conditions, the new state forming can be a metal,[27] a semiconductor with a different crystal modification,[31] an amorphous solid,[31] a heterophase structure,[19,21,30,32] and so on.[33,34]

In the present paper the dynamics of a photoinduced phase transition in a Peierls system is investigated under the assumption that the characteristic phonon and interband electronic relaxation times are constant. In contrast to Refs. 27–34, where the density of nonequilibrium electrons in the conduction band was as an externally controllable parameter and the method for producing this density either was not considered or was considered at a qualitative level, in the present paper the mechanism of the interaction of the electronic subsystem with the electromagnetic field is specified and the possibility of the nonequilibrium carrier density changing in the course of the phase transition is taken into account. The light intensity is chosen as an externally controllable parameter. This is more justified from the physical standpoint. The light field is treated as a quasimonochromatic stationary random process, in which photostimulated generation of nonequilibrium electron–hole pairs occurs as a result of the electric-dipole interaction of the photons with the electronic subsystem of the semiconductor. An increase of this density via the electron–phonon interaction gives rise to readjustment of the crystal structure and the electronic spectrum of the Peierls system.

In the present paper the expressions describing the time-dependence of the band gap in the course of a phase transition at the initial stage of evolution during the development of an instability of the initial, stable (in the absence of a light field) phase and at the final stage near the new stationary state of dynamic equilibrium are obtained. An interpretation of the experimental data of Ref. 35 on a photostimulated semiconductor–metal phase transition in a vanadium dioxide film irradiated with powerful laser radiation is given on the basis of the theory developed.

2. HAMILTONIAN AND ELECTRONIC SPECTRUM OF THE SYSTEM

Let us consider a chain of atoms with each atom containing a single external electron. We write the Hamiltonian of the electronic subsystem in the form

$$H_e=\sum_{n,m(m>0)} B_{n,n+m}(a_n^+ a_{n+m}+a_{n+m}^+ a_n), \tag{2.1}$$

where n is the number of the atom in the chain, $B_{n,n+m}$ is the overlap integral between the wave functions of the nth and $n+m$-st atoms, and a_n^+ and a_n are operators creating and annihilating an electron on the nth atom.

For narrow-gap systems (specifically, for the Peierls model considered here) the distance $r_{n,n+1}$ between the nearest-neighbor atoms is several times greater than the effective radius R of the atomic wave function of an electron. In this case the overlap integral $B_{n,n+m}$ is determined approximately by the relation[36]

$$B_{n,n+m}\sim\exp(-r_{n,n+m}/R). \tag{2.2}$$

We write the coordinate x_n of the an nth atom in a chain with pairwise convergence of the atoms in the form

$$x_n=nr_0+\frac{\cos(\pi n)R\xi}{2}, \tag{2.3}$$

where r_0 is the interatomic distance in the metallic phase and ξ is the period-doubling parameter of a one-dimensional crystal (the order parameter of a metal–semiconductor phase transition). Then we obtain for the distance $r_{n,n+m}$

$$r_{n,n+m}=mr_0+(-1)^n R\xi\frac{((-1)^m-1)}{2}. \tag{2.4}$$

Taking account of Eq. (2.4), the overlap integral $B_{n,n+m}$ (2.2) becomes

$$B_{n,n+m}=b\exp\left(-\chi(m-1)+\frac{(-1)^n(1-(-1)^m)}{2}\xi\right), \tag{2.5}$$

where $\chi=r_0/R$ is a dimensionless parameter characterizing the relative separation of the nearest-neighbor atoms, and b is the overlap integral of the wave functions of the nearest-neighbor atoms in the metallic phase ($\xi=0$). The phases of the atomic wave functions in the form (2.1) are chosen so that b in Eq. (2.5) is a real quantity.

To diagonalize the Hamiltonian (2.1) we employ Bogolyubov's method of canonical transformations.[37] We switch to collective Fermi second-quantization operators c_k and c_k^+ according to the formula

$$a_n=\frac{1}{\sqrt{N}}\sum_k c_k e^{ikn}, \tag{2.6}$$

where N is the number of atoms in the chain, $k=0, \pm 2\pi/N, \ldots, \pm\pi$, $c_{k+2\pi}=c_k$. In the new operator representation the Hamiltonian (2.1) becomes

$$H_e=\sum_k b(Q_k c_k^+ c_k+iR_k c_k^+ c_{k-\pi}), \tag{2.7}$$

where

$$Q_k=\frac{\cosh\xi\cos k(\exp(2\chi)-1)+\cos(2k)\exp\chi-\exp(-\chi)}{\cosh(2\chi)-\cos(2k)}, \tag{2.8}$$

$$R_k=-\frac{\sinh\xi\sin k(\exp(2\chi)-1)}{\cosh(2\chi)-\cos(2k)}. \tag{2.9}$$

Let us perform in Eq. (27) another canonical transformation to Fermi operators α_k, α_k^+ in accordance with the formula

$$c_k=\frac{\alpha_k+i\varphi_k\alpha_{k-\pi}}{\sqrt{1+\varphi_k^2}}. \tag{2.10}$$

The function φ_k in Eq. (2.10) is chosen so that the Hamiltonian obtained is diagonal in the new variables α_k, α_k^+:

$$H_e=\sum_k \varepsilon_k\alpha_k^+\alpha_k. \tag{2.11}$$

Substituting the expression (2.10) into Eq. (2.7) and equating to zero the off-diagonal elements, we find φ_k and the dispersion law ε_k:

$$\varphi_k=\frac{Q_k-Q_{k-\pi}-\mathrm{sign}(Q_k-Q_{k-\pi})\sqrt{(Q_k-Q_{k-\pi})^2+4R_k^2}}{2R_k}, \tag{2.12}$$

$$\varepsilon_k=\frac{b}{2}(Q_k+Q_{k-\pi}+\mathrm{sign}(Q_k-Q_{k-\pi})\sqrt{(Q_k-Q_{k-\pi})^2+4R_k^2}). \tag{2.13}$$

It is evident from the relations (2.13), (2.8), and (2.9) that the spectrum ε_k with $\xi\neq 0$ has two bands, the lower band being completely filled in the ground state and the upper band being empty (semiconductor phase). For $\xi=0$ the spectrum (2.13), (2.8), and (2.9) consists of a single half-filled band (metallic phase).

3. DIPOLE MOMENT OPERATOR

The dipole moment operator of a Peierls system is determined by the relation

$$\mathbf{d}=\sum_{n,m(m>0)}(\mathbf{d}_{n,n+m}a_n^+a_{n+m}+\mathbf{d}_{n,n+m}^*a_{n+m}^+a_n), \tag{3.1}$$

where the dependence of $\mathbf{d}_{n,n+m}$ on ξ is similar to the expression (2.5):

$$\mathbf{d}_{n,n+m}=-e\int\psi_n^*(\mathbf{r})\mathbf{r}\psi_{n+m}(\mathbf{r})d\mathbf{r}=(\mathbf{d}_1+i\mathbf{d}_2)\times\exp\left(-\chi(m-1)+(-1)^n(1-(-1)^m)\frac{\xi}{2}\right). \tag{3.2}$$

Here $\psi_n(\mathbf{r})$ is the atomic wave function of an electron located at the nth site and e is the electron charge. The choice of the phases of the wave functions $\psi_n(\mathbf{r})$ ensuring that the overlap integral (2.5) is a real quantity uniquely determines $\mathbf{d}_1$ and $\mathbf{d}_2$ in Eq. (3.2).

Substituting the expression (2.6) into Eq. (3.1) and using Eq. (3.2), we find

$$\mathbf{d}=\sum_k[(\mathbf{d}_1Q_k+\mathbf{d}_2P_k)c_k^+c_k+i(\mathbf{d}_1R_k+\mathbf{d}_2S_k)c_k^+c_{k-\pi}], \tag{3.3}$$

where

$$P_k=-\frac{\cosh\xi\sin k(\exp(2\chi)+1)+\sin(2k)\exp\chi}{\cosh(2\chi)-\cos(2k)}, \tag{3.4}$$

$$S_k=-\frac{\sinh\xi\cos k(\exp(2\chi)-1)}{\cosh(2\chi)-\cos(2k)}. \tag{3.5}$$

Switching in Eq. (3.3) to Fermi operators α_k and α_k^+ and using Eqs. (2.10) and (2.12) we obtain, finally,

$$\mathbf{d}=\sum_k\left[\left(\mathbf{d}_1\frac{\varepsilon_k}{b}+\mathbf{d}_2\frac{P_k+P_{k-\pi}\varphi_k^2-2S_k\varphi_k}{1+\varphi_k^2}\right)\alpha_k^+\alpha_k+i\mathbf{d}_2\frac{\varphi_k(P_k-P_{k-\pi})+S_k(1-\varphi_k^2)}{1+\varphi_k^2}\alpha_k^+\alpha_{k-\pi}\right]. \tag{3.6}$$

It should be noted that the operators (2.1) and (3.1) with $\mathbf{d}_2=0$ are formally similar. Therefore, the operators (2.11) and (3.6) are also similar.

Let the total dipole moment of the system be zero in the absence of an external electric field. Then it follows from Eqs. (3.6) and (2.13) that $\mathbf{d}_1=0$. Therefore the choice of phases of the wave functions $\psi_n(\mathbf{r})$ for which the integral $B_{n,n+m}$ in Eq. (2.1) is a real quantity makes in the present case the intersite matrix element of the dipole moment operator $\mathbf{d}_{n,n+m}$ in Eq. (3.1) imaginary. The case $\mathbf{d}_1\neq 0$, apparently, can be realized in systems possessing ferroelectric properties. Such systems are not studied in the present work. In the limit $\xi\to 0$, as is evident from Eq. (2.12), $\varphi_k\to 0$ for all $k\neq\pm\pi/2$. Therefore in Eq. (3.6) $\mathbf{d}_{k,k-\pi}\to 0$, and all dipole transitions are forbidden. If $\xi\neq 0$, then in Eq. (3.6) $\mathbf{d}_{k,k-\pi}\neq 0$, and the corresponding dipole transitions are allowed. Since in this case the first Brillouin zone is the region $k\in[-\pi/2,\pi/2]$, these transitions are vertical interband transitions in the spectrum (2.13).

4. DYNAMICAL EQUATION FOR THE ORDER PARAMETER OF A THERMODYNAMICALLY NONEQUILIBRIUM PEIERLS SYSTEM

Let us examine the behavior of the low-temperature phase of a Peierls system when nonequilibrium electron–hole pairs are excited in it. We shall assume that the excitation is due to induced transitions of electrons from the valence into the conduction band on account of, for example, a dipole electron–photon interaction with the incident radiation. It is known that the characteristic intraband relaxation time of electrons $\tau_e\sim 10^{-14}$ s is much shorter than the interband relaxation time $\tau\sim 10^{-11}$ s.[38] For this reason, when the system is irradiated with light with a constant amplitude, it can be assumed approximately[39] that thermodynamic equilibrium of electrons with their own Fermi quasilevel corresponding to the given band is established within each electronic band. The breakdown of thermodynamic equilibrium caused between the bands by the external irradiation is expressed by the difference of the corresponding Fermi quasilevels between one another.

The above-examined approach to the description of a thermodynamically nonequilibrium system, consisting of a collection of weakly interacting thermodynamically equilibrium subsystems, can also be extended to the case where the incident radiation has an adiabatically slowly varying amplitude A (the change ΔA in the field amplitude over the time $\tau_e\sim 10^{-14}$ s is much less than A). This is due to the fact that

the electronic subsystem within each band can follow completely the change in the field and therefore at any moment in time it is in a state of thermodynamic equilibrium. This situation is similar, to some extent, to the situation ordinarily encountered in the description of thermodynamically equilibrium systems with adiabatically slow variation of external parameters. In what follows, we shall confine our attention to constructing a theory for this particular case. In so doing, we shall not discuss the transient processes occurring in each electronic band over a characteristic time not exceeding the intraband electronic relaxation time $\tau_e \sim 10^{-14}$ s.

The free energy F_j of the electronic subsystem of the jth band ($j=1,2$) is determined by the relation

$$F_j = \mu_j N_j - k_B T \sum_k \ln\left(1+\exp\left(\frac{\mu_j-\varepsilon_k}{k_B T}\right)\right), \tag{4.1}$$

where μ_j and N_j are, respectively, the Fermi quasilevel and the number of electrons in the jth band. The summation over k in Eq. (4.1) extends over the range of the jth band of the spectrum (2.13) ($|k|<\pi/2$ for $j=1$ and $\pi/2<|k|<\pi$ for $j=2$).

Treating the parameter ξ as a generalized coordinate, we write the dynamic equation of the thermodynamically nonequilibrium Peierls system

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\xi}} - \frac{\partial L}{\partial \xi} = Q, \tag{4.2}$$

where

$$L = \sum_n \frac{m(\dot{x}_n)^2}{2} - F_1 - F_2 - F_c \tag{4.3}$$

is the Lagrangian function, Q is a generalized dissipative force [see Eq. (5.17)] characterizing the relaxation of the order parameter ξ of a metal–semiconductor phase transition to a stable position of equilibrium. In Eq. (4.3) m is the mass of an atom,

$$F_c = A\xi^2/2 \tag{4.4}$$

is the free energy of the crystal lattice, written in the static molecular-field approximation,[1] and in the harmonic approximation, taking account of only the first nonvanishing term in the Taylor series expansion in terms of the order parameter ξ of the metal–semiconductor phase transition with expansion coefficient A.

Substituting the relation (4.3) into Eq. (4.2) and using Eqs. (4.4), (4.1), and (2.3), we obtain

$$\frac{d^2\xi}{dt^2} = \frac{4}{NmR^2}\left(\sum_{|k|\leqslant\pi/2} \frac{\partial \varepsilon_k}{\partial \xi}\tanh\left(\frac{\varepsilon_k-\mu}{2k_B T}\right) - A\xi + QN\right). \tag{4.5}$$

The expression (4.5) is the dynamical equation of the Peierls system, determining the behavior of the order parameter ξ of the metal–semiconductor phase transition with excitation of nonequilibrium electron–hole pairs. In addition to the relation (4.5), we shall write an equation expressing electrical neutrality, relating the density n of electron–hole pairs and the Fermi quasilevel μ:

$$n = \frac{N}{2} - \sum_{|k|\leqslant\pi/2} \tanh\left(\frac{\varepsilon_k-\mu}{2k_B T}\right). \tag{4.6}$$

The expansion coefficient A in Eqs. (4.4) and (4.5) can be expressed in terms of the critical temperature T_0 of the thermodynamically equilibrium (in the absence of a light field) metal–semiconductor phase transition and other characteristics of the system. The condition of an equilibrium metal–semiconductor phase transition is instability of the metallic phase:

$$\partial^2 F(T_0, \xi=0)/\partial \xi^2 = 0,$$

where F is the free energy of the thermodynamically equilibrium Peierls system. In the absence of a light field, in the state of thermodynamic equilibrium there is no dissipation ($Q=0$), and the Fermi quasilevels of the valence and conduction bands of the spectrum (2.13) are zero:

$$\mu_{1,2} = \mp\mu = 0.$$

Hence we have, using Eqs. (4.5) and (2.13),

$$A = 2\sum_{|k|\leqslant\pi/2}\left(\frac{\partial^2 \varepsilon_k}{\partial \xi^2}\tanh\left(\frac{\varepsilon_k}{2k_B T_0}\right)\right)\Bigg|_{\xi=0}. \tag{4.7}$$

Calculating the sum in Eq. (4.7), we find approximately

$$A = \frac{4bN}{\pi}\left(\ln\left(\frac{\pi b}{2k_B T_0}\right)+1\right). \tag{4.8}$$

Thus, we have obtained Eq. (4.6), expressing the dependence of the density n of electron–hole pairs at the Fermi quasilevel μ and the order parameter ξ of the metal–semiconductor phase transition, i.e., $n(\mu,\xi)$, as well as the dynamical equation (4.5), describing the relation between the parameter ξ of the Peierls system and μ, i.e., $\xi(\mu)$. In turn, the Fermi quasilevel μ is determined by the degree of to which the light affects the system. The equation describing this effect should depend on the specific mechanism of the interaction of the radiation with the electronic subsystem.

5. RELAXATION OF THE ORDER PARAMETER ξ

To calculate the generalized dissipative force Q in Eq. (4.5), we note that the order parameter ξ (2.3) characterizes a phonon mode with wave vector $\mathbf{q}$ at the edge of the Brillouin zone ($q=\pi, \mathbf{q}$ is parallel to the principal axis of the Peierls crystal). Assuming that the relaxation of the phonon mode to the thermodynamically equilibrium value is due to phonon–phonon and phonon–electron interactions, we write the Hamiltonian of the system in the form

$$H = H_0 + H_e + V + U, \tag{5.1}$$

where

$$H_0 = \sum_{\mathbf{k}} \hbar\omega_{\mathbf{k}}(b_{\mathbf{k}}^+ b_{\mathbf{k}} + 1/2), \quad H_e = \sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} \alpha_{\mathbf{k}}^+ \alpha_{\mathbf{k}} \tag{5.2}$$

are, respectively, the Hamiltonians of the noninteracting phonons and electrons ($\omega_{\mathbf{k}}$ is the phonon spectrum, $\varepsilon_{\mathbf{k}}$ is the electron spectrum, $b_{\mathbf{k}}$, $b_{\mathbf{k}}^+$ ($\alpha_{\mathbf{k}}$, $\alpha_{\mathbf{k}}^+$) are, respectively, the operators creating and annihilating a phonon (electron) with wave (quasiwave) vector $\mathbf{k}$, and it is assumed that the sum-

mation in Eqs. (5.2) and subsequent formulas in this section extends over all branches of the phonon and electron spectra);

$$V=\sum_{\mathbf{k},\mathbf{q}} (V_{\mathbf{k},\mathbf{q}} b^{+}_{\mathbf{k}+\mathbf{q}} b_{\mathbf{k}} b_{\mathbf{q}} + \text{h.c.}) \tag{5.3}$$

is the phonon–phonon interaction operator, in writing which we confined ourselves only to three-phonon processes, which occur in a crystal lattice with cubic anharmonicity[31,40] ($V_{\mathbf{k},\mathbf{q}}$ is the matrix element for two phonons with wave vectors $\mathbf{k}$ and $\mathbf{q}$ to merge into a phonon with wave vector $\mathbf{k}+\mathbf{q}$);

$$U=\sum_{\mathbf{k},\mathbf{q}} U_{\mathbf{k},\mathbf{q}} \alpha^{+}_{\mathbf{k}+\mathbf{q}} \alpha_{\mathbf{k}} (b_{\mathbf{q}} - b^{+}_{-\mathbf{q}}) \tag{5.4}$$

is the Fröhlich electron–phonon interaction Hamiltonian[1,41] ($U_{\mathbf{k},\mathbf{q}}$ is the matrix element of the transition of an electron with quasiwave vector $\mathbf{k}$ and a phonon with wave vector $\mathbf{q}$ into an electron with quasiwave vector $\mathbf{k}+\mathbf{q}$).

The time variation of the operator $f_{\mathbf{k},\mathbf{p}} \equiv b^{+}_{\mathbf{p}} b_{\mathbf{q}}$ at the kinetic stage of evolution in second order perturbation theory in $V+U$ can be described by the equation[42]

$$\frac{df_{\mathbf{k},\mathbf{p}}}{dt} = i\hbar \,\mathrm{Tr}\,\rho[H_0, f_{\mathbf{k},\mathbf{p}}] + L^{(1)}_{\mathbf{k},\mathbf{p}} + L^{(2)}_{\mathbf{k},\mathbf{p}}, \tag{5.5}$$

where

$$L^{(1)}_{\mathbf{k},\mathbf{p}} = i\hbar \,\mathrm{Tr}\,\rho[U+V, f_{\mathbf{k},\mathbf{p}}], \tag{5.6}$$

$$L^{(2)}_{\mathbf{k},\mathbf{p}} = -\hbar \int_{-\infty}^{0} d\tau\, e^{\eta\tau} \mathrm{Tr}\,\rho \Bigg[U(\tau)+V(\tau), [U+V, f_{\mathbf{k},\mathbf{p}}] + i\sum_{\mathbf{q},\mathbf{s}} f_{\mathbf{q},\mathbf{s}} \frac{\partial L^{(1)}_{\mathbf{k},\mathbf{p}}}{\partial \gamma_{\mathbf{q},\mathbf{s}}} \Bigg], \quad \eta \to +0. \tag{5.7}$$

Here

$$U(\tau)+V(\tau) = \exp\left\{ i\frac{H_0+H_e}{\hbar}\tau \right\} (U+V) \times \exp\left\{ -i\frac{H_0+H_e}{\hbar}\tau \right\} \tag{5.8}$$

is the operator $U+V$ in the interaction representation, ρ is a statistical operator (density matrix) of an ideal nonequilibrium gas of phonons and electrons, which determines the c-number functions $\gamma_{\mathbf{k},\mathbf{p}}$ by means of the equations

$$\mathrm{Tr}\,\rho = 1, \quad \mathrm{Tr}\,\rho f_{\mathbf{k},\mathbf{p}} = \gamma_{\mathbf{k},\mathbf{p}}. \tag{5.9}$$

Substituting U and V from Eqs. (5.3) and (5.4) into Eqs. (5.5)–(5.8) and using Wick's theorem,[43] we obtain an equation for the number $N_{\mathbf{k}}$ of phonons in a state with wave vector $\mathbf{k}$:

$$\frac{dN_{\mathbf{k}}}{dt} = 4\pi \sum_{\mathbf{q}} |V_{\mathbf{q},\mathbf{k}-\mathbf{q}}|^2 \{ N_{\mathbf{q}} N_{\mathbf{k}-\mathbf{q}} (1+N_{\mathbf{k}}) - (1+N_{\mathbf{q}})(1+N_{\mathbf{k}-\mathbf{q}}) N_{\mathbf{k}} \} \delta(\omega_{\mathbf{q}} + \omega_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{k}}) + 8\pi \sum_{\mathbf{q}} |V_{\mathbf{q},\mathbf{k}}|^2 \{ (1+N_{\mathbf{q}})(1+N_{\mathbf{k}}) N_{\mathbf{q}+\mathbf{k}} - N_{\mathbf{q}} N_{\mathbf{k}} (1+N_{\mathbf{q}+\mathbf{k}}) \} \delta(\omega_{\mathbf{q}} + \omega_{\mathbf{k}} - \omega_{\mathbf{q}+\mathbf{k}}) + 2\pi \sum_{\mathbf{q}} |U_{\mathbf{q},\mathbf{k}}|^2 \{ (1-n_{\mathbf{q}}) n_{\mathbf{q}+\mathbf{k}} (1+N_{\mathbf{k}}) - n_{\mathbf{q}} (1-n_{\mathbf{q}+\mathbf{k}}) N_{\mathbf{k}} \} \delta(\omega_{\mathbf{k}} - (\varepsilon_{\mathbf{q}+\mathbf{k}} - \varepsilon_{\mathbf{q}})/\hbar), \tag{5.10}$$

where $n_{\mathbf{k}}$ is the number of electrons in a state with quasiwave vector $\mathbf{k}$.

Using the expression (5.10), we find for the deviation $\delta N_{\mathbf{k}} = N_{\mathbf{k}} - N^{(0)}_{\mathbf{k}}$ of the number of phonons $N_{\mathbf{k}}$ from the thermodynamically equilibrium value $N^{(0)}_{\mathbf{k}}$

$$\frac{d\,\delta N_{\mathbf{k}}}{dt} = -\frac{\delta N_{\mathbf{k}}}{\tau_p}, \tag{5.11}$$

where

$$\frac{1}{\tau_p} = 4\pi \sum_{\mathbf{q}} |V_{\mathbf{q},\mathbf{k}-\mathbf{q}}|^2 (N^{(0)}_{\mathbf{q}} + N^{(0)}_{\mathbf{k}-\mathbf{q}} + 1) \delta(\omega_{\mathbf{q}} + \omega_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{k}}) + 8\pi \sum_{\mathbf{q}} |V_{\mathbf{q},\mathbf{k}}|^2 (N^{(0)}_{\mathbf{q}} - N^{(0)}_{\mathbf{k}+\mathbf{q}}) \delta(\omega_{\mathbf{q}} + \omega_{\mathbf{k}} - \omega_{\mathbf{q}+\mathbf{k}}) + 2\pi \sum_{\mathbf{q}} |U_{\mathbf{q},\mathbf{k}}|^2 (n^{(0)}_{\mathbf{q}} - n^{(0)}_{\mathbf{k}+\mathbf{q}}) \delta(\omega_{\mathbf{k}} - (\varepsilon_{\mathbf{q}+\mathbf{k}} - \varepsilon_{\mathbf{q}})/\hbar) \tag{5.12}$$

is the reciprocal of the phonon relaxation time. Here

$$N^{(0)}_{\mathbf{q}} = (\exp\{\hbar\omega_{\mathbf{q}}/k_B T\} - 1)^{-1} \tag{5.13}$$

is the Bose–Einstein distribution,

$$n^{(0)}_{\mathbf{q}} = (\exp\{(\varepsilon_{\mathbf{q}} - \mu)/k_B T\} + 1)^{-1} \tag{5.14}$$

is the Fermi–Dirac distribution, and μ is the Fermi quasilevel. The first and second terms on the right-hand side of Eq. (5.12) correspond to the phonn–phonon relaxation and the third term corresponds to phonon–electron relaxation.

It follows from Eq. (5.11) that in the classical limit $N_{\mathbf{k}} \gg N^{(0)}_{\mathbf{k}}$ the statistical average value of the energy of a phonon mode (2.3)

$$\langle W \rangle = \frac{mR^2}{4} \langle \dot{\xi}^2 \rangle \tag{5.15}$$

satisfies the equation

$$\frac{d\langle W\rangle}{dt} = -\frac{\langle W\rangle}{\tau_p} = \langle Q\dot{\xi}\rangle. \tag{5.16}$$

The second equality in Eq. (5.16), taking account of Eq. (5.15), holds if the generalized dissipative force Q satisfies

$$Q = -\frac{mR^2}{4\tau_p}\dot{\xi}, \tag{5.17}$$

characterizing the linear damping of the phonon mode.

6. INTERACTION WITH RADIATION

We shall describe the interaction of the system with a light field by means of the operator V_1, which in the dipole approximation has the form

$$V_1 = -\mathbf{d}\cdot\mathbf{E}(t) = -\mathbf{d}\cdot\int \mathbf{E}_\omega \exp(-i\omega t)\,d\omega, \tag{6.1}$$

where $\mathbf{E}_\omega$ and ω are, respectively, the amplitude and frequency of the spectral component of the light field.

Let us consider the case where the incident radiation $\mathbf{E}(t)$ is a linearly polarized (along the axis of the crystal) quasimonochromatic stationary random process.[44] Then all spectral components $\mathbf{E}_\omega$ are statistically independent[44,45]

$$\langle \mathbf{E}_\omega \cdot \mathbf{E}_{\omega_1}\rangle = G(\omega)\,\delta(\omega+\omega_1). \tag{6.2}$$

Here $G(\omega)$ is the spectral density of the light field, which for a quasimonochromatic signal can be represented as[44]

$$G(\omega) = Ig(|\omega|-\omega_0), \tag{6.3}$$

where ω_0 is the carrier frequency, and $g(x)$ is a nonnegative bell-shaped function with a maximum at the point $x=0$ and satisfies the normalization condition

$$\int g(x)\,dx = 1. \tag{6.4}$$

The width $\Delta\omega$ of the spectrum $G(\omega)$ satisfies the inequality $\Delta\omega \ll \omega_0$. The quantity

$$I = \int G(\omega)\,d\omega/2 \tag{6.5}$$

characterizes the light intensity (in a Gaussian system, to within the factor $cn/2\pi$, where in the present case c is the speed of light and n is the index of refraction of the medium).

Using the Liouville equation[46]

$$i\hbar\frac{\partial\rho}{\partial t} = [H_e + V_1, \rho], \tag{6.6}$$

taking account of Eqs. (6.1) and (6.2), we find an equation for the diagonal elements ρ_{kk} of the density matrix ρ of the electronic subsystem in second-order perturbation theory in V_1:

$$\frac{\partial\rho_{kk}}{\partial t} = \frac{2\pi}{\hbar^2}\sum_s |\mathbf{d}_{ks}|^2 G\left(\frac{\varepsilon_s-\varepsilon_k}{\hbar}\right)(\rho_{ss}-\rho_{kk}), \tag{6.7}$$

where $\mathbf{d}_{ks}$ is the matrix element of the dipole moment operator (3.6). In the particular case of a monochromatic light field

$$\mathbf{E}(t) = \mathbf{E}_0\cos(\omega_0 t+\varphi) \tag{6.8}$$

with a uniformly distributed phase φ, the spectral density $G(\omega)$ has the form

$$G(\omega) = \frac{\mathbf{E}_0^2(\delta(\omega-\omega_0)+\delta(\omega+\omega_0))}{4}. \tag{6.9}$$

Then Eq. (6.7) becomes the well-known Fermi golden rule for the probability of induced transitions:[45]

$$\frac{\partial\rho_{kk}}{\partial t} = \frac{\pi}{2\hbar}|\mathbf{E}_0\cdot\mathbf{d}_{kk}|^2\,\delta(2\varepsilon_k-\hbar\omega_0). \tag{6.10}$$

In Eq. (6.10), it is assumed that the lower level (with quasiwave number $k-\pi$) is filled, and the upper level (with quasiwave number k) is empty.

Taking account of Eqs. (3.6) and (2.13), we obtain from Eq. (6.7)

$$\frac{\partial\rho_{kk}}{\partial t} = \frac{2\pi}{\hbar^2}d_k^2 G\left(\frac{2\varepsilon_k}{\hbar}\right)\tanh\left(\frac{\varepsilon_k-\mu}{2k_BT}\right), \tag{6.11}$$

where

$$d_k = \left| d_2\frac{\varphi_k(P_k-P_{k-\pi})+S_k(1-\varphi_k^2)}{1+\varphi_k^2}\right|. \tag{6.12}$$

Taking into consideration the relation (2.13), we find from Eq. (6.11) the kinetic equation for the density $n = 2\Sigma_{|k|<\pi/2}\rho_{kk}$ of electron–hole pairs in a Peierls system:

$$\frac{\partial n}{\partial t} = \frac{4\pi}{\hbar^2}\sum_{|k|\leqslant\pi/2} d_k^2\tanh\left(\frac{\varepsilon_k-\mu}{2k_BT}\right)G\left(\frac{2\varepsilon_k}{\hbar}\right) - \frac{n-n_0}{\tau}, \tag{6.13}$$

where n_0 is the density of electron–hole pairs in the absence of light and, using the relation (4.6), can be written in the approximate form

$$n_0 = n(\mu=0) = \begin{cases} N\sqrt{\dfrac{k_BT\sinh\xi}{\pi b}}\exp\left\{-\dfrac{2b\sinh\xi}{k_BT}\right\}, & b\xi \gg k_BT, \\ \dfrac{N}{\pi}\left(\dfrac{k_BT}{b}\ln 2 - \dfrac{b}{2k_BT}\xi^2\right), & b\xi \ll k_BT. \end{cases} \tag{6.14}$$

The last term on the right-hand side of Eq. (6.13) takes account of interband electronic relaxation with characteristic interband electronic relaxation time τ. The time τ for radiative and nonradiative recombination of electron–hole pairs and for recombination with participation of impurities and defects depends, in the general case, on the density of electrons and holes (see Ref. 39, p. 297), but in what follows we shall neglect this dependence, assuming τ to be a prescribed parameter. The expression (6.13) shows that the change in the density n of electron–hole pairs is due to the interaction of the electronic subsystem with the spectral components of the light which have the frequency $\omega_k = 2\varepsilon_k/\hbar$.

The relations (4.6), (4.5), and (6.13), taking account of Eq. (5.17), form a closed system of dynamical equations for the internal parameters μ, n, and ξ of a thermodynamically nonequilibrium Peierls system with prescribed external parameters T, I, ω_0, and so on.

7. DYNAMICAL EQUATION FOR A NONDEGENERATE OR WEAKLY DEGENERATE SEMICONDUCTOR WITH EXCITATION OF ELECTRON–HOLE PAIRS INTO THE CONDUCTION BAND BY MONOCHROMATIC LIGHT

We shall analyze Eqs. (4.5) and (4.6) under the assumption that the Peierls system is a nondegenerate or weakly degenerate semiconductor:

$$\mu - 2b\sinh\xi < 2k_BT. \tag{7.1}$$

The relation (7.1), which imposes a restriction on the range of variation of the Fermi quasilevel μ, taking account

of Eq. (4.6), is equivalent to an approximate inequality limiting the density n of electron–hole pairs from above:

$$n<n_1=\frac{8N}{3\pi}\sqrt{\frac{k_BT\sinh\xi}{b}}\,. \tag{7.2}$$

Using the formulas (4.5) and (4.6), and taking account of Eq. (5.17), we find an approximate equation for the order parameter ξ of a metal–semiconductor phase transition:

$$\frac{d^2\xi}{dt^2}+\frac{1}{\tau_p}\frac{d\xi}{dt}=\frac{4}{NmR^2}\left(\frac{4bN}{\pi}\sinh\xi K(\sqrt{1-\tanh^2\xi})\right.$$
$$\left.-4bn\cosh\xi-A\xi\right), \tag{7.3}$$

where $K(x)$ is the complete normal elliptic integral of the first kind.

Since in real physical systems the density of electron–hole pairs satisfies $n\ll N$, and the order parameter of the metal–semiconductor phase transition of the system $\xi\leqslant 0.5$,[1–4] from Eq. (7.3) we obtain approximately the following equation:

$$\frac{d^2\xi}{dt^2}+\frac{1}{\tau_p}\frac{d\xi}{dt}=\frac{16b}{NmR^2}\left(\frac{N\xi}{\pi}\ln\left|\frac{\xi_0}{\xi}\right|-n\,\mathrm{sign}\,\xi\right), \tag{7.4}$$

where

$$\xi_0=\frac{\pi}{2}\exp\left\{\mathrm{Arcsin}\left(\frac{\pi}{4}\right)-\frac{A\pi}{4bN}\right\} \tag{7.5}$$

is the order parameter of the metal–semiconductor phase transition for $n=0$.

We note that Eq. (7.4), obtained assuming the inequality (7.1) [or under the assumption (7.2), which is equivalent to (7.1)], is valid, as our analysis shows, even in the important particular case where the temperature $T=0$ and the density n of electron–hole pairs is arbitrary.

Let us analyze the relation (6.13), describing the change in the density n of electron–hole pairs under irradiation, for monochromatic light, where the form factor $g(x)$ in Eq. (6.3) is determined by the relation

$$g(x)=\delta(x).$$

We switch in Eq. (6.13) from summation to integration, i.e., $\Sigma\to\int dE$, keeping in mind the fact that the electron density of states $\nu(E)$ corresponding to the spectrum (2.13) for the case $\chi\gg 1$ has the form

$$\nu(E)=\frac{2N|E|}{\pi\sqrt{(4b^2\cosh^2\xi-E^2)(E^2-4b^2\sinh^2\xi)}}\,, \tag{7.6}$$

and the matrix element of the dipole moment operator d_k (6.12) with $\varepsilon_k=E$ is determined, in accordance with Eqs. (2.12), (2.13), (3.4), and (3.5), by the relation

$$d(E)\equiv d_k(\varepsilon_k=E)=\frac{4bd_2\cosh\xi\sinh\xi}{E}. \tag{7.7}$$

Then, assuming that there is no saturation of interband optical transitions, so that the relation

$$\tanh((\hbar\omega_0/2-\mu)/2k_BT)=1, \tag{7.8}$$

holds approximately, we obtain from Eq. (6.13)

$$\frac{dn}{dt}=-\frac{n-n_0}{\tau}+\frac{512Ib^2d_2^2\cosh^2\xi\sinh^2\xi}{\hbar^2\omega_0\sqrt{(16b^2\cosh^2\xi-(\hbar\omega_0)^2)((\hbar\omega_0)^2-16b^2\sinh^2\xi)}}\,. \tag{7.9}$$

Let us consider the case where the frequency ω_0 of the incident light is such that optical transitions of electrons occur from the valence band into the conduction band, Then we can set approximately in Eq. (7.9)

$$16b^2\cosh^2\xi-(\hbar\omega_0)^2\approx(\hbar\omega_0)^2-16b^2\sinh^2\xi. \tag{7.10}$$

Then we find approximately

$$\frac{dn}{dt}=-\frac{n-n_0}{\tau}+\frac{64Id_2^2N\xi^2}{\hbar^2\omega_0}. \tag{7.11}$$

Thus, we have obtained Eq. (7.11), describing the behavior of the density $n(t)$ of electron–hole pairs when the system is irradiated with monochromatic light with intensity $I(t)$, and Eq. (7.4), determining the time dependence $\xi(t)$ of the order parameter with a variation of the density n. These equations (7.4) and (7.11) are the basic dynamical equations for a nondegenerate or weakly degenerate [see Eq. (7.1)] semiconductor state of a Peierls system with optical excitation of electrons in the system into the allowed band.

8. STATIONARY SOLUTION AND ITS STABILITY

Let us consider first the case where a Peierls semiconductor is irradiated with light with constant intensity. Setting in the system of equations (7.4) and (7.11)

$$\frac{dn}{dt}\equiv\frac{d\xi}{dt}\equiv 0, \tag{8.1}$$

we find its stationary solution ξ_s, n_s as a function of the intensity I in the implicit form

$$I=\frac{\hbar^2\omega_0}{64d_2^2N\xi_s^2\tau}\left(\frac{N|\xi_s|}{\pi}\ln\left|\frac{\xi_0}{\xi_s}\right|-n_0(\xi_s,T)\right), \tag{8.2}$$

$$n_s=\frac{N|\xi_s|}{\pi}\ln\left|\frac{\xi_0}{\xi_s}\right|. \tag{8.3}$$

In order that the stationary solution (8.2) and (8.3) satisfy the

approximation (7.2), used above, in the entire range of variation of ξ ($0<\xi\leqslant\xi_0$), the relation limiting the temperature T of the system from below

$$\left(\frac{3}{4e}\right)^2 \xi_0 b < k_B T. \tag{8.4}$$

must be satisfied.

At lower temperatures, where the relation (8.4) is invalid, the solution (8.2) and (8.3) satisfies the condition (7.2) only in the region

$$\xi\in(0,\xi_1)\cup(\xi_2,\xi_0],$$

where ξ_1, ξ_2 ($\xi_1<\xi_2$) are the roots of the transcendental equation

$$\sqrt{\xi}\ln\left|\frac{\xi_0}{\xi}\right| = \frac{8}{3}\sqrt{\frac{k_B T}{b}}\,. \tag{8.5}$$

Analysis of the system of equations (7.4) and (7.11) shows that the solution (8.2) and (8.3) is stable if

$$\xi_s > \xi_c\,, \tag{8.6}$$

where

$$\xi_c = \xi_0 \exp\left\{-1-\frac{\pi m R^2}{16 b \tau_p \tau}\right\}. \tag{8.7}$$

Using the solution (8.2) and (8.3), we find a condition, equivalent to the inequality (8.6), that imposes an upper bound on the light intensity I

$$I < I_c\,, \tag{8.8}$$

where

$$I_c = \frac{\hbar^2\omega_0}{64 d_2^2 N \xi_c^2 \tau}(n_c - n_0(\xi_c, T)), \tag{8.9}$$

$$n_c = \frac{N\xi_0}{\pi}\left(1+\frac{\pi m R^2}{16 b\tau_p\tau}\right)\exp\left\{-1-\frac{\pi m R^2}{16 b\tau_p\tau}\right\}. \tag{8.10}$$

Thus, for adiabatically slow increase of intensity I of the incident monochromatic light from zero to the value I_c, determined in Eq. (8.9), the density n of electron–hole pairs in the accordance with the description (8.2) and (8.3) increases from the thermodynamically equilibrium (no irradiation) value n_0 (6.14) to the critical value n_c determined by Eq. (8.10), and the order parameter ξ decreases from a value approximately equal to ξ_0 to ξ_c determined by Eq. (8.7). When $I>I_c$, the stationary semiconductor state corresponding to Eqs. (8.2) and (8.3) becomes unstable, and there is no other stable stationary solution of the system of equations Eqs. (7.4) and (7.11) corresponding to a semiconductor phase.

We shall make numerical estimates of n_c and I_c for vanadium dioxide, whose one-dimensional electronic conduction band is formed by overlapping of the $3d$ wave functions of vanadium atoms arranged in the form of chains parallel to the crystal axis c.[4] At temperatures below the critical value $T_0 = 340\,\mathrm{K}$ the vanadium atoms in the chain converge in pairs, and a gap forms in the electronic spectrum at the Fermi level, so that the low-temperature phase of VO_2 can be regarded as a one-dimensional (quasi-one-dimensional) Peierls semiconductor.[2–4]

Taking characteristic values of the physical quantities for VO_2 $b\approx 0.3\,\mathrm{eV}$, $N\approx 10^{23}\ \mathrm{cm}^{-3}$, $\xi_0\approx 0.5$,[3,4] $\tau\approx 3\times 10^{-11}\,\mathrm{s}$, $\tau_p\sim 10^{-13}\,\mathrm{s}$,[38] $d_2\sim 10^{-18}\,\mathrm{CGS}$,[45,47] $\hbar\omega_0 = 1.17\,\mathrm{eV}$,[35] $m\approx 8.5\times 10^{-23}\,\mathrm{g}$, $R\approx 0.5\times 10^{-8}\,\mathrm{cm}$, we obtain from Eqs. (8.9) and (8.10), taking account of Eq. (8.7),

$$n_c\sim 10^{21}\ \mathrm{cm}^{-3}, \quad I_c\sim 10^8\,\mathrm{W/cm}^2. \tag{8.11}$$

The condition (8.4) gives $T>100\,\mathrm{K}$.

9. DYNAMICS OF THE SYSTEM AT SHORT TIMES

We shall investigate the behavior of a Peierls system irradiated by monochromatic light whose intensity I is a step function of the time:

$$I(t) = \begin{cases} 0, & t<0, \\ I=\mathrm{const}, & t\geqslant 0. \end{cases} \tag{9.1}$$

Let us examine the solution at the initial stage of evolution, when the order parameter ξ of the metal–semiconductor phase transition is close to its initial value ξ_0. Then the system of equations (7.4) and (7.11) can be approximately linearized, and in the approximation[48]

$$\left|\frac{d^2\xi}{dt^2}\right| \ll \frac{1}{\tau_p}\left|\frac{d\xi}{dt}\right| \tag{9.2}$$

it can be written as

$$\frac{d\xi}{dt} = \frac{16 b\tau_p}{N m R^2}\left(\frac{N}{\pi}(\xi_0-\xi)-n\right), \tag{9.3}$$

$$\frac{dn}{dt} = -\frac{n-n_0}{\tau} + \frac{64 d_2^2 N\xi_0}{\hbar^2\omega_0} I\xi. \tag{9.4}$$

The initial conditions for (9.3) and (9.4) are

$$\xi(t=0) = \xi_0 - \frac{\pi n_0}{N}, \quad n(t=0) = n_0. \tag{9.5}$$

Solving the problem (9.3)–(9.5), we obtain

$$\xi = \xi_1 + \frac{\xi_0-\xi_1-\pi n_0/N}{\lambda_2-\lambda_1}(\lambda_2 e^{\lambda_1 t} - \lambda_1 e^{\lambda_2 t}), \tag{9.6}$$

$$n = n_1 - \frac{(\xi_0-\xi_1)N-\pi n_0}{\lambda_2-\lambda_1}\left(\lambda_2\left(\frac{\lambda_1 m R^2}{16 b\tau_p}+\frac{1}{\pi}\right)e^{\lambda_1 t} - \lambda_1\left(\frac{\lambda_2 m R^2}{16 b\tau_p}+\frac{1}{\pi}\right)e^{\lambda_2 t}\right), \tag{9.7}$$

where

$$\xi_1 = \left(\xi_0 - \frac{\pi n_0}{N}\right)\frac{\hbar^2\omega_0}{\hbar^2\omega_0 + 64\pi\tau I\xi_0 d_2^2}, \tag{9.8}$$

$$n_1 = \left(\frac{N\xi_0}{\pi} - n_0\right)\frac{64\pi\tau I\xi_0 d_2^2}{\hbar^2\omega_0 + 64\pi\tau I\xi_0 d_2^2}, \tag{9.9}$$

$$\lambda_{1,2}=-\frac{8b\tau_p}{\pi mR^2}\left[1\pm\sqrt{1-\frac{\pi mR^2}{4b\tau_p\tau}\left(1+\frac{64\pi\tau I\xi_0 d_2^2}{\hbar^2\omega_0}\right)}\right]. \tag{9.10}$$

Taking account of the characteristic numerical values of the basic parameters (see below or the numerical estimates at the end of Sec. 8) and confining ourselves to the case $I \ll 10^{10}\,\mathrm{W/cm^2}$, we find approximately from Eq. (9.10)

$$\lambda_1=-\frac{16b\tau_p}{\pi mR^2}, \tag{9.11}$$

$$\lambda_2=-\frac{1}{\tau}-\frac{64\pi I\xi_0 d_2^2}{\hbar^2\omega_0}. \tag{9.12}$$

It is evident from Eqs. (9.11) and (9.12) that for the characteristic numerical values of the parameters (see the end of Sec. 8) the inequality $|\lambda_1|\gg|\lambda_2|$ is satisfied. For this reason, in accordance with the expression (9.6), the characteristic time τ_0 of a transition into a new phase is determined approximately by the relation

$$\tau_0=\frac{1}{|\lambda_2|}=\frac{\hbar^2\omega_0\tau}{\hbar^2\omega_0+64\pi I\xi_0 d_2^2\tau}. \tag{9.13}$$

The condition for the applicability of our approximation (9.2), taking account of the expression (9.6), can be written in the form $\tau_0\gg\tau_p$. Hence, using the relation (9.13), we find an inequality that gives an upper bound on the light intensity I:

$$I\ll\frac{\hbar^2\omega_0}{64\pi\xi_0 d_2^2}\left(\frac{1}{\tau_p}-\frac{1}{\tau}\right). \tag{9.14}$$

Substituting into Eq. (9.14) the characteristic numerical values of the parameters for vanadium dioxide, specifically, $\xi_0\approx 0.5$,[3,4] $\tau\approx 3\times10^{-11}$ s, $\tau_p\sim 10^{-13}$ s,[38] $d_2\sim 10^{-18}$ CGS,[45,47] and $\hbar\omega_0=1.17$ eV,[35] we obtain the condition of applicability of the approximation (9.2): $I\ll 10^{10}\,\mathrm{W/cm^2}$, which is the same as the approximation which we used in the derivation of Eqs. (9.11) and (9.12).

In the experiment of Ref. 35, where a vanadium dioxide film was irradiated with powerful laser radiation with intensity $I\simeq 7\cdot10^8\,\mathrm{W/cm^2}$, a photoinduced semiconductor–metal phase transition, occurring over a characteristic time $\tau_0\simeq 10^{-12}$ s, was observed. A numerical estimate using Eq. (9.13) gives $\tau_0\simeq 1.6\times10^{-12}$ s. Therefore the theoretical value τ_0 calculated on the basis of the theory developed agrees well with the experimental data of Ref. 35.

10. DYNAMICS OF THE SYSTEM AT LONG TIMES

We shall now investigate the behavior of a Peierls system under irradiation by monochromatic light, whose intensity I is a step function (9.1) of the time, at the final stage of evolution when the order parameter ξ of the metal–semiconductor phase transition is close to its new stable value corresponding to a position of dynamic equilibrium, $\xi\approx\xi_s(I)$ [see Eq. (8.2)]. We shall confine ourselves to the intensities $I<I_c$, where the final state of the system is the semiconductor phase (8.2), (8.3).

The system of equations (7.4) and (7.11), linearized near the stationary solution ξ_s, n_s, determined by Eqs. (8.2) and (8.3), has in the approximation (9.2) the form

$$\frac{d\xi}{dt}=\frac{16b\tau_p}{NmR^2}\left[\frac{N}{\pi}\ln\left|\frac{\xi_0}{e\xi_s}\right|(\xi-\xi_s)-(n-n_s)\right], \tag{10.1}$$

$$\frac{dn}{dt}=\frac{128Id_2^2N\xi_s}{\hbar^2\omega_0}(\xi-\xi_s)-\frac{n-n_s}{\tau}. \tag{10.2}$$

The solution of the system of equations (10.1) and (10.2) can be written as

$$\begin{pmatrix}\xi\\ n\end{pmatrix}=\begin{pmatrix}\xi_s\\ n_s\end{pmatrix}+\begin{pmatrix}a_{11}a_{12}\\ a_{21}a_{22}\end{pmatrix}\begin{pmatrix}\exp\{\lambda_1 t\}\\ \exp\{\lambda_2 t\}\end{pmatrix}, \tag{10.3}$$

where

$$\lambda_{1,2}=\frac{8b\tau_p}{\pi mR^2}\ln\left|\frac{\xi_0}{e\xi_s}\right|-\frac{1}{2\tau}\pm\sqrt{\left(\frac{8b\tau_p}{\pi mR^2}\ln\left|\frac{\xi_0}{e\xi_s}\right|-\frac{1}{2\tau}\right)^2+\frac{16b\tau_p}{\pi mR^2\tau}\ln\left|\frac{\xi_0}{e\xi_s}\right|-\frac{2^{11}b\tau_p d_2^2N\xi_s I}{\hbar^2\omega_0}}, \tag{10.4}$$

and the constant coefficients a_{ij} in the matrix are determined from Eqs. (10.1) and (10.2) and the initial conditions.

It is evident from the relation (10.4) that when the condition (8.6) is satisfied [or the inequality (8.8) equivalent to it] the solution (10.3) has the form of a process in which ξ and n relax to a stationary state ξ_s, n_s of stable dynamic equilibrium determined by Eqs. (8.2)–(8.3). For low light intensities I, when $\xi_s=\xi_s(I)$ [see Eq. (8.2)] satisfies the approximate inequality

$$\xi_0>\xi_s(I)>\xi_c\left(1+\sqrt{\frac{\pi mR^2}{2b\tau\tau_p}}\right), \tag{10.5}$$

the radicand in Eq. (10.4) is positive, and therefore, as $t\to\infty$, the asymptotic behavior of $\xi(t)$ and $n(t)$ has the exponential form

$$\begin{cases}\xi-\xi_s\propto\exp(-\gamma t),\\ n-n_s\propto\exp(-\gamma t)\end{cases} \tag{10.6}$$

with relaxation constant $\gamma=-\lambda_1>0$ (10.4) (in Eq. (10.3) at long times t the exponential with the smaller decay constant plays the main role).

The inequality (10.5) for ξ_s is equivalent to the condition bounding the light intensity I from above

$$I<I_1, \tag{10.7}$$

where

$$I_1\equiv I\left(\xi_0\left(1+\sqrt{\frac{\pi mR^2}{2b\tau\tau_p}}\right)\exp\left\{-1-\frac{\pi mR^2}{16b\tau\tau_p}\right\}\right), \tag{10.8}$$

and the functional dependence $I(\xi_s)$ is determined by the expression (8.2).

If the light intensity I satisfies the condition

$$I_1<I<I_c, \tag{10.9}$$

for which for $\xi_s=\xi_s(I)$ [see Eq. (8.2)] the approximate inequality

$$\xi_c<\xi_s(I)<\xi_c\left(1+\sqrt{\frac{\pi mR^2}{2b\tau\tau_p}}\right) \tag{10.10}$$

is satisfied, then the radicand in Eq. (10.4) is negative, and at long times t the behavior of $\xi(t)$ and $n(t)$ has the character of damped oscillations

$$\begin{cases}\xi-\xi_s\propto\exp(-\gamma t)\cos(\omega t+\varphi_1),\\ n-n_s\propto\exp(-\gamma t)\cos(\omega t+\varphi_2)\end{cases} \tag{10.11}$$

with circular frequency

$$\omega=\frac{8b\tau_p}{\pi mR^2}\sqrt{\frac{\pi mR^2}{2b\tau\tau_p}-\ln^2\left|\frac{\xi_s}{\xi_c}\right|} \tag{10.12}$$

and decay constant

$$\gamma=\frac{8b\tau_p}{\pi mR^2}\ln\left|\frac{\xi_s}{\xi_c}\right|. \tag{10.13}$$

It is evident from Eqs. (10.12) and (10.13) that as the intensity I approaches the critical value I_c from below (and $\xi_s\to\xi_c+0$), the frequency ω (10.12) of the oscillations of the order parameter ξ of the metal–semiconductor phase transition and the density n of electron–hole pairs increases to its maximum value ω_m:

$$\omega_m=\sqrt{\frac{32b\tau_p}{\pi mR^2\tau}}, \tag{10.14}$$

and the damping coefficient γ decreases to zero.

Let us make some numerical estimates. Taking the numerical values characteristic for the physical parameter of VO_2, specifically, $b\approx0.3$ eV, $N\approx10^{23}$ cm^{-3}, $\xi_0\approx0.5$,[3,4] $\tau\approx3\times10^{-11}$ s, $\tau_p\sim10^{-13}$ s,[38] $d_2\sim10^{-18}$ CGS,[45,47] $\hbar\omega_0=1.17$ eV,[35] $m\approx8.5\times10^{-23}$ g, and $R\approx0.5\times10^{-8}$ cm, we obtain from Eqs. (10.14) and (10.13)

$$\omega(I=I_1)=0,\quad \gamma(I=I_1)\cong\sqrt{\frac{32b\tau_p}{\pi mR^2\tau}}\approx3\times10^{12}\ \mathrm{s}^{-1},$$

$$\omega_m=\omega(I=I_c)\approx3\times10^{12}\ \mathrm{s}^{-1},\quad \gamma(I=I_c)=0. \tag{10.15}$$

Therefore near the critical point $I\approx I_c$ ($I<I_c$) the characteristic relaxation time γ^{-1} (10.13) of a Peierls system to the position of stable dynamic equilibrium $\xi_s\approx\xi_c$ ($\xi_s>\xi_c$) becomes anomalously long ($\gamma^{-1}\to\infty$ for $I\to I_c-0$).

11. DISCUSSION OF THE EXPERIMENT

To check experimentally the solution (8.2)–(8.3) describing the behavior of the order parameter ξ of the metal–semiconductor phase transition and the density n of electron–hole pairs in the stationary state or with adiabatically slow variation of the light intensity I, when the characteristic time of the smooth increase of the intensity in a pulse is much longer than the characteristic relaxation time of the process $\tau_0\sim10^{-12}$ s (9.13), it is desirable to use a Peierls-unstable material in the form of a thin film placed into a material which is transparent at the frequency of the laser radiation and possesses good heat removal (for example, in superfluid helium). This makes it possible to avoid excessive heating of the system even for quite high irradiation intensities I close to I_c [see Eq. (11)].

For a light field (9.1) switched on abruptly (or for irradiation with a square laser pulse of long duration), the experimental check of the formula (9.13) for the characteristic time τ_0 of the photoinduced transition into a new state of dynamic equilibrium can be performed under ordinary conditions. Specifically, the behavior of a vanadium dioxide film on an aluminum substrate irradiated with a laser pulse with intensity $I\approx7\cdot10^8$ W/cm^2 and duration $6\cdot10^{-12}$ s was studied experimentally in Ref. 35. It was found that when the energy of the exciting photons $\hbar\omega_0=1.17$ eV, VO_2 passes from the semiconductor into the metallic state in time $\tau_0\sim10^{-12}$ s after the onset of the pulse. After this, the metallic phase remains stable for a quite long time t ($t>10^{-9}$ s). If the energy of the exciting photons $\hbar\omega_0=2.34$ eV, the semiconductor state of vanadium dioxide becomes unstable $t\sim10^{-9}$ s after the onset of irradiation.

To explain the experimentally observed phenomenon described above, we shall make a numerical estimate of the maximum possible temperature change ΔT in the sample under the action of the laser irradiation. For this, we shall examine a very thin (~1 nm) most strongly heated region of the film near the surface, neglecting heat transfer and assuming that there is enough time for all of the absorbed radiation energy to be converted into heat (at short times $t<10^{-12}$ s this assumption is quite nominal). Using the approximate formula

$$\Delta T=\frac{\alpha I\Delta t}{c\rho}, \tag{11.1}$$

where α, c, and ρ are, respectively, the optical absorption coefficient, the specific heat, and the density of the fill material, and Δt is the irradiation time, and taking the numerical values of the parameters characteristic for VO_2 $\alpha\sim10^5$ cm^{-1},[4] $c\sim1$ J/gK, $\rho\approx10$ g/cm^3, $\Delta t=\tau_0\sim10^{-12}$ s, we obtain $\Delta T\approx10$ K. Since the temperature of the thermodynamically equilibrium semiconductor–metal phase transition in vanadium dioxide is $T_0=340$ K, we conclude that at room temperature a photoinduced transition into the metallic

phase, occurring in a time $\tau_0 \sim 10^{-12}$ s, cannot be explained from the standpoint of a thermal mechanism.

At the same time, the stability of the metallic phase of vanadium dioxide after the passage of the entire pulse ($\Delta t \sim 6 \times 10^{-12}$ s) appears to be due to the quite high temperature T of the film ($T > T_0 = 340$ K), since in accordance with Eq. (11.1) $\Delta T \approx 60$ K in this case.

We note that the time delay $t \sim 10^{-9}$ s, observed with energy of exciting photons $\hbar\omega_0 = 2.34$ eV, in the photoinduced semiconductor–metal phase transition from our standpoint is due to the long lifetime τ_π of nonequilibrium electrons and holes in the π^* band,[35,4] which delays by a corresponding time the process leading to heating of the film. In contrast to Ref. 35, we assume that in this case the photoinduced transition into the metallic phase is due to a purely thermal mechanism, since $\tau_\pi \sim 10^{-9}$ s $\gg \tau \sim 10^{-11}$ s,[38] and for this reason the density of nonequilibrium electron–hole pairs in the d band is negligibly small ($n \ll n_c$).

We shall interpret the experimental results of Ref. 35 on the photoinduced semiconductor–metal phase transition, occurring in a characteristic time $\tau_0 \sim 10^{-12}$ s, from the standpoint of the theory developed in the present paper for the electron–phonon instability of the semiconductor phase of a Peierls system. We note first that the intensity of the laser pulse in the experiment of Ref. 35 $I \approx 7 \times 10^8$ W/cm^2 exceeds the critical value I_c (8.11) and therefore, as shown in Secs. 8–10, a transition should occur into the metallic phase, which was in fact observed in Ref. 35. The experimentally measured transition time $\tau_0 \sim 10^{-12}$ s is close to the value calculated on the basis of the present theory (see the numerical estimate for Eq. (9.13) at the end of Sec. 9).

12. CONCLUSIONS

In summary, in the present paper the dynamics of a nonstationary thermodynamically nonequilibrium photoinduced phase transition in a one-dimensional system with a Peierls instability was investigated under the assumption that the characteristic interband electronic relaxation time τ and the phonon relaxation time τ_p are constant and greatly exceed the intraband electron relaxation time τ_e.

It was shown that when the system is irradiated with light of intensity $I < I_c$ (8.9) the Peierls semiconductor passes into a new semiconductor state with a smaller band gap in the electronic spectrum. This transition is an aperiodic relaxation process [for $I < I_1$ (10.8)] or it has the form of damped oscillations (for $I_1 < I < I_c$).

For irradiation intensities $I > I_c$ the system passes into the metallic phase, and the slow (with characteristic time $\tau \sim 10^{-11}$ s) interband relaxation of electrons becomes rapid (with characteristic time $\tau \sim 10^{-14}$ s) intraband relaxation. If in the process the temperature of the sample increases above the critical temperature T_0 of the thermodynamically equilibrium metal–semiconductor phase transition, then the metallic phase is stable. In the opposite case, the question of the stability of the metallic state requires additional analysis, taking account of the deviation of the electron energy distribution in the conduction band from the quasi-Fermi distribution, and is not studied in the present paper.

*)E-mail: semenov@quant.univ.simbirsk.su

[1] L. N. Bulaevskiĭ, Usp. Fiz. Nauk **115**, 263 (1975) [Sov. Phys. Usp. **18**, 131 (1975)].
[2] A. L. Semenov, Fiz. Tverd. Tela (St. Petersburg) **39**, 925 (1997) [Phys. Solid State **39**, 826 (1997)].
[3] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Vestn. Mosk. Universiteta, Ser. 3, Fiz., Astron. **30**, 52 (1989).
[4] A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *Metal–Semiconductor Phase Transition and Its Applications* (Nauka, Lenigrad, 1979), p. 28.
[5] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Fiz. Tverd. Tela (Leningrad) **31**, 261 (1989) [Sov. Phys. Solid State **31**, 1803 (1989)].
[6] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Vestn. Mosk. Universiteta, Ser. 3, Fiz., Astron. **31**, 99 (1990).
[7] A. L. Semenov, Fiz. Tverd. Tela (St. Petersburg) **36**, 1974 (1994) [Phys. Solid State **36**, 1079 (1994)].
[8] A. L. Semenov and S. V. Sukhov, Izv. Ross. Akad. Nauk, Ser. Fiz. **6**, 120 (1996).
[9] I. V. Krive and A. S. Rozhavskiĭ, Zh. Éksp. Teor. Fiz. **81**, 1811 (1981) [Sov. Phys. JETP **54**, 959 (1981)].
[10] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Vestn. Mosk. Universiteta, Ser. 3, Fiz., Astron. **29**, 98 (1988).
[11] V. I. Emel'yanov, N. L. Levshin, S. Yu. Poroĭkov, and A. L. Semenov, Vestn. Mosk. Universiteta, Ser. 3, Fiz., Astron. **32**, 63 (1991).
[12] V. I. Emel'yanov and A. L. Semenov, Fiz. Tverd. Tela (Leningrad) **32**, 3083 (1990) [Sov. Phys. Solid State **32**, 1790 (1990)].
[13] S. N. Artemenko, A. F. Volkov, and S. V. Zaĭtsev-Zotov, Usp. Fiz. Nauk **166**, 434 (1996).
[14] G. Gruner, Rev. Mod. Phys. **60**, 1129 (1988).
[15] S. A. Brazovskiĭ and I. E. Dzyaloshinskiĭ, Zh. Éksp. Teor. Fiz. **71**, 2338 (1976) [Sov. Phys. JETP **44**, 1233 (1976)].
[16] L. P. Gor'kov and E. N. Dolgov, Zh. Éksp. Teor. Fiz. **77**, 396 (1979) [Sov. Phys. JETP **50**, 203 (1979)].
[17] S. N. Artemenko and A. F. Volkov, Zh. Éksp. Teor. Fiz. **80**, 2018 (1981) [Sov. Phys. JETP **53**, 1050 (1981)].
[18] S. N. Artemenko and A. F. Volkov, Zh. Éksp. Teor. Fiz. **81**, 1872 (1981) [Sov. Phys. JETP **54**, 992 (1981)].
[19] K. F. Berggren and B. A. Huberman, Phys. Rev. B **18**, 3369 (1978).
[20] V. F. Elesin, V. V. Kapaev, and Yu. V. Kopaev, Zh. Éksp. Teor. Fiz. **71**, 714 (1976) [Sov. Phys. JETP **44**, 375 (1976)].
[21] V. V. Kapaev, Yu. V. Kopaev, and S. N. Molotkov, Mikroelektronika **12**, 499 (1983).
[22] A. L. Semenov, Fiz. Tverd. Tela (St. Petersburg) **40**, 2113 (1998) [Phys. Solid State **40**, 1915 (1998)].
[23] A. L. Semenov, Zh. Éksp. Teor. Fiz. **114**, 1407 (1998) [JETP **87**, 764 (1998)].
[24] A. L. Semenov, Zh. Éksp. Teor. Fiz. **111**, 1398 (1997) [JETP **84**, 774 (1997)].
[25] A. L. Semenov, Zh. Éksp. Teor. Fiz. **111**, 2147 (1997) [JETP **84**, 1171 (1997)].
[26] H. Gibbs, *Optical Bistability: Controlling Light with Light* (Academic Press, New York, 1985) [Russian translation, Mir, Moscow, 1988, p. 108].
[27] Yu. V. Kopaev, V. V. Menyaĭlenko, and S. N. Molotkov, Zh. Éksp. Teor. Fiz. **89**, 1404 (1985) [Sov. Phys. JETP **62**, 813 (1985)].
[28] I. M. Suslov, JETP Lett. **39**, 670 (1984).
[29] V. B. Stopachinskiĭ and I. M. Suslov, Zh. Éksp. Teor. Fiz. **91**, 314 (1986) [Sov. Phys. JETP **64**, 143 (1986)].
[30] V. I. Emel'yanov and I. F. Uvarova, Zh. Éksp. Teor. Fiz. **94**, 255 (1988) [Sov. Phys. JETP **67**, 1663 (1988)].
[31] V. I. Emel'yanov and D. V. Babak, Laser Phys. **7**, 514 (1997).
[32] R. F. Mamin, Zh. Éksp. Teor. Fiz. **111**, 1465 (1997) [JETP **84**, 808 (1997)].
[33] S. A. Akhmanov, V. I. Emel'yanov, N. I. Koroteev, and V. N. Seminogov, Usp. Fiz. Nauk **147**, 675 (1985) [Sov. Phys. Usp. **28**, 1084 (1985)].
[34] E. Schöll, *Self-Organization in Semiconductors* (Russian translation, Mir, Moscow, 1991).
[35] A. A. Bugaev, V. V. Gudyalis, V. P. Zakharchenya *et al.*, JETP Lett. **34**, 430 (1981).
[36] O. Madelung, *Solid-State Physics: Localized States* (Russian translation, Nauka, Moscow, 1985, p. 144).
[37] N. N. Bogolyubov and N. N. Bogolyubov, Jr., *Introduction to Quantum Statistical Mechanics* (Nauka, Moscow, 1984), p. 282.
[38] N. R. Belashenkov, V. B. Karasev, A. A. Solunin *et al.*, Fiz. Tverd. Tela

(St. Petersburg) **36**, 2475 (1994) [Phys. Solid State **36**, 1347 (1994)].
[39] V. L. Bonch-Bruevich and S. G. Kalashnikov, *The Physics of Semiconductors* (Nauka, Moscow, 1977), p. 255.
[40] E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics* (Pergamon Press, Oxford, 1981) [Russian original, Nauka, Moscow, 1979, p. 342].
[41] N. N. Bogolyubov, Jr., B. I. Sadovnikov, and A. S. Shumovskiĭ, *Mathematical Methods of the Statistical Mechanics of Model Systems* (Nauka, Moscow, 1989), p. 180.
[42] A. I. Akhiezer and S. V. Peletminskiĭ, *Methods of Statistical Physics* (Nauka, Moscow, 1977), p. 227.
[43] E. M. Lifshitz and L. P. Pitaevskiĭ, *Statistical Physics* (Pergamon Press, New York) [Russian original, Nauka, Moscow, 1978, Part 2, p. 65].
[44] S. A. Akhmanov, Yu. E. D'yakov, and A. S. Chirkin, *Introduction to Statistical Radio Physics and Optics* (Nauka, Moscow, 1981), p. 42.
[45] D. N. Klyshko, *The Physical Principles of Quantum Electronics* (Nauka, Moscow, 1986), p. 22.
[46] A. S. Davydov, *Solid-State Physics* (Nauka, Moscow, 1976), p. 296.
[47] A. V. Andreev, V. I. Emel'yanov, and Yu. A. Il'inskiĭ, *Cooperative Phenomena in Optics: Superradiance, Bistability, Phase Transitions* (Nauka, Moscow, 1988), p. 256.
[48] Yu. E. Kuzovlev, T. K. Soboleva, and A. É. Filipov, Zh. Éksp. Teor. Fiz. **103**, 1742 (1993) [JETP **76**, 858 (1993)].

Translated by M. E. Alferieff

Nature of the Mesyats effect—pulsed superheating of microscopic spikes on a cathode

A. A. Valuev and G. É. Norman*)

Moscow Physicotechnical Institute, 141700 Dolgoprudny, Moscow Region, Russia
(Submitted 25 June 1999)
Zh. Éksp. Teor. Fiz. **116**, 2176–2181 (December 1999)

A study is made of an effect observed experimentally by Mesyats which involves, prior to the electrical explosion, as such, the accumulation of an energy on the order of a few times the sublimation energy by a microscopic cathode spike during explosive emission from a cathode in a vacuum or gaseous discharge. The same effect is observed during electrical explosion of a wire. Simple estimates by various authors imply that the temperature of the wire should rise to 10^5 K. In reality, when energy is applied very rapidly the wire cannot expand and it is superheated into a metastable state (essentially to the crystal-liquid spinodal). When the temperature rises above 10^4 K, the specific heat of the metal increases as electronic degrees of freedom are unfrozen. Thus, the temperature attained prior to an electrical explosion does not exceed 17000 K. © *1999 American Institute of Physics.* [S1063-7761(99)02312-4]

1. INTRODUCTION

Mesyats[1] has developed some new ideas on the explosive character of emission and introduced the concept of ectons, or explosion microcenters. It has been shown that a variety of experimental data on cathode processes can be explained by making the simple hypothesis that, before they explode, microscopic cathode spikes absorb an energy equal to several times the sublimation energy. This can be interpreted as superheating of the microscopic spikes to temperatures on the order of 10^5 K if the simplest equation of state is used. Because of the paradoxical nature of this conclusion, special experiments have been done on the electrical explosion of thin wires.[2] The results were in complete agreement with Mesyats' hypothesis. Here the effect was enhanced when the rate of energy input was raised. The maximum superheating was obtained at the highest rate of energy input, 10^{12} W/g.

Heating of wires to temperatures of $\approx 10^5$ K during electrical explosion was noted in the early experiments of Kvartskhava, Plyutto, *et al.*[3] and later by Tucker.[4] (See the review by Kotov, *et al.*[5]) No convincing explanation of this effect, however, has been given.

The transition of a solid object into a superheated metastable state was noted by Urlin[6] in an analysis of experiments on powerful shock waves in solids. The transition into a metastable state during rapid ($\approx 10^{-6}$ s) Joule heating has been studied by Baĭkov and Shestak.[7] For a copper wire with a diameter of $3 \cdot 10^{-2}$ cm, a superheating $\vartheta = (T - T_m)/T_m \approx 0.13$ was observed, where T_m is the melting temperature. The kinetics of the bulk melting of metals has been studied theoretically by Motorin and Musher.[8] Their estimate of a maximum superheating ϑ of up to 20% or, in absolute numbers, $T - T_m = 150 - 200$ K, was in agreement with the earlier work.[7] Computer simulations of metastable superheated infinite crystalline solids have been done by molecular dynamics and Monte Carlo methods.[9–14] The results of Refs. 7–14 were not drawn upon in the analysis of high temperature superheating in Refs. 1–6.

In this paper we show that the experimental data of Mesyats, *et al.*[1,2] have an entirely rational explanation. The Mesyats effect occurs when the rate of energy input is so high that a wire cannot expand significantly. The high magnetic pressure at high current densities also inhibits expansion. The possible heating regimes for copper are examined in Sec. 2. Two extreme scenarios for this process are considered: constant pressure and constant volume. The thermal and caloric equations of state[15] are used in analyzing both regimes. The final temperatures were less than 17000 K in both scenarios. Temperatures below 17000 K correspond to a liquid state of copper if equilibrium melting is assumed. The possibility of maintaining copper in a metastable solid state during rapid heating[1,2] is examined in Section 3, drawing on the data of Refs. 7–14. It is shown that a wire can enter a superheated metastable state and continue to absorb energy, while remaining at near its initial specific volume. Even if melting sets in later, an increase in the volume (radius) of the wire may not be noticed experimentally prior to an electrical explosion (sudden increase in the volume and electrical resistivity of the wire). The heating processes at constant pressure and volume are compared in Section 4.

2. HEATING REGIMES

As an example of an estimate of the final state of a wire immediately prior to an electrical explosion, we have chosen one of the intense energy input regimes described in Ref. 2 for copper. In particular, the radius of the wire is $r = 3 \times 10^{-3}$ cm, the current density is $j \approx 10^8$ A/cm^2, and the energy input prior to an electrical explosion is roughly 9 kJ/g. Then the magnetic pressure $P_M = \mu_0 \mu (jr)^2/8$ ($\mu_0 \mu$ is the magnetic permeability) at the boundary of the wire is ~ 0.1 Mbar.

Since the data given in Ref. 2 are inadequate for an accurate determination of the heating dynamics up to the

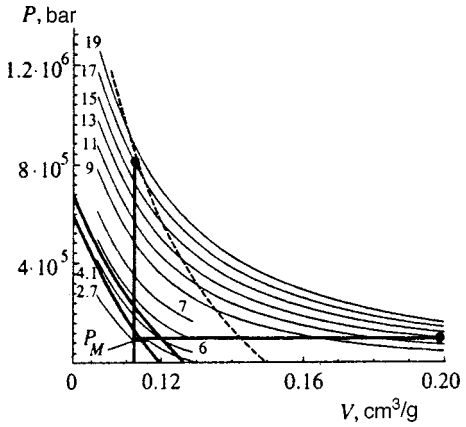

FIG. 1. Pulsed heating regimes. The vertical to P_M plus the horizontal line correspond to scenario A, and plus the vertical line, to regime B; the points denote the final states preceding an electrical explosion. The thermal equation of state for copper: the numbers on the isotherms denote the temperature in units of 10^3 K, the thick lines are the solid-liquid binodal, and the dashed curve is the estimated spinodal for the superheated solid.

time of the explosion, here we consider two limiting scenarios for the process:

(A) heating along an isochore until a pressure P_M is reached, followed by expansion along an isobar until the time of an explosion. This scenario initially seems more natural, so it will not be justified further, and

(B) heating along an isochore leading to a superheated, metastable state of the solid. Arguments in favor of this scenario will be presented in the next section. This scenario, as well as the scenarios intermediate between A and B are only possible as a result of rapid heating and the inertia of expansion.

The final states are determined graphically from diagrams for the energy-volume $(E-V)$, enthalpy-volume $(I-V)$, and pressure-volume $(P-V)$ isotherms constructed in accordance with the equations of state of copper.[15] For example, for an isochore on the $E-V$ diagram, we seek the point of intersection of the horizontal and vertical lines corresponding to the known experimental values of the specific internal energy and volume. The temperature at the intersection point is determined by interpolation from tabular data surrounding this point. The final state (pressure) corresponds to the intersection of the experimental isochore and the isotherm found on the $P-V$ diagram.

Both heating regimes are represented in Fig. 1, together with the equation of state for copper[6,15–17] and our estimates of the spinodal for solid copper.

3. SUPERHEATED CRYSTALS

Superheated inert gas crystals have been studied by computer simulation using Monte Carlo[9–11] and molecular dynamics[12,13] methods. Calculations were done along the isotherm[9,10] and along the isobar.[12,13] Calculations along the isotherm have been done for solids up to the range of parameters where the derivative of the pressure with respect to volume, $(\partial P/\partial V)_T$, becomes greater than zero. The point where $(\partial P/\partial V)_T=0$ is the point where the spinodal crystal-liquid phase transition takes place. The specific volume of the crystal at the spinodal exceeded that on the binodal by 25%.[10] Note that the difference in specific volumes at the binodal between the crystal and liquid was only 10% in the calculations of Ref. 10.

The spinodal of superheated crystalline copper was plotted for a 25% excess. Note that, for a model crystal formed by charges of the same sign against a compensating background, Monte Carlo calculations show that the spinodal is shifted relative to the binodal by 70% in volume for constant temperature or by 20% in temperature for constant volume.[14] A similar shift in the temperature at the isobars of inert gas crystals has been obtained by molecular dynamics methods[12,13] using a mechanical stability criterion.

Thus, the results of computer simulations obtained by different authors using different schemes differ little from one another. They are also in agreement with analytic estimates.[8] Therefore, we may expect that the estimated position of the spinodal in the figure is suitable as a starting point for further derivations. The fundamental difficulty lies elsewhere.

In fact, since the formation of the liquid phase at the surface of a melting object does not involve the expenditure of work to form the new surface, superheating of a solid is, in general, impossible.[18] It has, however, been pointed out[18] that the situation changes if the body is heated from within and its surface is kept at a temperature below the melting point. This situation is close to the experimental conditions in Refs. 1 and 2, where the magnetic pressure (and, therefore, the melting point for most metals) increases at the surface of the sample, the surface is cooled by radiation, there is no skin effect, and heating is uniform throughout the entire volume of the wire. Note that the large amount of experimental data on the melting of metals by high power laser radiation have nothing to do with the theme of this paper, since melting takes place within a thin surface skin layer in those experiments. Here, as in Refs. 1–8, we are concerned with volume melting. Note that the computer simulations of Refs. 9–14 involve uniformly heated infinite crystals (without an open surface).

With heating from inside, the probability of forming nuclei of the liquid phase depends on elastic deformations accompanying the appearance of droplets of liquid inside the solid.[18] It has been shown,[8] however, that the energy of elastic deformations makes a significant contribution to the free energy of nucleation only for small amounts of superheating. Then the nucleation rate per unit volume, ω, depends on the relaxation of elastic stresses and is determined by the diffusion time for the vacancies formed during nucleation. Here the rate ω is low, and, for pulsed energy input into a small volume, cannot reach the critical level (where the nucleation probability ~ 1) while satisfying the requirement of a small amount of superheating. The crystal melts with pulsed heating only for large amounts of superheating, when the elastic deformation of the crystal lattice can be neglected. In this case, the rate of growth of microscopic droplets is determined by heat conduction and ω is substantially higher than for small amounts of superheating.

Therefore, we can expect that melting sets in only for large amounts of superheating, and that the nucleation prob-

ability is highest on the axis of the conductor, where there is no magnetic pressure. Even if a nucleus is formed, the propagation time of the melting front is extremely long compared to the energy input time in this case. The velocity of the melting front has been calculated for copper by the molecular dynamics method.[19] It was up to 100 m/s for 20% superheating. Thus, over the characteristic times $\sim 10^{-8}$ s for the Mesyats effect, the melting front can only move a distance $10^{-2}R$.

An additional argument in favor of scenario B is the fact that, for these energy inputs, the exploding wire effect is natural, as the destruction of a metastable state of a solid near the spinodal when the system is perturbed. On the other hand, in the case of scenario A, it is not clear why an explosion should happen, when a wire that is in an equilibrium liquid state, still far from the liquid-vapor binodal, still has a metallic conductivity. Note that the effect of the shift in the phase equilibrium for a liquid-vapor phase transition examined in Ref. 20, and caused by an inhomogeneity of the magnetic pressure in the different phases, is far weaker for a solid-liquid transition owing to the much smaller difference in the conductivity of the phases.

4. DISCUSSION

In this paper we have only made qualitative estimates. For example, the pressure is assumed constant over the radius of the wire, the time dependence of the energy input is neglected, and the difference between the equilibrium and metastable isotherms is neglected, both in the thermal and the caloric equations of state, although the energies in these states are different. The specific heat of the metastable state increases near the spinodal, as noted in the case of laser heating of superheated liquids.[21]

The figure shows a $P-V$ diagram for copper in the parameter region corresponding to the onset of an electrical explosion. Initially, scenarios A and B are indicated by a single vertical line, and then by horizontal and vertical lines, respectively. In these estimates we have used caloric equations of state for the internal energy and enthalpy[15] and the melting curve of Ref. 6. This curve is consistent with contemporary data[16,17] at the triple point.

It is clear that the final temperature in both scenarios is less than 17000 K. This resolves the paradox mentioned in the introduction. This reduction in the temperature is so low compared to the estimates based on the sublimation energy because, at temperatures of $\sim 10^4$ K, electrons begin to make a significant contribution to the specific heat. The outer shells of the copper atom have a $d^{10}s^1$ structure, so the electronic structure of crystalline copper has six Brillouin zones: a half filled s-zone and five filled d-zones. The energy levels of all the d-zones lie within 1 eV of the Fermi energy. Thus, at temperatures on the order of 10^4 K, all 11 electrons begin to contribute to the electronic specific heat, which, of course, has been taken into account in the data[15] we have used. The earlier estimates[1–5] of superheating to 10^5 K were a consequence of an incorrect extrapolation of the low temperature equations of state to high temperatures.

Note that electron emission causes cooling of the electrons in the metal. (This fact was pointed out to us by S. Barengol'ts.) Since the electrons make a substantial contribution to the specific heat under these conditions, electron emission not only cools the electrons, but also causes additional cooling of the metal, itself, i.e., of the surface of the microscopic spikes on the cathode, which, as noted above, is important for the existence of a superheated crystal.

The end point of scenario B lies near the estimated spinodal for the solid. Thus, it is more or less clear why electrical explosions occur in scenarios close to B (the entry into a superheated state of the solid) at these energy input levels. On the other hand, is not at all clear why an electrical explosion (assuming equilibrium melting) should take place in scenario A at these energy input levels. We note, again, that the end point for this scenario is still far from the binodal of the liquid-vapor phase transition.

A more exact calculation of the initial stage of the electrical explosion process will have to avoid the approximations mentioned above.

We thank G. A. Mesyats, whose brought our attention to this problem, S. A. Barengol'ts for discussing the results, and K. V. Khishchenko who kindly provided the equation of state for copper.

*)E-mail: henry@aha.ru

[1] G. A. Mesyats, *Ectons* [in Russian], Ural'skaya izd. firma Nauka, Ekaterinburg (1993); Usp. Fiz. Nauk **165**, 601 (1995); G. A. Mesyats, *Explosive Electron Emission*, Ekaterinburg, URO-Press (1998).
[2] L. I. Chemezova, G. A. Mesyats, V. S. Sedoi, B. N. Semin, and V. V. Valevich, *Proc. XVIII Int. Symp. Disch. and Elec. Insul. in Vacuum*, Eindhoven (1998), p. 48.
[3] I. F. Kvartskhava, A. A. Plyutto, A. A. Chernov, and V. V. Bondarenko, Zh. Éksp. Teor. Fiz. **30**, 42 (1956) [Sov. Phys. JETP **3**, 40 (1956)].
[4] T. J. Tucker, J. Appl. Phys. **32**, 1894 (1961).
[5] Yu. A. Kotov, V. S. Sedoĭ, and L. I. Chemezova, Preprint No. 41, ISÉ, Tomsk (1986).
[6] V. D. Urlin, Zh. Éksp. Teor. Fiz. **49**, 485 (1965) [Sov. Phys. JETP **22**, 341 (1965)]; V. D. Urlin and R. F. Trunin, ed., in *Properties of Condensed Materials at High Pressures and Temperatures* [in Russian], MAÉP, Moscow (1992), p. 377.
[7] A. P. Baĭkov and A. F. Shestak, Pis'ma Zh. Tekh. Fiz. **5**, 1335 (1979) [Sov. Tech. Phys. Lett. **5**, 562 (1979)].
[8] V. I. Motorin and S. L. Musher, J. Chem. Phys. **81**, 465 (1984).
[9] M. Ross and B. Alder, Phys. Rev. Lett. **16**, 1077 (1966).
[10] V. G. Baĭdakov, S. P. Protsenko, and V. P. Skripov, Trudy MOPI, Izd-vo. vysshcheĭ shkoly, Moscow (1976).
[11] V. M. Zamalin, G. É. Norman, and V. S. Filinov, *Monte Carlo Methods in Statistical Thermodynamics* [in Russian], Nauka, Moscow (1977), p. 88.
[12] V. G. Baĭdakov, A. E. Galashev, and V. P. Skripov, Fiz. Tekhn. Poluprovodn. **22**, 2681 (1980) [*sic*].
[13] J. Solca, A. J. Dyson, G. Steinebrunner, B. Kirchner, and H. Huber, J. Chem. Phys. **108**, 4107 (1998).
[14] H. E. DeWitt and W. L. Slattery, in *Strongly Coupled Coulomb Systems*, ed. by G. Kalman, M. Rommel, K. Blagoev, New York, Plenum Press (1998), p. 1.
[15] V. E. Fortov, K. V. Khishchenko, P. R. Levashov, and I. V. Lomonosov, Nucl. Instrum. Methods Phys. Res. A **415**, 604 (1998).
[16] E. Yu. Tonkov, *Phase Transitions of Compounds at High Pressures* [in Russian], Vol. 1, Metallurgiya, Moscow (1988).
[17] V. I. Severin, Yu. A. Priselkov, A. V. Tseplayeva *et al.*, Teplofiz. Vys. Temp. **36**, 577 (1998).

[18] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part 1, L. P. Pitaevskii, ed., (1995), p. 601.
[19] S. R. Phillpot, J. F. Lutsko, D. Wolf, and S. Yip, Phys. Rev. B **40**, 2841 (1989).
[20] V. S. Vorob'ev and S. P. Malyshenko, Phys. Rev. E **56**, 3959 (1997).
[21] A. A. Samokhin, Trudy IOFAN **13**, 3 (1988).

Translated by D. H. McNeill

Magnetic properties of granular ferromagnetic materials

E. Z. Meĭlikhov*)

Institute of Molecular Physics, Russian Scientific Center ''Kurchatov Institute,'' 123182 Moscow, Russia
(Submitted 5 July 1999)
Zh. Éksp. Teor. Fiz. **116**, 2182–2191 (December 1999)

Granular ferromagnetic metals are usually regarded as systems with weakly interacting superparamagnetic particles whose magnetization is described by a Langevin model. It is shown that this approach is inadmissible. A correct description of the magnetic properties of granular ferromagnetic metals requires that the magnetic anisotropy of the granules, the spread in their orientations, and the variety of their shapes be taken into account. A model with magnetically anisotropic granules, as opposed to the Langevin model, predicts a weak temperature dependence for the magnetization of granular ferromagnetic metals that is in agreement with experiment and provides averaged information on the grain shapes. The glassy nature of the magnetic state of these systems is demonstrated. The magnetic correlations owing to the dipole interaction of the magnetic moments of the granules are examined. © *1999 American Institute of Physics.* [S1063-7761(99)02412-9]

Experiments have shown that the magnetization of the nanocomposite $Fe_x(SiO_2)_{1-x}$ with an iron content $x \approx 0.5$ close to the critical value for a dielectric-metal transition has a very weak dependence on temperature within the range $T = 4-300$ K.[1] Thus, a description of the magnetization of such a system using the Langevin function[2,3] is inadmissible. This means that the conventional model of single domain, superparamagnetic (and noninteracting with one another) Fe-granules is not applicable to this case. There are several reasons for this, but the main reason is the large magnetic energy W_A associated with the (crystalline and geometrical) anisotropy of the granules. The weak temperature dependence of this energy (at least, far from the Curie point) causes the magnetization to be temperature independent. On the other hand, this anisotropy can show up as glassy behavior in granular ferromagnetic metals, especially in the way their magnetization depends on the prehistory and rate of measurements and in magnetization creep, i.e., its dependence on time when the external conditions are constant. This paper is an attempt to examine all these questions in terms of a simple model where the magnetic anisotropy of the granules is related to their nonspherical shape and the distributions of the orientations and shapes of the granules in real systems are taken into account statistically. This paper continues the examination of the magnetic properties of nanocomposites with spherical ferromagnetic granules begun in Ref. 4.

1. MAGNETIZATION OF NANOCOMPOSITES AT LOW TEMPERATURES

We first examine the applicability of the superparamagnetic approximation. The magnetization of this system can be independent of temperature only when the energy W_A required to change the orientation of the magnetic moments of the individual single domain granules is much greater than the thermal energy kT. In the absence of an external magnetic field, the orientation of these moments is determined by the intragranular crystalline anisotropy or by the asymmetric geometric shape of the granules. In these cases, $W_A \sim VK_1$ or $W_A \sim VI_s^2 \nu$, respectively, where V is the granule volume, K_1 is the crystalline anisotropy constant ($K_1 \sim 10^5$ erg/cm^3 for iron), I_s is the saturation magnetization ($I_s \approx 1700$ G for iron), and $\nu \sim 1$ is the difference in the demagnetization coefficients for nonspherical granules. For granules of size $a \sim 10^{-6}$ cm we have $W_A \sim 10^3$ K (for the crystalline anisotropy) and $W_A \sim 10^4$ K (for the geometric anisotropy), which exceeds the experimental temperature.

In order to establish which of the two anisotropy mechanisms really occurs, it is necessary to compare the measured fields $H \sim 10$ kOe at which the magnetization begins to saturate (regardless of the temperature!) with the effective anisotropy field $H_A = W_A / VI_s$, which amounts to $H_A \sim K_1 / I_s \sim 1$ kOe (for a crystalline anisotropy) or $H_A \sim I_s^2 \nu \sim 10$ kOe (for a geometric anisotropy). It is evident that, in this case, the field dependence of the magnetization is determined by the anisotropy associated with the nonspherical shape of the granules.

In the following we shall assume that the Fe-granules are in the shape of prolate ellipsoids of rotation with semiaxes $a > b = c$. In the absence of an external magnetic field the magnetic moment of each single domain granule will be directed along its major axis. An external magnetic field which does not coincide in direction with this axis will tend to turn the magnetic moment of the granule so that its direction approaches that of the field. We define the equilibrium orientation of the magnetic field by the angle θ between its direction and that of the major axis of a granule. Then the anisotropy energy can be written in the form[5]

$$W_A = \frac{1}{2} I_s^2 \nu \sin^2 \theta,$$

where $\nu = N_b - N_a$, and N_b and N_a are the demagnetization

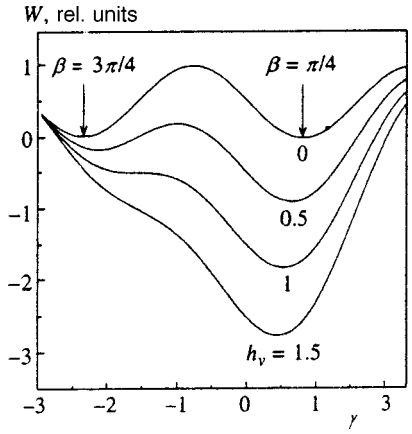

FIG. 1. The magnetic energy W of a single domain ellipsoidal granule as a function of the angle γ between its magnetic moment and the external magnetic field for different values of the reduced field h_ν. The initial ($h_\nu=0$) orientation of the magnetic moment is: $\beta=\pi/4$ (the magnetic moment always ''sits'' in the right hand minimum of the $W(\gamma)$ curve) or $\beta=3\pi/4$ (in a field $h_\nu\approx 1$ the magnetic field ''jumps'' from the left hand minimum to the one on the right).

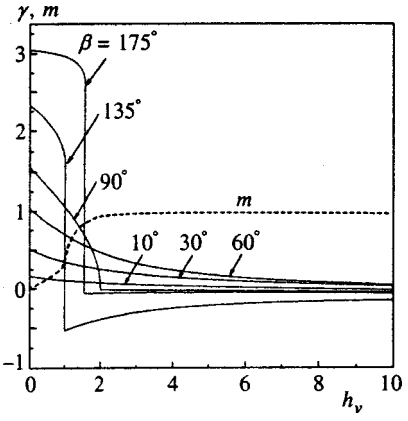

FIG. 2. The angle γ of inclination of the magnetic moment of ellipsoidal single domain granules as a function of the field. The initial direction of the moment is determined by the orientation of the major axis of the ellipsoid (the angle β). The monotonically increasing curve is the total magnetic moment of a system of ellipsoidal granules with a uniform random distribution of the orientation angles β.

coefficients of the ellipsoid along the b and a axes. The total magnetic energy W of the ellipsoid in an external magnetic field is given by

$$\frac{W}{V}=\frac{1}{2}I_s^2\nu\sin^2\theta-HI_s\cos\gamma, \tag{1}$$

where the last term corresponds to the energy of the interaction of the magnetic moment of the granule with the magnetic field and γ is the angle between the magnetic moment of the granule and the magnetic field. Since in equilibrium (and neglecting thermal fluctuations; see below), the magnetic field, major axis of an ellipsoidal granule, and its magnetic moment lie in a single plane, we have $\gamma=\beta-\theta$, where β is the angle between the magnetic field and the major axis of the ellipsoid.

The equilibrium angle $\gamma(H,\beta)$ of orientation of the magnetic moment corresponds to the minimum of the energy (1) and can be found using the equation

$$\frac{\sin 2(\beta-\gamma)}{\sin\gamma}=h_\nu,\quad h_\nu=\frac{2H}{I_s\nu}, \tag{2}$$

which determines the field dependence of this angle for granules with a specified (by the angle β) orientation. It is clear that, as the field increases, the direction of the magnetic moment should approach that of the magnetic field. However, the corresponding functions $\gamma(H,\beta)$ are different in two cases: $|\beta|<\pi/2$ (initial magnetic moment inclined at an acute angle to the ''future'' magnetic field) and $|\beta|>\pi/2$ (the same, but with an obtuse angle). This is related to the presence of two minima in the magnetic energy (1) of a granule in a weak magnetic field (Fig.1). As Fig. 2 shows, in the first case the angle γ approaches zero monotonically and, in the second, with a sudden change of sign corresponding to a ''jump'' of the magnetic moment from one minimum of the magnetic energy (1) to the other.

If the granules had the same shape ($\nu=$const) and their axes were distributed uniformly in space, then the field dependence of the total magnetic moment M of the system (normalized to its saturation value M_s) could be found by simple summation (over the orientations):

$$m_\nu(h_\nu)=\frac{M}{M_s}=\frac{1}{\pi}\int_0^\pi\cos\gamma(h_\nu,\beta)d\beta. \tag{3}$$

The resulting $m_\nu(h_\nu)$ is also shown in Fig. 2. It can be seen that it still preserves a ''memory'' of the sudden ''jumps'' of the magnetic moments in half the granules at fields $h_\nu\sim 1$. Clearly, this behavior of the system magnetization, which is inconsistent with experiment, is related to the simplifying assumption that all the granules are similar ($\nu=$const). In real systems, this condition is not, in general, satisfied, since they consist of granules with different shapes. Assuming for simplicity that all the granules are ellipsoidal, as before, but that the distribution of their volume fractions with respect to the parameter ν has a distribution $\psi(\nu)$, then instead of Eq. (3) we obtain

$$m(H)=\int_{\nu_{\min}}^{\nu_{\max}}m_\nu\frac{2H}{I_s\nu}\psi(\nu)d\nu, \tag{4}$$

where the limits of integration are determined by the distribution of granule shapes.

As an example, let us consider the uniform distribution $\psi(\nu)=(\nu_{\max}-\nu_{\min})^{-1}$ for the case in which the shape of the granules varies from spherical ($\nu_{\min}=0$) to prolate ellipsoidal with an axis ratio of $a/b=10$ ($\nu_{\min}=5.9$).[1)] The $m(H)$ dependence obtained for this case is shown in Fig. 3.[2)] Also shown there is the corresponding experimental dependence (curve *1*) for the magnetization of the $Fe_x(SiO_2)_{1-x}$ ($x\approx 0.5$) system, measured at $T=77$ K (Ref. 1). It is clear that the theoretical dependence is the same as the experimental one everywhere except at intermediate fields $H\sim 10$ kOe, where the calculated magnetizations are somewhat higher (up to $\sim 20\%$) than the experimental values. This model essentially applies to $T=0$. Thus, the agreement obtained here, which is not based on using any fit parameters, indicates that a Langevin model for the thermal fluc-

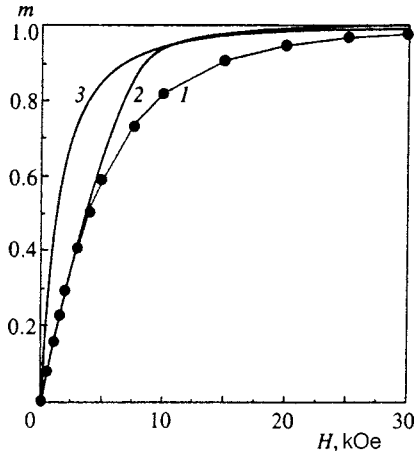

FIG. 3. Field dependences of the magnetization of a granular ferromagnetic material (Fe): *1*—experimental dependence for the dielectric nanocomposite $Fe_x(SiO_2)_{1-x}$ ($x \approx 0.5$, $T=77$ K); *2*—theoretical ''fast'' dependence for a system of ellipsoidal Fe-granules with uniform random distributions of their orientations and axis ratios (see text); *3*—theoretical ''infinitely slow'' (thermodynamic) dependence for a system of ellipsoidal Fe-granules of equal volume with uniform random distributions of their orientations and axis ratios (see text) corresponding to a reduced temperature of $T_1 = kT/VI_s^2 = 0.1$.

tuations of the magnetic moment is not applicable and that the magnetic anisotropy associated with the shape of the granules must be taken into account.

The model considered here can be used to obtain simple formulas which might make it possible to obtain information on the ''average'' granule shape beginning with experimental dependences for the magnetization of the granular metals. For this, we can use the simple solutions of Eq. (1) for the two limiting cases corresponding, respectively, to the initial segment of the magnetization curve (low fields) and to its saturation region (high fields). In the first case, the solution of Eq. (1) has the form

$$\cos\gamma \approx \cos\beta + (1/2)h_\nu \sin^2\beta,$$

so that

$$m_\nu = \frac{1}{\pi}\int_0^\pi \cos\gamma(\beta)\,d\beta = \frac{H}{2I_s\nu},$$

$$m(H) = \int_0^\infty m_\nu(H/\nu)\,\psi(\nu)\,d\nu = \frac{H}{2I_s}\langle\nu^{-1}\rangle.$$

This last equation makes it possible to determine the moment $\langle\nu^{-1}\rangle$ of the distribution. In the second case, the solution of Eq. (1) has the form

$$\cos\gamma \approx 1 - \frac{1}{2h_\nu^2}\sin^2 2\beta,$$

so that

$$m_\nu = 1 - \nu^2\left(\frac{I_s}{4H}\right)^2, \quad m(H) = 1 - \langle\nu^2\rangle\left(\frac{I_s}{4H}\right)^2,$$

which makes it possible to determine the moment $\langle\nu^2\rangle$ of the distribution.

The accuracy of the experiments in the magnetization saturation region, however, is usually not sufficient for a reliable determination of $\langle\nu^2\rangle$. Thus, it makes sense to apply the above recipe only to the initial segment of the experimental curve in Fig. 3, which yields $\langle\nu^{-1}\rangle \approx 0.5$.

As for the discrepancies between the calculated and experimental $m(H)$ curves at intermediate fields, they are related, first, to thermal fluctuations in the magnetic moment and, second, to the fact that, at these fields, a granular metal is a magnetic ''glass.''

2. NANOCOMPOSITES AS MAGNETIC GLASSES

In the Langevin model, the magnetic anisotropy is neglected and the granule energy is described by the second (Zeeman) term of Eq. (1). In this case, the depth VI_sH of the potential well (for the magnetic moment) is large compared to the thermal energy only in strong magnetic fields $H \gg kT/VI_s$. In low fields, this well is shallow, so that the fluctuation thermal spread in the orientation angles γ of the magnetic moments is large, and this leads to a strong temperature dependence for the magnetization in low magnetic fields. A nonspherical granule with a large magnetic anisotropy is another matter. Here, as noted above, two situations are possible: the initial (for $H=0$) magnetic moment of a granule directed along the large axis of an ellipsoid can be inclined to an applied field at an acute or obtuse angle. In the first case ($\beta < \pi/2$, see Fig. 1), that minimum in the orientational dependence of the magnetic energy (1) in which the magnetic moment of the granule ($\gamma \approx \beta$) was initially ''arranged'' is either separated from the symmetric second minimum ($\gamma \approx \beta - \pi$) by a high energy barrier of height $\sim VI_s^2 \gg kT$ (in low magnetic fields) or it is the only minimum (in high magnetic fields). In both cases, it is a single deep minimum with a small effective spread $\Delta(\cos\gamma)$ in the orientations of the magnetic moment owing to thermal fluctuations:

$$\Delta(\cos\gamma) \sim \begin{cases} (kT/VI_s^2)^{1/2} \ll 1, & H \ll I_s \\ (kT/VI_sH)^{1/2} \ll 1, & H \gg I_s \end{cases}. \tag{5}$$

The magnetization of granules of this sort is essentially temperature independent.

In the second case ($\beta > \pi/2$), the situations for low and high magnetic fields are similar to the first case ($\beta < \pi/2$), but there is a region of intermediate fields (the same one where the magnetic moment of a granule jumps from one initial energy minimum into the other) where the two minima in the energy W are separated by a low barrier (whose height goes entirely to zero for some value of H; cf. Fig. 1). Here the initial minimum also becomes shallow. In this range of fields, the granule magnetic moment can no longer be regarded as localized in one of the energy minima. It can be in either with finite probability. The result of a measurement of the magnetization will depend on the time allotted to the measurement. In ''fast'' measurements, the magnetic moment will not be able to undergo a transition from one minimum to the other, and the effective spread in the actual values of $\Delta(\cos\gamma)$ is related only to the initial minimum. In ''slow'' measurements, the magnetic moment is ''smeared out'' (owing to thermal fluctuations) over the two minima.

Thus, the magnetization obtained from fast measurements should differ from that obtained from slow measurements. The characteristic time for the jump in the magnetic moment is

$$\tau(H) \sim \tau_0 \exp[\delta W(H)/kT],$$

where $\tau_0 \sim mc/eH$ is the period of the precession in the magnetic moment and $\delta W(H)$ is the barrier height between the energy minima, which depends on the field.[6] The height of the initial barrier is $\delta W(0) \sim VI_s^2 \sim 10^4 - 10^5$ K and the height of the final barrier (in a magnetic field $H \gg I_s \sim 1$ kOe) is $\delta W(\infty) \sim VHI_s \sim 10^4 - 10^5$ K, so that any measurements in low or high fields are fast. A jump time of $\tau \sim 1$ s is obtained for $\delta W(H) \sim 20kT \ll \delta W(0)$, which occurs only within a narrow range of intermediate fields, $H = 1 - 10$ kOe (see Fig. 1). However, even in this case, the measurements will not really be slow, since when there is a large spread in the orientations, shapes, and volumes of the granules, the condition for slowness at each value of the field will be satisfied only for an isolated group of granules.

This all means that it is almost impossible to actually observe a slow (i.e., thermodynamically equilibrium) field dependence of the magnetization of a nanocomposite. However, it is possible to try to detect magnetization creep (i.e., a slow change in the magnetization with time) in the course of prolonged measurements in a constant field. The magnitude of the possible creep can be estimated by comparing the above fast dependence with the limiting form of the magnetization as a function of field for infinitely slow measurements. In the latter case, the field dependence of the magnetization $m_{\beta\nu}(H)$ of a group of granules whose orientation and shape are characterized by the parameters β and ν is given by

$$m_{\beta\nu}(H) \propto \int_{\gamma=0}^{\pi} \int_{\phi=0}^{2\pi} \exp\left[-\frac{W(\gamma,\phi)}{kT}\right] \cos\gamma d\Omega, \qquad (6)$$

where one of the arguments in Eq. (1) for the granule energy (the angle θ) is replaced, for convenience in taking the integral, by the azimuthal angle of the magnetic moment, ϕ, which is related to the angles θ, γ, and β by $\cos\theta = \cos\gamma\cos\beta + \sin\gamma\sin\beta\cos\phi$ and $d\Omega = \sin\gamma d\gamma d\phi$ is the solid angle. (In the Langevin model, which neglects the magnetic anisotropy, the energy W depends only on γ, so that Eq. (6) reduces to the Langevin function.)

Equation (6) is valid for a system of granules in thermodynamic equilibrium. Averages can be taken over the orientations and shapes of the granules, as before, using Eqs. (3) and (4). Since the effective temperature $T_1 = kT/VI_s^2$ depends on the granule volume V in this case, we have taken $V = \text{const}$ for all the granules in order to simplify the calculations. The infinitely slow field dependence of $m(H)$ obtained in this way for $T_1 = 0.1$ (for $V = 10^{-19}$ cm^3 this corresponds to $T \approx 200$ K) is plotted in Fig. 3 (curve *3*). The above remarks imply that it should not agree with experiment. However, as might be predicted, the thermodynamic $m(H)$ dependence (curve *3*) does lie above the experimental curve everywhere and, for fields $1 < H < 10$ kOe, where the magnetic moment is spread out between the two energy minima in a real experiment, it also lies above the theoretical fast field dependence (curve *2*). The difference between curves *3* and *1*, $\simeq 50\%$ for a field of $H \simeq 5$ kOe, provides an estimate of the change in the magnetization owing to creep (over an infinite time).

3. THE ROLE OF CORRELATIONS IN THE MAGNETIZATION

In intermediate magnetic fields ($H = 1 - 10$ kOe for the system under consideration here), the magnetic anisotropy no longer plays a decisive role, and the ordering influence of the external magnetic field no longer leads to a ''forced'' alignment of the magnetic moments of all the granules along the field. Thus, the magnetic moment of each granule can, more or less, freely align itself along the direction of the local magnetic field produced at that granule by the rest of the granules. This leads to correlation effects which influence the total magnetic moment of the system.[3)]

For simplicity, we shall consider a system consisting of identical granules in the following. Since the shape anisotropy is now unimportant, we can assume them to be identical spheres with magnetic moments μ of different magnitudes. The magnetic moment of the ith granule, μ_i, precesses about the direction of the magnetic field at a frequency $\omega \sim eH_i/m_e c$, where $H_i = H + \delta H_i$ is the sum of the external field and the field δH_i created by all the other granules at the site of the granule being considered. Including the field δH_i means taking the interaction between the magnetic moments of the granules into account. For granules in a dielectric matrix, the most important interaction is the long range dipole interaction, whose energy (on the order of μ^2/l^3, where l is the average distance between granules) can be fully comparable to the thermal energy kT. It has been shown[7] that the dipole interaction among granules consisting of a large number of ''magnetic'' atoms can be calculated by replacing them with effective magnetic moments located at the centers of mass of the granules.

The disorientation angle θ between the directions of the magnetic moments of neighboring granules, which determines the probability of intergranular electron tunnelling, is related to their orientation angles $\gamma_{1,2}$ and $\phi_{1,2}$ by

$$\cos\theta = \cos\gamma_1\cos\gamma_2 + \sin\gamma_1\sin\gamma_2\cos(\phi_1 - \phi_2). \qquad (7)$$

Because of random perturbations of thermal origin, the precessional motion of the magnetic moments of the granules is subject to random interruptions which, in general, change the angles of inclination of the precession ($\gamma_{1,2}$), as well as its phase (the angles $\phi_{1,2}$). Since changing the angle of inclination of a magnetic moment requires a much greater expenditure of energy than changing the phase of the precession, it is natural to assume that the time τ_ϕ over which coherent precession is maintained is much shorter than the time τ_γ over which the angles of inclination are maintained constant.[4)] Thus, assuming that the angles $\phi_{1,2}$ are not correlated, we can average Eq. (7) over a time $\tau_\phi \ll t \ll \tau_\gamma$ and obtain

$$\overline{\cos\theta} = \overline{\cos\gamma_1}\,\overline{\cos\gamma_2}. \qquad (8)$$

Averaging Eq. (8) over times $t \gg \tau_\gamma$ is equivalent to averaging over space (i.e., over different granules):

$$\langle \cos\theta \rangle = \langle \cos\gamma_1 \cos\gamma_2 \rangle. \quad (9)$$

Here $\cos\gamma_{1,2} = m(H) + \delta_{1,2}$, where δ_i is the local fluctuation associated with the difference between the local field and the average (naturally, on the average, $\langle \delta_i \rangle = 0$ over the granules) and $m(H)$ is the reduced (normalized to unity) magnetization of the system.

The correlation of the fluctuations of this type for different granules can be described using the correlator

$$\langle \delta_i \delta_k \rangle = \langle \delta_i^2 \rangle \exp\left(-\frac{r_{ik}}{r_\gamma} \right),$$

where r_γ is the correlation radius and r_{ik} is the distance between granules i and k. It is easy to see that

$$\langle \delta_i^2 \rangle = \langle [\cos\gamma_i - m(H)]^2 \rangle = \langle \cos^2\gamma \rangle - m^2(H).$$

Finally,

$$\langle \cos\theta \rangle = m^2(H) + [\langle \cos^2\gamma \rangle - m^2(H)] \exp(-\ell / r_\gamma), \quad (10)$$

where ℓ is the distance between granules 1 and 2. This implies, in particular, that when there is no correlation between the angles of inclination $\gamma_{1,2}$ ($r_\gamma = 0$), the disorientation angle θ is given simply by $\langle \cos\theta \rangle = m^2(H)$.

The field at the ith granule is $H_i = H + \delta H_i$, where the collective contribution to the external field, $\delta H_i = \mu \Sigma \beta_{ik} \delta_k$, is determined by the ''geometry'' (through the ''geometric'' parameters β_{ik}) and the fluctuations δ_k of the neighboring granules. Since $\langle \cos\gamma_i \rangle = m(H_i)$, we have $\cos\gamma_i = m(H + \delta H_i) \approx m(H) + \delta H_i\, m'(H)$, which implies that

$$\delta_i = \mu m'(H) \sum \beta_{ik} \delta_k, \quad (11)$$

where $m'(H)$ is the derivative of the magnetization as a function of the field.

On the other hand, the fluctuations δ_i (and, thereby, the correlation radius r_γ) are determined by the intensity of the intergranular interaction and the geometry (i.e., the mutual positions of the granules):

$$\delta_i = \sum \lambda_{ik} \delta_k. \quad (12)$$

In the continuum approximation, which is applicable if $r_\gamma \gg \ell$ (otherwise, the correlations of the magnetic moments can be neglected), we can introduce a smooth function of position, $\delta = \delta(\mathbf{r})$, such that $\delta(r_{ik}) = \delta_k$. Then,

$$\delta \approx \delta\left(\sum \lambda_{ik} \right) + (\nabla\delta)\left(\sum \lambda_{ik} r_{ik} \right) + \frac{1}{2}(\nabla^2\delta) \sum \lambda_{ik} r_{ik}^2,$$

where, because of the isotropy of the system, the first two terms in the expansion equal zero. Thus, the spatial variation of $\delta(r)$ obeys the equation

$$\nabla^2\delta = \delta / r_\gamma^2, \quad r_\gamma^2 = \frac{1}{2} \sum \lambda_{ik} r_{ik}^2.$$

The solution of this equation has the form $\delta \propto \exp(-r/r_\gamma)$ and describes the spatial correlation of the magnetic moments of the granules. Equations (11) and (12) imply that $\lambda_{ik} = \mu m'(H) \beta_{ik}$, so that

$$r_\gamma^2 = \frac{1}{2} \mu m'(H) \sum \beta_{ik} r_{ik}^2. \quad (13)$$

The magnetic field dependence of the correlation radius r_γ is, therefore, determined by the derivative of the magnetization as a function of the magnetic field. Since the latter saturates in high fields, an increase in the magnetic field leads to a reduction in the correlation radius, so that in the magnetization saturation regime, the correlations of the magnetic moments of the granules can be neglected.

For estimating the correlation radius we note that $\beta_{12} \sim (1/\ell^3)$, so that $\Sigma \beta_{ik} r_{ik}^2 \sim 1/\ell$. As for the derivative $m'(H)$, Fig. 3 implies that $m'(H) = 10^{-4} - 10^{-5}\ \mathrm{Oe}^{-1}$. In a system consisting of granules with different sizes, the correlation of the magnetic moments is associated mainly with the large granules. They have large magnetic moments, $\mu = (10^3 - 10^4)\mu_B$, and ''align'' the magnetic moments of the smaller surrounding granules to themselves. For an average distance between the large and small granules of $\ell \sim 30$ Å we find $r_\gamma \sim 10^{-7} - 10^{-6}$ cm.

Correlation effects may be important in gigantic magnetic resistance effects in ferromagnetic materials made up of granulated metals.[4,8] A separate paper will be devoted to this question.

The author thanks B. A. Aronzon and V. V. Ryl'kov for useful discussions. This work was supported by the Russian Fund for Fundamental Research (Grant Nos. 96-02-18429-a and 98-02-17412-a) and the PICS-Russian Fund for Fundamental Research program (Grant No. 98-02-2237).

*)E-mail: meilikhov@imp.kiae.ru

1) Electron microscope studies show that the granules of a real composite are nonspherical, and the ratio of their maximum and minimum dimensions can vary by an order of magnitude. The variant examined here must be considered as purely illustrative. Note, however, that calculations for other, not very ''exotic'' distributions $\psi(\nu)$, are qualitatively similar to those shown in the text.

2) The calculated curve ''joins'' the absolute magnitudes of the magnetic field beginning with $I_s = 1700$ G.

3) Correlations in a system of superparamagnetic (Langevin) granules have been discussed in Ref. 2, which we shall follow here.

4) Note, however, that for ferromagnetic granules with large effective moments, the corresponding fluctuations are small.

[1] B. A. Aronzon, D. Yu. Kovalev, A. N. Lagar'kov, E. Z. Meĭlikhov, V. V. Rylkov, M. A. Sedova, N. Negre, M. Goiran, and J. Leontin, JETP Lett. **70**, 90 (1999).
[2] P. Allia, M. Knobel, P. Tiberto, and F. Vinai, Phys. Rev. B **52**, 15398 (1995).
[3] S. Honda, T. Okada, M. Nawate, and M. Tokumoto, Phys. Rev. B **56**, 14566 (1997); E. F. Ferrari, F. C. S. da Silva, and M. Knobel, Phys. Rev. **56**, 6086 (1997).
[4] E. Z. Meĭlikhov, JETP Lett. **69**, 623 (1999).
[5] C. Kittel, Rev. Mod. Phys. **21**, 541 (1949).
[6] W. J. Schuele, S. Shtrikman, and D. Treves, J. Appl. Phys. **36**, 1010 (1965).
[7] D. Altbir, P. Vargas, J. d'Albuquerque e Castro, and U. Raff, Phys. Rev. B **57**, 13604 (1998).
[8] J. Inoue and S. Maekawa, Phys. Rev. B **53**, R11927 (1996).

Translated by D. H. McNeill

Parametric interaction of magnetostatic waves with a nonstationary local pump

G. A. Melkov,*) A. A. Serga, A. N. Slavin,†) V. S. Tiberkevich, A. N. Oleĭnik, and A. V. Bagada

Taras Shevchenko Kiev University, 01033 Kiev, Ukraine
(Submitted 16 February 1999)
Zh. Éksp. Teor. Fiz. **116**, 2192–2211 (December 1999)

A solution is obtained for the general problem of the nonstationary interaction of backward volume magnetostatic waves in films of yttrium-iron garnet with local parametric pumping. In the case of a large pump region, $l \gg \lambda$, where λ is the wavelength of the backward volume magnetostatic waves, the problem reduces to a system of truncated equations for two packets of counter propagating waves. In the opposite case, $l < \lambda$, the exact problem of parametric interactions of the eigenmodes of a ferrite film (both counterpropagating and in the same direction) is solved numerically. Both cases are studied experimentally and good qualitative and quantitative agreement is obtained with the theory. For the first time, the reversal of a wave front and the time reversal of the shape of backward volume magnetostatic wave pulses are observed and a change in the propagation time for the peak of the signal pulse and a reduction in its width owing to pumping are recorded. Two operating regimes are identified for a nonstationary parametric backward volume magnetostatic wave amplifier with local pumping, which differ in the ratio of the duration of the pump pulse to the transit time for the wave through the local pump region, and the effect of the parametric excitation of two-dimensional spin waves on the interaction of backward volume magnetostatic waves with a local nonstationary parametric pump is determined. © *1999 American Institute of Physics.* [S1063-7761(99)02512-3]

1. INTRODUCTION

Wave interactions in solids are widely used to study the physical properties of solids, as well as for creating devices for a wide range of purposes. For example, an optical parametric interaction has been used to amplify waves and to reverse their wave fronts.[1–3] Here the limited spatial and temporal extents of the region where the waves interact with the pump, i.e., its localization and nonstationarity, are of some importance.

In general, every interaction is local, since it actually takes place within a limited region of space. However, localization does not always have a significant effect on the characteristics of the interaction. If we are concerned with a parametric wave interaction or Raman scattering, then localization becomes important only when the dimensions l of the interaction region are comparable to the mean free path of the waves. Then all the parameters of the interaction change, beginning with the threshold for wave excitation and ending with such technical characteristics as the transmission bandpass, which are now determined by the size l of the space (i.e., the active region) where it takes place. In optics, where the mean free path of the waves in fiber optics can be as long as tens of kilometers,[1] parametric pumping is always local and this fact is taken into account in the analysis of nonlinear interactions, as well as in their practical applications.[1–3] In magnetic crystals, especially the ferrites, two-dimensional exchange spin waves with a short mean free path ($\leqslant 10^{-2}$ mm) are most often excited and, for quite understandable reasons, the pump localization is usually not taken into account in the calculations.[4] In addition to two-dimensional spin waves, a rich spectrum of long wavelength, dipole-exchange spin, or (in accordance with the magnetostatic approximation often used to analyze them) magnetostatic waves, can be excited in the ferrites. In high quality ferrite films of yttrium-iron garnet (YIG), the mean free path for the latter is hundreds of times longer and the localization of the pump begins to play an important role. It was first taken into account in parametric excitation of spin waves in Ref. 5, and was, subsequently to Ref. 6, taken into account everywhere.[7–9] Until recently, this accounting was quite limited; only the change in the excitation threshold was included under stationary pump conditions. However, as noted repeatedly before that,[6] the pumping cannot be stationary during parametric amplification of long wavelength oscillations and waves in ferromagnetic materials; its duration must be considerably shorter than the characteristic relaxation times for exchange two-dimensional spin waves. This happens because of the multimode character of the spin-wave spectrum: for a single frequency in the crystal there is an infinite set of eigenmodes and waves with different magnitudes and directions of the wave vector. In order to prevent two-dimensional spin waves from being excited from the thermal level to levels detectable in an experiment, nonstationary pumping is necessary. Using nonstationary pumping makes it possible to operate beyond the threshold for generation of magnetostatic waves, which ensures much higher gains and conversion coefficients for these waves. A first brief discussion of the interaction of travelling magnetostatic waves with local nonstationary pumping, with the problem of the reversal of the

wave front of backward volume magnetostatic waves in YIG films as an example, was given by Gordon, *et al.*[10] It is known that backward volume magnetostatic waves propagate in ferrite films with saturation magnetization along the direction of the superposed magnetization field $\mathbf{H}_0$ and are characterized by opposite directions of the phase and group velocity vectors.[11] We have chosen them for experimental study because backward volume magnetostatic waves are reciprocal with respect to reversal of the direction of the wave vector and can also interact efficiently with longitudinal electromagnetic pumping at twice the frequency.[11] Backward volume magnetostatic waves have all the characteristic properties of magnetostatic waves and all the conclusions drawn here are directly applicable to any type of magnetostatic wave.

This paper is a generalization of Ref. 10 to the case of backward, as well as forward waves. Here we present the results of a theoretical and experimental study of the parametric interaction of incident magnetostatic waves with a local nonstationary, longitudinal pump with frequency ω_p whose variable magnetic field vector $\mathbf{h}_p$ is collinear with a constant magnetic field vector $\mathbf{H}_0$. This interaction causes amplification of a forward wave with frequency ω_1 at the exit from the active region, and an idling backward wave (parametric echo) counter to the forward wave appears at the entrance to the active region with the complex conjugate amplitude and a frequency ω_2 that obeys the condition for a parametric resonance, $\omega_2=\omega_p-\omega_1$. In the case of active regions with very short lengths l, on the order of or less than the magnetostatic wavelength $\lambda\equiv 2\pi/k$, a passing idling wave that obeys the same parametric resonance conditions may appear along with the counterpropagating wave but propagate in the same direction as the forward wave.

In Sec. 2 of this paper we describe the experimental setup. Section 3 contains the results of a theoretical analysis of the interaction of waves with local parametric pumping, including a model based on solving the truncated equations for coupled waves[12–15] that is valid for $l\gg\lambda$. An analytic solution of these equations for nonstationary pumping is obtained in terms of Bessel functions of an imaginary argument. These solutions are used to analyze the dependences of the gain coefficient K_1 for the forward wave and the conversion coefficient K_2 for conversion of the forward into a backward wave on the pump power P_p, pump pulse duration τ_p, and signal duration τ_s, as well as on the mutual position of the pulses. Some nontrivial consequences of the interaction of the incident wave with a nonstationary pump are found: compression of the signal pulse and a change in the propagation time for its peak. In the opposite case of a small active region, $l\leqslant\lambda$, the interactions of co- and counterpropagating magnetostatic waves are analyzed using a numerical solution of the equations of motion for the amplitudes of spin waves in momentum space. The results of an experimental study of the parametric interaction of magnetostatic waves with a local, nonstationary pump are presented in Sec. 4. The data are in good agreement with the theory.

Regarding the significance of these results we can say the following. The solution for the parametric interaction of magnetostatic waves with a local, nonstationary pump obtained here can be used to describe the parametric excitation of any other waves, including light, for which it is only necessary to change the interaction coefficients of the waves; all of the conclusions of this paper remain valid. The interaction of magnetostatic waves with a local parametric pump may be applied in spin-wave magnetoelectronics[16] for amplifying waves and reversing their wave fronts, for correlation processing of signals, and correcting the wave front and shape of pulses, as is done in adaptive optics,[3,17] for echo generation, and, finally, for physical research on nonlinear processes in magnetic and other materials. In fact, spin-wave studies are quite simple; they are done in the well explored and accessible microwave bands with extremely moderate pump powers of $P_p\sim 1$ W. Thus, they provide a convenient model for many nonlinear processes in other media and at other wavelengths, including optical wavelengths, that require complicated apparatus. For various reasons, some optical processes have not yet been observed, such as wave front reversal during first order parametric processes (owing to difficulties in satisfying the conservation laws). In optics, higher order processes, mainly of second order,[3] are studied. These difficulties do not arise when magnetostatic waves are used.[10] They have many other advantages, including the possibility of exciting forward and backward waves (with opposite group and phase velocities) and creating interaction regions with dimensions on the order of, or smaller than, the wavelength, i.e., simultaneous excitation of co- and counterpropagating waves.

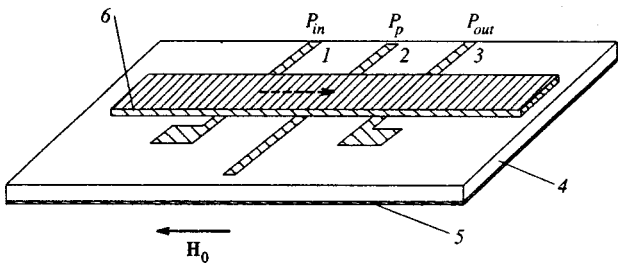

FIG. 1. The experimental model: *1*, *3*—input and output antennas; *2*—inductor for the pump magnetic field; *4*—polycor substrate; *5*—shield; *6*—YIG film on a substrate of gadolinium-gallium garnet. The dashed arrow indicates the direction of propagation of the input signal.

2. DESCRIPTION OF THE EXPERIMENTAL MODEL AND MEASUREMENT TECHNIQUES

In these experiments we used to experimental models with different lengths l of the active region. Model I (Fig. 1) is a structure consisting of three thin metallic conductors in direct contact with a rectangular sample of YIG film on a gadolinium-gallium garnet substrate, a waveguide for backward volume magnetostatic waves. The dimensions of the sample in the plane of the film were 1.6×18 mm^2 and the film thickness was 4.9 μm. The waveguide is oriented with its long side perpendicular to the axes of the conductors and parallel to the direction of the magnetization field $\mathbf{H}_0$. The conductors are made of silver wire with a diameter $d=50\,\mu$m and are placed 3.75 mm apart from one another on the free surface of a shielded polycor (Al_2O_3) plate. The plate is 0.5 mm thick. The outer conductors, which perform

the function of exciting and receiving the magnetostatic waves, serve as antennas for the magnetostatic waves with a separation between the input and output antennas of $L = 7.5$ mm. Antennas of this sort excite waves propagating along the normals to their axes most efficiently. In this case, these waves are backward volume magnetostatic waves with wave numbers $k \leqslant \pi/d$. The middle conductor is the pump inductor and serves to create a longitudinal, high frequency, magnetic pump field $\mathbf{h}_p \| \mathbf{H}_0$ in the film. It is made in the form of a half wave resonator, open at the ends, which is tuned to the pump frequency and has a load Q-factor of $Q \simeq 20$. The small width of the inductor, plus the fact that the resonant frequency of the latter is considerably higher than that of the signal, ensures that it has little effect on the propagation of the backward volume magnetostatic wave when there is no pumping. No reflection of the incident wave from this structure was observed.

During the measurements, microwave pulses with durations of 15 ns or more are incident on the antenna *1* (Fig. 1) and generate a packet of backward volume magnetostatic waves which propagate along the model. When the pump is turned on, a region develops around the inductor *2* in which the pump magnetic field $\mathbf{h}_p$ interacts parametrically with the spin subsystem of the ferromagnetic material. The effective length l of this region along the long side of the waveguide is determined mainly by the diameter d of the conductor and the thickness and dielectric constant of the substrate, and is usually several times d. The exact value of l can be determined by the method to be described in paragraph 10 of Sec. 3 of this paper. For this model, $l \simeq 220\ \mu$m. The amplification and reversal of the spin waves take place within this region (referred to as the active region below).

It was possible to increase the interaction length to 3.3 mm in model II by replacing the microcavity pump inductor (Fig. 1) by an open dielectric resonator. The parameters of the excitation, transmission, and receiver systems for the backward volume magnetostatic waves were identical to those for model I.

Before making the measurements associated the parametric effects, we determined the location of the operating point on the dispersion curve for the backward volume magnetostatic waves. It is known that, for $k = 0$, the frequency ω of the backward volume magnetostatic waves is bounded by the ferromagnetic resonance frequency $\omega_\perp = g\sqrt{(H_0 + 4\pi M)H_0}$ and that for $k \to \infty$, ω approaches a lower bound of $\omega_H = gH_0$. Here M is the saturation magnetization and g is the gyromagnetic ratio for the electron spin. H_0 was chosen so that, even for backward volume magnetostatic waves with frequencies near the upper bound $\omega_\perp$, the band where they exist excluded the possibility of a first order parametric decay. For this, it is enough that $\omega_\perp$ exceed twice the lower bound of the frequency band for the spin waves, ω_H. The upper frequency limit $\omega_\perp$ for the backward volume magnetostatic waves was determined from the amplitude-frequency characteristic of the experimental model and, for $H_0 = 1020$ Oe, it equalled $2\pi \cdot 4775$ MHz. For our film, $4\pi M = 1750$ G, and a calculation yields the above value of the frequency of the ferromagnetic resonance when H_0 is replaced by the effective (superposed) magnetization field $H_0 + 22$ Oe. This substitution is justified by the need to account for the influence of the anisotropy of the actual ferrite sample and corresponds to a rather typical value of the total field of the cubic and uniaxial anisotropies for YIG films grown in the (111) plane, -86 Oe.

In the following, it is precisely the effective magnetization field which was used for determining all the parameters of the backward volume magnetostatic wave. An elementary calculation indicates that here the working frequency $\omega = 2\pi \cdot 4720$ MHz corresponds to a wave number $k = 155\ \text{cm}^{-1}$, a group velocity $v = -2.21\ \text{cm}/\mu\text{s}$, and a delay of $T_L = 339$ ns for propagation from the input to the output antenna. Given the actual weak inhomogeneity of the magnetization field, the calculated delay is in good agreement with the experimental value of 336 ns. The shortest duration of the signal pulses at the input antenna was 15 ns in the experiments. The limited actual transmission bandwidth of the system and dispersion effects distort the backward volume magnetostatic wave pulse shape and transform it from a rectangular to a bell shape.

The width of the ferromagnetic resonance line for our samples was $2\Delta H \simeq 0.4$ Oe. The relaxation parameter for the backward volume magnetostatic waves was $\Gamma = g\Delta H = 3.52 \times 10^6\ \text{s}^{-1}$ and the calculated loss for a backward volume magnetostatic wave over a 1 mm path is 1.6 dB.

The experiment was done in a linear regime with respect to the power of the input signals. The pulsed power at the antenna *1* (Fig. 1) was less than $P_{\text{in}} = 5$ mW.

As a preliminary estimate of the effect of two-dimensional exchange spin waves on our results, we have measured the threshold for parametric generation of these waves. A weak continuous signal with a frequency $\omega \neq \omega_p/2$ was applied to the input of the model. After pumping above the threshold for two-dimensional spin waves was applied, the losses of the backward volume magnetostatic waves increased owing to scattering on two-dimensional spin waves that had been parametrically excited in the active region. Attainment of the threshold was indicated by the appearance of a dip in the continuous output signal during the time of the pump pulse, and the pump duration was increased to 10 μs for low power levels above critical in order to establish this time more precisely. The measurements showed, for example, that for model II the threshold for generation of two-dimensional spin waves is 10 mW for the maximum attainable pump power in our experiment, $P_{p\,\max} = 5$ W.

A word about the measurement technique and accuracy: Times (the signal delay T_L and its width and shift caused by the pump) were measured on a wide band analog oscilloscope. Time intervals shorter than 50 ns were determined with an accuracy of ± 0.5 ns. The relative amplitude of the output signals was measured by an insertion method using a precision polarization attenuator with an accuracy of ± 0.5 dB. Pump power with a pulse duration of from 10 ns to 10 μs at a repetition rate of 300 Hz was also varied using a polarization attenuator with an accuracy of better than 0.1 dB.

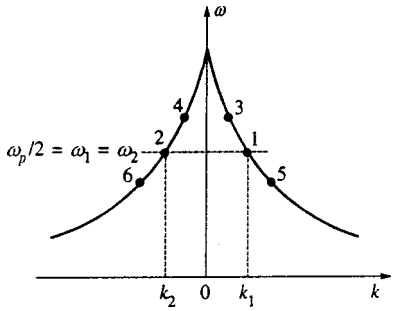

FIG. 2. Frequencies and wave numbers of backward volume magnetostatic waves which interact parametrically through nonstationary local pumping. The smooth curve is the dispersion curve of the backward volume magnetostatic wave.

3. THEORY

1. The equation of motion for the amplitudes c_k of waves travelling along the z axis acted on by a spatially inhomogeneous longitudinal pump can be written in the form[15]

$$\frac{\partial c_k}{\partial t} = -i\omega_k c_k - \Gamma_k c_k + \sum_{k'} h_{k+k'} \exp(-i\omega_p t) V_{kk'} c_{k'}^*, \tag{1}$$

where ω_k and Γ_k are, respectively, the frequency and damping parameter for the spin wave and $V_{kk'}$ is the parametric interaction coefficient for waves with wave vectors k and k'. For the case of parametric amplification of backward volume magnetostatic waves in a longitudinally magnetized YIG film of interest to us,

$$V_{kk'} = \frac{\omega_k \omega_{k'} - (gH_0)^2}{4H_0\sqrt{\omega_k \omega_{k'}}}, \tag{2}$$

where h_k is the kth Fourier harmonic of the nonuniform pump field $h_p(z)$,

$$h_k = \frac{1}{V}\int_V h_p(z)\exp(-ikz)dV, \tag{3}$$

and V is the volume of the ferrite film.

Equation (1) describes the parametric coupling of waves with wave numbers k and k'. The Fourier harmonic $h_{k+k'}$ in Eq. (1) can be regarded as the amplitude of a pump wave with an effective wave vector $k_p = k + k'$. According to Eq. (3), the amplitude of the effective pump wave depends on the scale length l for localization of the pumping. For uniform pumping, $l \to \infty$ and all the Fourier harmonics go to zero, except the uniform one with $k_p = 0$; this implies the possibility of a parametric coupling only for two waves, $k + k' = k_p = 0$, i.e., $k = -k'$. In principle, for nonuniform pumping $h_p(z)$, waves with arbitrary effective k_p exist, but, in reality, their amplitudes differ from zero only for $k_p \leqslant 2\pi/l$. This circumstance leads to a substantial change in the character of the parametric interaction of the waves as the pumping length l is varied. In order to explain this, we turn to the dispersion characteristics of the backward volume magnetostatic waves shown in Fig. 2. With homogeneous pumping, only waves 1 and 2, which propagate counter to one another with k_1 and $k_2 = -k_1$, interact. As l decreases, the number of waves that actually interacts rises, and now, besides 1 and 2, waves 3 and 6, and 4 and 5, for which $|k_{3,5} - k_{4,6}| \leqslant k_p \sim 2\pi/l$, also interact, but the interacting waves are, as before, counterpropagating. Finally, for $2\pi/l \geqslant |k_1| = |k_2|$, the interaction of copropagating waves moving in the same direction (e.g., 3 and 5, or 4 and 6 in Fig. 2) becomes possible. The above condition implies that, because of the inhomogeneity of the pumping, a copropagating interaction is possible only for sizes l of the pumping region that are on the order of the wavelength of the backward volume magnetostatic waves, $\lambda \equiv 2\pi/k$. It is this circumstance which makes it extremely difficult to observe these sorts of effects in nonlinear optics, where they have not even been studied theoretically.[3,14] The presence of a copropagating interaction along with the counterpropagating interaction leads to the propagation, in both the forward and backward directions, of two waves with different phases relative to the pump phase (for example, wave 4 may develop as a result of an interaction with wave 5 and with wave 6). In the experiment, because of the randomness in the pump phase, this leads to beating of the output signals.

It is clear from all of the above that a correct theoretical description of the parametric wave interaction under the influence of nonuniform pumping will require some extremely careful simplifications of the equation of motion (1). For example, the standard approach using truncated equations for the wave packets[12–15,18,19] is valid, strictly speaking, only when the copropagating interaction is included for active regions of length $l \gg \lambda$: only in this case can we assume that the spread of the wave numbers in the packet obeys $\Delta k \ll k$. For $l \leqslant \lambda$, the wave packet approximation is valid only for low pump levels (substantially below critical) in the presence of input signals considerably above the thermal level for the excited waves. In general, for $l \leqslant \lambda$, Eq. (1) can only be solved numerically, which we shall do, despite the tediousness and lack of clarity.

2. First we shall examine the case of long pumping ($l \gg \lambda$). As shown above, then it is possible to use a system of truncated equations for the wave packets.[14,18,19]

Truncated equations are obtained from Eq. (1) by transforming to the z-representation.[15] Assuming that two narrow packets of waves are excited near waves 1 and 2 (see Fig. 2) with wave vectors k_1 and k_2 and frequencies ω_1 and ω_2 such that $k_1 + k_2 = 0$ and $\omega_1 + \omega_2 = \omega_p$, and taking the inverse Fourier transform of Eq. (1), we obtain a system of equations for the slowly varying complex amplitudes (envelopes) of the signal $a_1(z,t)$ and idler $a_2(z,t)$ waves:

$$\begin{aligned} &\left(\frac{\partial}{\partial t} + \Gamma_1 + v_1\frac{\partial}{\partial z}\right) a_1 = h_p V_{1,2} a_2^*, \\ &\left(\frac{\partial}{\partial t} + \Gamma_2 - v_2\frac{\partial}{\partial z}\right) a_2^* = h_p V_{1,2}^* a_1. \end{aligned} \tag{4}$$

For the case of parallel pumping of backward volume magnetostatic waves of interest to us, $\Gamma_1 = \Gamma_{k_1} = \Gamma_2 = \Gamma_{k_2} \equiv \Gamma$, $v_1 = v_{k_1} = v_2 = v_{k_2} \equiv v$ is the group velocity of the waves in the packets, and $V_{1,2} = V_{k_1,k_2} \equiv V$.

The system of Eqs. (4) has been studied in most detail in Ref. 14. It has been solved in the case where the spatial pumping profile has the form of a hyperbolic secant, $h_p(z) = h_{p0}/\cosh(z/l)$. On the other hand, in the case of paramagnetic amplification of magnetostatic waves with a long active region, $l \gg \lambda$ (model II, where the active region is created using an open dielectric resonator), the spatial pumping profile has a shape closer to rectangular. Since the pumping shape has a strong influence on the output signal profile,[14] for best correspondence with the experimental situation we have assumed that the pump field amplitude h_p is locally uniform within an active region $0 \leqslant z \leqslant l$ and equal to zero outside it.

We shall also assume that the pump is applied to the film at time $t=0$ and that it has a rectangular shape of duration τ_p. Signal and idler waves with arbitrary shapes can be incident on the left ($z \leqslant 0$) and right ($z \geqslant l$). Thus, the initial and boundary conditions for the system of Eqs. (4) can be written in the form

$$\begin{aligned} &a_1(t=0)=a_1^0(z), \quad a_1(z=0)=a_1^S(t), \\ &a_2^*(t=0)=a_2^{0*}(z), \quad a_2^*(z=l)=a_2^{S*}(t). \end{aligned} \tag{5}$$

3. The solution of Eq. (4) with Eq. (5) was obtained by the Green function $G_{ij}(t,z,z')$ method;[20] here the amplitudes of the signal and idler waves are given by

$$\begin{aligned} a_1(t,z) &= \int_0^t dt' \int_0^l dz' [G_{11}(t-t',z,z')F_1(t',z') \\ &\quad + G_{12}(t-t',z,z')F_2(t',z')], \\ a_2^*(t,z) &= \int_0^t dt' \int_0^l dz' [G_{21}(t-t',z,z')F_1(t',z') \\ &\quad + G_{22}(t-t',z,z')F_2(t',z')], \end{aligned} \tag{6}$$

where

$$\begin{aligned} &F_1(t,z)=a_1^0(z)\,\delta(t)+v_1 a_1^S(t)\,\delta(z), \\ &F_2(t,z)=a_2^{0*}(z)\,\delta(t)+v_2 a_2^{S*}(t)\,\delta(l-z). \end{aligned}$$

Here $\delta(x)$ is the Dirac delta function.

Now, in place of Eqs. (4) and (5) we have to solve simpler equations for the four Green functions with zero boundary and initial conditions:

$$\begin{aligned} &\left(\frac{\partial}{\partial t}+\Gamma_1+v_1\frac{\partial}{\partial z}\right)G_{11}=Vh_pG_{21}+\delta(t)\,\delta(z-z'), \\ &\left(\frac{\partial}{\partial t}+\Gamma_2-v_2\frac{\partial}{\partial z}\right)G_{21}=Vh_pG_{11}, \\ &\left(\frac{\partial}{\partial t}+\Gamma_2-v_2\frac{\partial}{\partial z}\right)G_{22}=Vh_pG_{12}+\delta(t)\,\delta(z-z'), \\ &\left(\frac{\partial}{\partial t}+\Gamma_1+v_1\frac{\partial}{\partial z}\right)G_{12}=Vh_pG_{22}. \end{aligned} \tag{7}$$

It is clear from Eq. (7) that it is actually necessary to solve only one pair of equations, since the solution of the other is obtained automatically on switching the subscripts $1 \leftrightarrow 2$ and making the substitutions $z \to l-z$, $z' \to l-z'$.

The equations for G_{ij} were solved by the Laplace transform method in the time domain, which yielded a system of ordinary differential equations whose solution presents no special difficulty. After in inverse Laplace transform, we finally obtain

$$\begin{aligned} G_{11}(t,z,z') &= \frac{\rho}{v_1}G\left[\rho t+\eta\left(\frac{z}{l}-\frac{z'}{l}\right),\frac{z}{l},\frac{z'}{l}\right] \\ &\quad \times \exp(-\omega_r t+k_r(z-z')), \\ G_{21}(t,z,z') &= \frac{\rho}{\sqrt{v_1 v_2}}H\left[\rho t+\eta\left(\frac{z}{l}-\frac{z'}{l}\right),\frac{z}{l},\frac{z'}{l}\right] \\ &\quad \times \exp(-\omega_r t+k_r(z-z')). \end{aligned} \tag{8}$$

Here

$$\begin{aligned} H(\tau,\zeta,\zeta') &= \frac{\sigma}{2}\sum_{n=0}^{\infty}[P_{2n}(\tau,\alpha_n)-P_{2n}(\tau,\beta_n) \\ &\quad -P_{2n+2}(\tau,\gamma_n)+P_{2n+2}(\tau,\delta_n)], \\ G(\tau,\zeta,\zeta') &= \frac{\sigma}{2}\sum_{n=0}^{\infty}[P_{2n+1}(\tau,\alpha_n)-P_{2n+1}(\tau,\beta_n) \\ &\quad -P_{2n+1}(\tau,\gamma_n)+P_{2n+1}(\tau,\delta_n)] \end{aligned}$$

for $\zeta < \zeta'$ and

$$\begin{aligned} G(\tau,\zeta,\zeta') &= \delta(\tau-\alpha_0)+\frac{\sigma}{2}\sum_{n=0}^{\infty}[P_{2n-1}(\tau,\alpha_n) \\ &\quad -P_{2n+1}(\tau,\beta_n)-P_{2n+1}(\tau,\gamma_n) \\ &\quad +P_{2n+3}(\tau,\delta_n)] \end{aligned}$$

for $\zeta > \zeta'$. In these equations

$$P_n(\tau,\xi)=\theta(\tau-\xi)\left(\frac{\tau-\xi}{\tau+\xi}\right)^{n/2}I_n(\sigma\sqrt{\tau^2-\xi^2}),$$

$\theta(x)$ is the Heaviside unit function, $I_n(x)$ is the Bessel function with an imaginary argument, and

$$\begin{aligned} &\alpha_n=2n+|\zeta-\zeta'|, \quad \beta_n=2n+2-(\zeta+\zeta'), \\ &\gamma_n=2n+(\zeta+\zeta'), \quad \delta_n=2n+2-|\zeta-\zeta'|, \\ &\sigma=\frac{Vh_pl}{\sqrt{v_1v_2}}, \quad \rho=\frac{2v_1v_2}{(v_1+v_2)l}, \quad \eta=\frac{v_1-v_2}{v_1+v_2}, \\ &\omega_r=\frac{v_1\Gamma_2+v_2\Gamma_1}{v_1+v_2}, \quad k_r=\frac{\Gamma_2-\Gamma_1}{v_1+v_2}. \end{aligned}$$

4. The expressions for the Green functions (8) together with Eqs. (6) are the general solution for the parametric interaction of waves with nonstationary, locally-uniform parametric pumping for an interaction region $l \gg \lambda$. In deriving them, we have not imposed any restrictions on the form of the initial and boundary conditions, or on the duration and power of the pump pulse. It can be shown that our solution yields all the particular cases obtained before.[1–3,8,18,19] For example, an expression for the generation threshold follows from Eq. (8) in the limit $t \to \infty$. As a result, we obtain a

generalization of the threshold formula for parallel pumping first derived, as far as we know, in Ref. 5 for identical waves ($\nu_1=\nu_2\equiv\nu$; $\Gamma_1=\Gamma_2\equiv\Gamma$):

$$(h_{p\,th}V)^2=\frac{(\nu_1\Gamma_2+\nu_2\Gamma_1)^2}{4\nu_1\nu_2}+\nu_1\nu_2\chi^2, \tag{9}$$

where χ is the solution of the equation

$$2\nu_1\nu_2\chi=-(\nu_1\Gamma_2+\nu_2\Gamma_1)\tan(\chi l), \quad \pi/2l\leqslant\chi\leqslant\pi/l.$$

The gain coefficient for a signal passing from the input to the output of the active region in the stationary subthreshold regime ($t\to\infty, h_p<h_{p\,th}$) is given by

$$K_1=\left[\cos\sqrt{\sigma^2+\Omega^2}-i\frac{\Omega}{\sqrt{\sigma^2+\Omega^2}}\sin\sqrt{\sigma^2+\Omega^2}\right]^{-1}, \tag{10}$$

where $\Omega=(\Delta\omega+i\Gamma)(l/v)$, which expression is a generalization of the one obtained in Ref. 6 for the resonance $\Delta\omega\equiv\omega-\omega_p/2=0$.

Without going into detailed comparisons here, we note only that we have obtained the same results as Ref. 3 for stationary pumping and zero initial conditions, as Ref. 21 for zero initial conditions, as Refs. 8 and 19 for stationary initial and boundary conditions, as Ref. 14 for uniform ($l\to\infty$) pumping, and so on, which confirms the validity and generality of our results.

5. The purpose of this paper is to study parametric effects in detail, with backward volume magnetostatic waves in YIG ferrite films as an example. Thus, we shall begin by discussing the peculiarities of that case. First, let us examine the effect of the multiwave character of the ferrite sample. At the frequencies of the pump, signal, and idler, there is an infinite set of eigenmodes and waves that differ in the magnitude and direction of the wave vector **k**. Only in an ideal crystal and only for small amplitudes are these waves independent. When the wave amplitudes are finite, various kinds of nonlinear effects become significant, in particular, the parametric effects that form the basis of the effects considered in this paper. At the same time, the conservation laws for parametric interactions ($\mathbf{k}_1+\mathbf{k}_2=\mathbf{k}_p$ and $\omega_1+\omega_2=\omega_p$) simultaneously permit a parametric coupling between an infinite number of eigenmodes and waves in ferrite films; their excitation will be determined solely by the condition that the threshold amplitude $h_{p\,th}$ (9) be exceeded. Unfortunately, it seems that in ferrite films there are waves (exchange two-dimensional spin waves) whose excitation threshold, because their propagation velocity ($\nu\sim10^4$ cm/s) is low, lies considerably below the threshold for excitation of the backward volume magnetostatic waves, whose velocity is usually on the order of 10^6 cm/s. According to Eq. (9), for the typical dimensions of the active region, $l\simeq0.1$-1 mm, even when the damping parameters of the two-dimensional spin waves and the backward volume magnetostatic waves are equal ($\Gamma\sim(2-4)\cdot10^6\,\mathrm{s}^{-1}$), the excitation threshold for the former is more than order of magnitude below that for the latter. If, on the other hand, we note that the parametric effects increase significantly only near the threshold, then it becomes clear that in order to achieve significant gain for backward volume magnetostatic waves with parametric pumping, the amplitudes of the variable magnetic pump fields must be an order of magnitude higher than the excitation thresholds for the two-dimensional spin waves. Here the amplitudes of the two-dimensional spin waves increase exponentially, suppressing (owing to phase limiting and other nonlinear mechanisms[11,15]) the parametric coupling of the pump with other degrees of freedom and damping them further, so that the excitation threshold for the backward volume magnetostatic waves cannot be reached at all, even for pump fields at these levels.[6,22] These undesired effects can be eliminated without letting the amplitudes of the two-dimensional spin waves reach significant values, by, for example, using a pulsed pump oscillator[6] with a maximum pulse length $\tau_{p\,\max}\ll\Gamma^{-1}$. $\tau_{p\,\max}$ certainly depends on the pump power (the higher P_p is, the more rapidly the amplitude of the two-dimensional spin waves increases) and, as we shall see below, for real YIG films and $P_p\sim1$ W, $\tau_{p\,\max}$ should not exceed a few tens of nanoseconds.

Strong dispersion is yet another feature of backward volume magnetostatic waves.[4] Their diffusion length is only a few millimeters.[23] Thus, as a signal pulse propagates from the input converter to the active region, it loses its initial rectangular shape and becomes bell shaped.[23] This is facilitated by the finite bandpass of the input converters, which is usually less than 100 MHz, comparable to the reciprocal of the duration of the nanosecond pulses. Thus, besides rectangular signal pulse shapes, we have examined bell shaped ones with a hyperbolic secant form, $1/\cosh(1.76t/\tau_s)$, where τ_s is the full width of the bell shaped pulse at a level of approximately 0.7 times the maximum. In practice, τ_s is close to the width of the initial rectangular pulse.

6. Because of the complexity of Eqs. (6) and (8), at first we shall carry out the subsequent analysis for two limiting cases corresponding to the two operating regimes of a parametric amplifier with local pumping. The first regime is quasistationary, with a pump pulse duration τ_p that is considerably longer than the transit time τ_l for the signal across an active region of length l, i.e., $\tau_p\gg\tau_l=l/v$. Here, however, the restrictions on τ_p and l should be kept in mind: because of the multiwave property of the ferrite film, $\tau_p\leqslant\tau_{p\,\max}\ll\Gamma^{-1}$, while $l\gg\lambda$ in the case where there is no copropagating interaction and the truncated Eqs. (4) are not generally valid. Thus, this regime is referred to as quasistationary, rather than stationary. The second regime is nonstationary and is characterized by the opposite inequality, $\tau_p\ll\tau_l$. The difference in these regimes is that, in the first case of a small pump region l and long pump pulse τ_p ($\tau_p\gg\tau_l$), the parametrically coupled waves (see Fig. 2, waves 3–6 or 4 and 5) are rigidly coupled by the frequency condition for a parametric resonance ($\omega_3+\omega_6=\omega_4+\omega_5=\omega_p$), but their wave numbers k_3,k_6 or k_4,k_5 (see Fig. 2) can differ by an amount $\Delta k\leqslant1/l$. In the second case of a long pump region and a short pump pulse ($\tau_l\gg\tau_p$), on the other hand, the condition for conservation of the wave number of the parametrically interacting waves must be satisfied rigorously (waves 3, 4 or 5, 6 in Fig. 2; $k_3=k_4$, $k_5=k_6$), but, because of the pump duration τ_p, the frequency condition has a set of spectral components within the frequency intervals $\omega_p\pm1/\tau_p$, and

can be smeared out by that amount; i.e., $\omega_3+\omega_4=\omega_5+\omega_6=\omega_p\pm 1/\tau_p$. Even from this simple analysis, it is clear that in the quasistationary regime the frequency bands and transition processes in a parametric amplifier will be determined mainly by the length l of the interaction region and the transit time for a wave across this region, $\tau_l=l/v$. In the second, nonstationary regime the decisive parameter will be the pump pulse duration τ_p. Of course, in general, the characteristics of a parametric amplifier with local pumping will be determined by both quantities (τ_l and τ_p), but the shorter the characteristic time is, the more effect it will have.

7. Let us first consider the quasistationary regime, for which $\tau_l \ll \tau_p$, in more detail. The gain and bandpass can be obtained from Eq. (10) in this case. The resonant (at $\omega=\omega_p/2$) gain for the forward wave with $\Gamma^{-1}\gg\tau_l,\tau_s,\tau_p$ is given by

$$K_1\left(\omega=\frac{\omega_p}{2}\right)=K_{10}=\frac{1}{\cos\sigma}=\frac{1}{\cos\left(\frac{\pi}{2}\right)\left(\frac{h_p}{h_{p\,th}}\right)}$$

$$=\frac{1}{\cos\left(\frac{\pi}{2}\right)\sqrt{\left(\frac{P_p}{P_{p\,th}}\right)}}, \tag{11}$$

where P_p and $P_{p\,th}$ are, respectively, the pump power and its threshold value. The bandpass for $K_{10}\geqslant 5$ is

$$\Delta f\simeq\frac{1}{2K_{10}\tau_l},\quad K_{10}\Delta f\simeq\frac{1}{2\tau_l}. \tag{12}$$

As might be expected, all the characteristics of the quasistationary regime depend significantly on the length of the interaction region. Similar results hold for the backward wave, and these can be obtained easily from Eqs. (11) and (12) using the Manley-Rowe relations, which state that the energy inputs from the pump into the circuits of the signal and idler waves are the same if their frequencies are equal.[24] Thus, $K_{20}^2=K_{10}^2-1$ and the resonant reversal coefficient takes the form

$$K_2\left(\omega=\frac{\omega_p}{2}\right)=K_{20}=\tan\sigma. \tag{13}$$

All the parameters of the quasistationary regime depend significantly on the pump power. According to Eqs. (11) and (12), as the pump power approaches the threshold, the gain coefficient increases and the passbands become narrower, so that the pulse shape is strongly distorted. Thus, the quasistationary regime can be used in parametric amplifiers with local pumping only for small gains, when the amplified pulses are not significantly distorted. The pump power in this regime must not exceed the threshold $P_{p\,th}$.

8. We now turn to an analytic study of the other limiting case, for which $\tau_p\ll\tau_l$, i.e., to the nonstationary regime. An especially simple expression for the resonant gain K_{10} is obtained here for $\tau_p\leqslant\tau_s/2$ (in this case, each spectral component of the amplified signal corresponds to a frequency doubled spectral component of the pump):

$$K_{10}=\cosh(Vh_p\tau_p),\quad \tau_p\leqslant\tau_s/2. \tag{14}$$

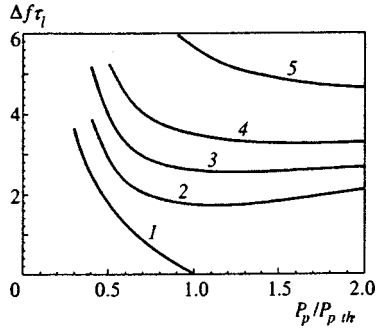

FIG. 3. The relative passband $\Delta f\tau_l$ as a function of the relative power $P_p/P_{p\,th}$ for different values of the ratio τ_p/τ_l: *1*—∞; *2*—3; *3*—2; *4*—1.5; *5*—1.

For $\tau_p\gg\tau_s$ (actually, for $\tau_p\geqslant 5\tau_s$),

$$K_{10}=I_0(Vh_p\sqrt{2\tau_s\tau_p}),\quad \tau_p\gg\tau_s. \tag{15}$$

Here the Manley-Rowe relations can also be used to obtain the conversion coefficients for the backward wave. For example, for $\tau_p\leqslant\tau_s/2$, we have $K_{20}=\sinh(Vh_p\tau_p)$.

It is most interesting that the pulse shapes in an essentially nonstationary regime are practically independent of the pump power. In this case, the threshold power does not have the same significance as in the quasistationary case. This can be confirmed by comparing Eqs. (11)–(13) and (14), (15). For example, at high pump powers $Vh_p\tau_p\gg 1$, the time to reach the amplitude is now on the order of $5\tau_p/\sqrt{Vh_p\tau_p}=5\tau_p/\sqrt{\ln(2K_{10})}$, i.e., it even decreases as the power is raised. In this regard, the nonstationary regime is of practical interest for signal processing in parametric amplifiers with local pumping. In this regime, as opposed to the quasistationary regime, the pump power can exceed the threshold $P_{p\,th}$. It should be kept in mind that increasing the pump power has its negative aspects: in order to reduce the influence of parasitic two-dimensional spin waves, it is necessary to reduce the pump pulse length τ_p. As we shall see below, in the nonstationary regime we actually have $\tau_{p\,\max}\simeq 30-40$ ns.

An exact analysis of the transition of an amplifier from the quasistationary regime into the nonstationary regime was made with the aid of Eqs. (6) and (8). Figure 3 shows the relative passbands $\Delta f\tau_l$ as functions of $P_p/P_{p\,th}$ for different pump pulse durations τ_p. Curve *1* for the quasistationary regime was constructed using Eq. (10). It is quite clear from the figure that, while the passband approaches zero as $P_p\to P_{p\,th}$ in the quasistationary regime, when the pump pulse duration is reduced it becomes possible to maintain a finite passband for arbitrary powers. By $\tau_p\leqslant 3\tau_l$ the passband $\Delta f\geqslant 2/\tau_l$ for arbitrary powers, although the passband becomes narrower near the threshold $P_{p\,th}$. For $\tau_p\sim\tau_l$ the passband increases by more than a factor of two, and the power $P_{p\,th}$ ceases altogether to play any special role in the plots of the transmission bandwidth as a function of pump power. Here the gains and conversion coefficients exceed 10 dB.

9. In the case $\tau_p\ll\tau_l$ studied here, there is the possibility of obtaining a backward wave with a signal whose time variation is reversed in time relative to the time t_p the pump power pulse is applied.[3] In other words, the leading edge of the pulse becomes the trailing edge, and *vice versa*. In the

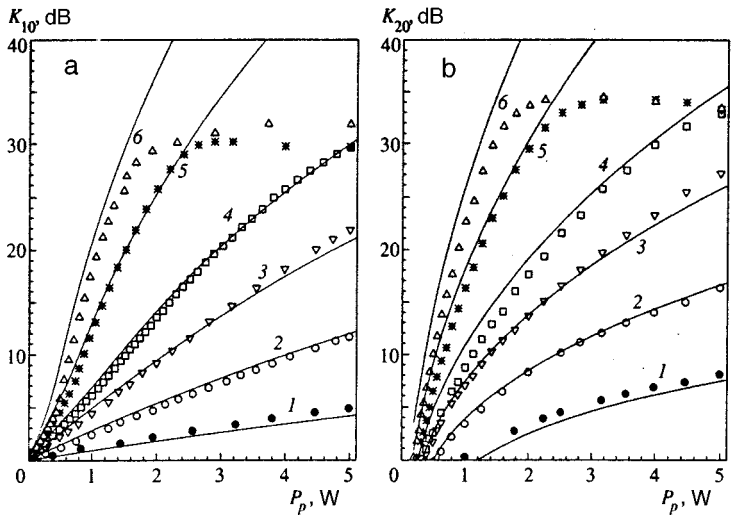

FIG. 4. The resonant gain K_{10} for the incident wave (a) and the resonant conversion coefficient K_{20} for the incident wave into a backward wave (b) as functions of pump power for different pump pulse durations τ_p: *1*—10 ns; *2*—20 ns; *3*—30 ns; *4*—40 ns; *5*—60 ns; *6*—80 ns; $\tau_l = 150$ ns, $\tau_s = 20$ ns, $P_{p\,th} = 40$ mW. The smooth curves are theoretical calculations for the same parameters.

case of a pump pulse with a δ-function form ($Vh_p \sim \delta(t_p - t)$), the signal amplitude of the reversed wave has the form[3]

$$a_2^*(z=0,t) = \text{const} \cdot a_1(2t_p - t), \tag{16}$$

i.e., a signal pulse that is in the active region at time t_p begins to move backwards from the point where the δ-function shaped pump pulse found it and, after a time t_p reaches the beginning of the active region, but now is the trailing edge of a pulse in the input converter.

We have studied the case of non-δ-function pump pulses lasting a finite time $\tau_p \geqslant 0$, while remaining, as before, in the nonstationary approximation, i.e., $\tau_p \ll \tau_l$. Here, instead of an ideal reversal of the time evolution of the pulse of Eq. (16), distortions arise owing to the finite pump duration τ_p:

$$a_2^*(z=0,t) = \sinh(Vh_p\tau_p)\left[a_1(2t_p + \tau_p - t) + \left(\frac{Vh_p\tau_p}{\tanh(Vh_p\tau_p)} - 1 \right) \frac{1}{2V^2h_p^2} \frac{\partial^2 a_1}{\partial t^2} \right]. \tag{17}$$

The distortions in the pulse shape are caused by the second term in the square brackets of Eq. (17); they are proportional to the steepness of the pulse fronts of the signal wave and, as expected, go to zero for $\tau_p = 0$. An analysis of Eq. (17) shows that the duration of the transition processes is on the order of τ_p in this case, as before.

10. We have studied the passage of real bell-shaped signals with shapes $\propto 1/\cosh(1.76t/\tau_s)$ through a locally pumped parametric amplifier. Here we note some nontrivial features of the amplified signal profiles.

One of these is a change in the time $\Delta\tau_l$ for propagation of the maximum of the output signal as the pumping is varied, which usually shows up as an apparent slowing down of the signal. This effect occurs because the trailing part interacts with the large amplitude idler wave which increases exponentially with time at each point in the active medium. The time $\Delta\tau_l$ depends on the pump power and the pulse durations τ_s and τ_p, and increases as these quantities increase.

Yet another feature of the output signal profiles is that the duration of the output signals can be reduced compared to the input signals in certain cases. This means that the spectral composition of the input signal can be broadened during nonstationary parametric amplification, while it can only be cut off during ordinary linear amplification because of the finiteness of the transmission bandpass. The physical reason for this narrowing is quite clear: if the pump pulse is shorter than the signal pulse, while the former, in turn, is longer than the interaction length τ_l, then only part of the signal pulse will be amplified, so that the output signal is narrowed. One nontrivial result of solving Eqs. (6) and (8) is that for high pump powers the signal can be narrowed even for $\tau_s < \tau_l$. For example, when $\tau_s = 0.5\tau_l$, $\tau_p = 0.2\tau_l$, and $P_p = 10P_{p\,th}$, the output signal is narrowed by a factor of 1.3. The backward wave pulse can also be narrowed. It turns out to be shorter than τ_s.

The limiting case of a long, quasimonochromatic signal ($\tau_s \to \infty$), where the total duration of the output signal equals $\tau_p + \tau_l$, while its amplitude increases resonantly as $\omega \to \omega_p/2$, is of special interest. In fact, this is a filter with a resonant frequency that is determined by the pump frequency and with an output pulse shape that is independent of the signal duration. In addition, by measuring the total duration $\tau_p + \tau_l$ of the output signal experimentally, in this case we can determine the time τ_l and, from it, the length $l = v\tau_l$ of the interaction region, which plays a decisive role in all the theoretical calculations.

We have only shown the results of the nontrivial theoretical inferences here. As for the conventional dependences, such as the dependence of the gains on the pump power, they will be presented below during a comparison of the theoretical and experimental results. (See Fig. 4, for example.)

11. We now consider the case of a small pump region, $l \leqslant \lambda$, where the co-propagating interaction begins to play a larger role alongside the counter-propagating interaction. As opposed to optics, this case is easy to realize in practice for backward volume magnetostatic waves. In our experiments, for example, $\lambda = 2\pi/k \simeq 400\ \mu$m, while l could be reduced to 200 μm by using a wire resonator. As indicated above, the

approximation of wave packets can be used only for low pumping, so we do not show an analytic solution of Eq. (1) by this method. We note only that the modulation coefficients for the forward and backward signals are the same and equal to $2\sigma\exp(-2kl)$ ($\sigma \ll 1$), i.e., they decrease exponentially as the size l of interaction region increases and increase linearly with the pumping. It turned out, as well, that a copropagating interaction changes the generation threshold, reducing it by roughly a factor of $[1-\exp(-2kl)]^{-1}$. In the case of arbitrary pumping, the system of Eqs. (1) must be solved numerically; in the calculations we included the three hundred primary waves of the lowest mode of the backward volume magnetostatic waves within the wave vector interval $\pm 300\,\mathrm{cm}^{-1}$. The results of these calculations differ substantially from those of the wave packet approximation, even for $\sigma \geqslant 0.5$.

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section primary attention is devoted to quantitative comparison of the preceding theoretical results with experiment. All the major conclusions of the theory were either confirmed experimentally, or the reason for any disagreement was understood.

1. A large pump region, $l \gg \lambda = 2\pi/k$, was realized in model II, with $l \simeq 3.3$ mm and $\tau_l = 150$ ns. In this case it was easy to create the most tolerable nonstationary regime for parametric amplification ($\tau_p \ll \tau_l$) and there was almost no beating of the output signals, which corresponds just to counterpropagating parametric interactions of the waves. Here the truncated equations (4) for coupled wave packets can be used for a theoretical description.

Figure 4 shows the measured resonant gain coefficient K_{10} for the transmitted wave and the resonant conversion coefficient K_{20} for conversion of the input signal into a backward wave as functions of the pump power P_p for different pump pulse durations τ_p. The signal pulse duration in the pump region is $\tau_s = 20$ ns. Theoretical calculations using Eqs. (6) and (8) are also shown as smooth curves for the following parameters of the experimental model: $\tau_l = 150$ ns, $v = 2.2 \cdot 10^6$ cm/s, $l = \tau_l v = 3.3$ mm, and $P_{p\,th} = 40$ mW. The time τ_l for the signal wave to cross the active region and the length of the latter were determined experimentally from the measured duration of the output pulse ($\tau_p + \tau_l$) in regimes with a continuous signal and pulsed pumping, as mentioned in paragraph 10 of the previous section. The pump power was lowered in order to minimize the effect of parametrically excited two-dimensional spin waves on the result.

There is good agreement between the experimental data and theoretical curves in Fig. 4. The saturation of the $K_{10}(P_p)$ and $K_{20}(P_p)$ curves at a level of 30–35 dB is apparently caused by the effect of parametric two-dimensional spin waves whose amplitude reaches levels of experimental significance during the time τ_p the pump is acting. The higher the pump power P_p is, the shorter the time τ_p required for this to happen. According to Fig. 4, for this model the maximum pump pulse duration $\tau_{p\,\max}$ for operation of a parametric amplifier without two-dimensional spin waves at the maximum pump power of $P_p = 5$ W was $\tau_{p\,\max} \simeq 40$ ns. When the pump power is lowered to 1.5 W, the pump pulse duration can be raised to 80 ns.

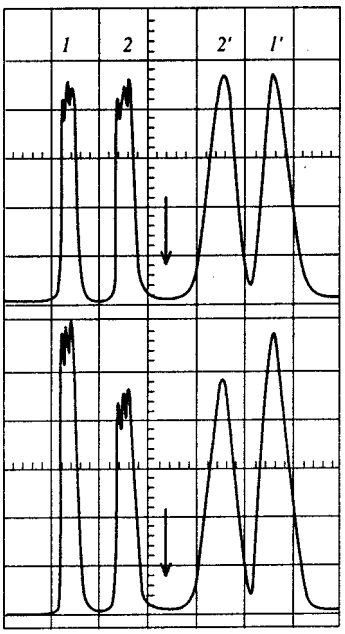

FIG. 5. Time reversal of an input signal waveform: *1, 2*—signals reflected from the input of the experimental model without having interacted with the pump; *1′, 2′*—signals at the input to the experimental model formed by a reversed wave. The amplitudes of the signals *1* and *2* are attenuated by 20 dB relative to the signals *1′* and *2′*. The large divisions represent 50 ns. The arrows denote the time the pump is applied.

Figure 5 shows the results of using the phenomenon of reversing the wave front of magnetostatic waves to perform the operation of time reversal of a signal shape. In the experiment we used a sequence of two pulses with different amplitudes, whose ordering in time in the wave reversed by the model (Fig. 5, the two pulses on the right) was opposite to their order in a signal that had not interacted with the pump and had been reflected from the input to the model (the pulses on the left), which was actually a copy of the input signal (given the imperfect reflection). Prior to experiment, a sequence of two signal pulses of equal duration and amplitude was fed to the input, while the time of delivery and the duration of the pump were chosen so as to ensure equality of the amplitudes of the reversed signals. This procedure is necessary in order to compensate the effect of damping of the backward volume magnetostatic waves in the YIG on the shape of the reversed signal (in our case, the ratio of the amplitudes of the backward pulses). For the case in Fig. 5, this compensation was possible using a pump signal with $\tau_p = 30$ ns. As discussed in the theoretical section, the use of such a long pump pulse led to significant distortions of the output pulses relative to the input pulses, but in Fig. 5 the reversal of the shape in the backward wave is clearly visible for the case of input pulses with different amplitudes.

2. The opposite case of a short active region with $l \leqslant \lambda = 2\pi/k$ was realized using model I with a wire half-wave pump resonator. Here the length of the active region ($l \simeq 220\,\mu$m, $\tau_l = 10$ ns) was more than an order of magnitude shorter than in the case of an open dielectric resonator. This

led to significant changes in the properties of the paramagnetic amplifier for magnetostatic waves. First, because of the enhanced role of the copropagating interaction (see the previous section), beating of the amplitudes of the forward and backward waves took place. The modulation coefficient (i.e., the ratio of the difference between the maximum and minimum amplitude of the beats to their sum) was as high as 80–90% for the highest pumping levels. This indicates that the copropagating interaction could be of the same order of magnitude as the counterpropagating interaction. Second, according to Eq. (9), there was a sharp (about an order of magnitude) increase in the threshold for excitation of backward volume magnetostatic waves owing to an increased role for energy losses from the interaction space. The threshold for backward volume magnetostatic waves increased to 1.9 W while the threshold for excitation of two-dimensional spin waves was essentially unchanged (~ 10 mW), since the mean free path for the latter was, as before, smaller than the size of the active region, even with $l \simeq 220\,\mu$m, and energy loss processes were insignificant. All this led to an increased contribution from the two-dimensional spin waves and to a further reduction in the duration of the maximum possible pump pulses, $\tau_{p\,\max}$, to 10–20 ns.

The experimental data on the maximum (because of beating, the readout was taken from the upper edge of the output pulse that had been smeared out by the beating) resonant gain and conversion coefficients as functions of the pump power are in good agreement with the theoretical curves obtained by numerical solution of Eq. (1). Here the best agreement with experiment was obtained for a pump field distribution that was not rectangular, as assumed for solving the truncated equations, but had a Lorentz form

$$h_p = h(z) = \frac{h_{p\,\max}}{1+(2z/l)^2}$$

that was closer to that in the experiment. In the experiment with $\tau_p = 30$ ns, saturation of the $K_{10}(P_p)$ and $K_{20}(P_p)$ curves was observed. As before (Fig. 4), this was caused by the parametric excitation of spin waves. We do not show these curves here, since they are fundamentally the same as those in Fig. 4, except that the gain is substantially lower (by roughly 10 dB). We show only the change in the time evolution of the output signal under the influence of the pump, in order to illustrate clearly two nontrivial features of the nonstationary parametric amplifier: the change in the propagation time for the signal peak, $\Delta\tau_l$, and the reduced duration of the output signals compared to the input. As noted in the theoretical section, a real narrowing of the pulses is observed only when $\tau_p, \tau_l \leqslant \tau_s$, which, for the experimental possibilities available to us, could be easily achieved in the case of model I with its small active region.

Figure 6 shows the time evolutions of the forward signal passing through the active region measured at the output antenna for various pump powers P_p. (See Fig. 1.) Curve *1*, measured without pumping ($P_p = 0$), represents the initial shape of the signal pulse with a duration of $\tau_s \sim 30$ ns. It is quite clear from this figure that after pumping starts, the delay to the peak of the pulse is increased (to $\Delta\tau_l \leqslant 15$ ns)

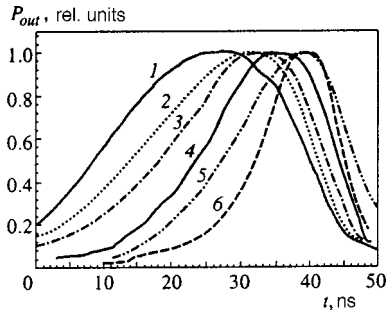

FIG. 6. The variation in the signal profile at the output antenna as the pump power P_p is raised: *1*—0 W; *2*—0.3 W; *3*—0.63 W; *4*—1.25 W; *5*—2 W; *6*—5 W; $\tau_p = 20$ ns, $\tau_l = 10$ ns.

and it is narrowed substantially (from 30 ns to 12 ns). This last effect may be important for amplifying the envelopes of solitons, for which an increase in the pulse amplitude must be accompanied by a simultaneous reduction in its duration. In the case of an open dielectric pump resonator considered above, there was also a narrowing of the signal, but only for $\tau_s \geqslant 100$ ns, which was related to the longer length of the active region in model II.

5. CONCLUSIONS

The interaction of travelling backward volume magnetostatic waves with a local nonstationary electromagnetic pump at twice their frequency in YIG films has ben studied theoretically and experimentally. In our experiments we have, for the first time, been able to observe the reversal of a wave front and the time reversal of the shape of backward volume magnetostatic wave signals. During reversal and amplification of backward volume magnetostatic wave pulses the shape of the output signals is found to have changed significantly, with narrowing under certain conditions, and an increase in the time delay of the pulse maximum when pumping is applied. Experimental dependences of the gain and of the conversion coefficient for conversion of backward volume magnetostatic waves into backward waves as functions of the power and duration of the pump signal have been obtained.

In order to explain the experimental results we have solved the general problem of the interaction of a travelling linear wave with a spatially localized parametric pulsed pump. In the case of a large region where the backward volume magnetostatic waves interact with the pump, when the length l of this region greatly exceeds the wavelength λ of the backward volume magnetostatic waves ($l \gg \lambda$), the problem was reduced to a system of two truncated equations for the wave packets that only take counterpropagating interactions (i.e., with oppositely directed wave vectors) of the waves into account. A Green function method was used to obtain a general solution for this system in terms of the Bessel function with an imaginary argument. Because the features of the interaction between backward volume magnetostatic waves and a variable magnetic pump field only show up through the constant coupling parameter, this solution can easily be extrapolated to any type of waves in various media with only a change in the coupling parameter. In the opposite

limiting case of a small pumping region ($l \leqslant \lambda$), numerical methods were used to solve the exact problem of the interaction of the eigenmodes of a ferrite film with a nonuniform nonstationary pump taking both counterpropagating and copropagating (with the wave vectors of the interacting waves in the same direction) interactions of the backward volume magnetostatic waves into account. Both of these cases were realized in the experiments.

Two amplification regimes have been found to exist: quasistationary and nonstationary. In the first case, the pump power P_p is below the threshold for excitation of backward volume magnetostatic waves, $P_{p\,th}$, while $\tau_p > \tau_l$, where τ_l is the time for a magnetostatic wave to cross the pump region. Here the usual behavior of a parametric amplifier is manifested, in particular, there is an unlimited growth in the gain as $P_p \to P_{p\,th}$ and the gain bandwidth then goes to zero. In the nonstationary regime, $\tau_p < \tau_l$, and the pump power can exceed the threshold power, i.e., $P_p > P_{p\,th}$. In this case, both the gain and the passband increase when the pump power is raised and they become greater than the values for the first regime. Yet another feature of the nonstationary amplification regime is the possibility of narrowing the output pulse, i.e., expanding its frequency spectrum, while a stationary linear amplifier can only cut off the signal spectrum.

Essentially all of the measured experimental dependences are in qualitative, but also good quantitative agreement with the corresponding theoretical curves at low powers and for short pump pulse durations. The discrepancies which arise when these parameters are increased can be explained by the influence of parasitic excitation of volume two-dimensional spin waves which are degenerate in frequency with the backward volume magnetostatic wave signals.

Some problems that require further study include the gain processes for nonlinear magnetostatic waves, in particular the formation and amplification of soliton envelopes in forward and backward waves.

This work was partially supported by the Fund for Fundamental Research of Ukraine (Grant No. 2.4/707), as well as by a U. S. NSF grant (Grant DMR-9701640) and a special supplementary grant from the Eastern European Program of the U. S. NSF (Grant RUI-DMR-9701640).

*)E-mail: melkov@boy.rpd.univ.kiev.ua
†)Department of Physics, Oakland University, Rochester, MI-48309, USA

[1] G. Agrawal, *Nonlinear Fiber Optics* [Russian translation], Mir, Moscow (1996).
[2] A. Yariv and P. Juch, *Optical Waves in Crystals* [Russian translation], Mir, Moscow (1987).
[3] B. Ya. Zel'dovich, R. F. Pilipetskiĭ, and V. V. Shkunov, *Wave Front Reversal* [in Russian], Nauka, Moscow (1985).
[4] A. G. Gurevich and G. A. Melkov, *Magnetic Oscillations and Waves* [in Russian], Nauka, Moscow (1994).
[5] V. S. L'vov and A. M. Rubenchik, IAÉ SOAN SSSR, Preprint, No. 31 (1976).
[6] G. A. Melkov and S. V. Sholom, Zh. Tekh. Fiz. **60**, 118 (1990) [*sic*].
[7] A. V. Bagada, G. A. Melkov, A. A. Serga, and A. N. Slavin, Phys. Rev. Lett. **79**, 2137 (1997).
[8] Yu. V. Gulyaev, P. E. Zel'berman, and A. V. Lugovskoĭ, Zh. Éksp. Teor. Fiz. **111**, 199 (1997) [JETP **84**, 109 (1997)].
[9] P.A. Kolodin, P. Kabos, C. E. Patton, B. A. Kilinikos, N. G. Kovshikov, and M. P. Kostylev, Phys. Rev. Lett. **80**, 1972 (1998).
[10] A. L. Gordon, G. A. Melkov, A. A. Serga *et al.*, JETP Lett. **67**, 913 (1998).
[11] A. G. Gurevich and G. A. Melkov, *Magnetization Oscillations and Waves*, CRC Press, New York (1996).
[12] R. V. Khokhlov, Radiotekh. Elektron. (Moscow) **6**, 917 (1961).
[13] N. Bloembergen, *Nonlinear Optics* [Russian translation], Mir, Moscow (1966).
[14] A. P. Sukhorukov, *Nonlinear Wave Interactions in Optics and Radio Physics* [in Russian], Nauka, Moscow (1988).
[15] V. S. L'vov, *Nonlinear Spin Waves* [in Russian], Nauka, Moscow (1987).
[16] A. V. Vashkovskii, V. S. Stal'makhov, and Yu. P. Sharaevskii, *Magnetostatic Waves in Microwave Electronics* [in Russian], Izd-vo. Saratovskogo Universiteta, Saratov (1993).
[17] D. M. Pepper, Sci. Am. **254**, 74 (1986).
[18] I. R. Shen, *Principles of Nonlinear Optics* [Russian translation], Nauka, Moscow (1989).
[19] A. M. Gorbunov, Zh. Éksp. Teor. Fiz. **67**, 869 (1998) [Sov. Phys. JETP **40**, 689 (1974)].
[20] V. S. Vladimirov, *The Equations of Mathematical Physics* [in Russian], Nauka, Moscow (1988).
[21] B. R. Suydam and R. A. Fisher, Opt. Eng. (Bellingham) **21**, 184 (1982).
[22] Ya. Monosov, *Nonlinear Ferromagnetic Resonance* [in Russian], Nauka, Moscow (1971).
[23] M. Chen, M. A. Tsankov, J. M. Nash, and C. E. Patton, Phys. Rev. B **49**, 12773 (1994).
[24] W. Louisell, *Coupled Mode and Paramagnetic Electronics*, Wiley, New York (1963).

Translated by D. H. McNeill

ERRATA

Erratum: The ($a\gamma\gamma\gamma$) vertex and three-photon axion decay in external magnetic fields [JETP 89, 13–17 (July 1999)]

V. V. Skobelev

Moscow State Industrial University, 109280 Moscow, Russia
Zh. Éksp. Teor. Fiz. **116**, 2212 (December 1999)

[S1063-7761(99)02612-8]

This article contains the following important misprints:

1. The second formula in Eq. (11) on p. 14 should read $(\tilde{\gamma}^5)^2=1$.
2. The last term in Eq. (12a) on p. 14 should be $\tilde{g}^{\alpha\rho}\tilde{g}^{\beta\sigma}$.
3. The last term in the second formula of Eq. (12d) on p. 14 should be $\varepsilon^{\alpha\beta}\tilde{g}^{\rho\sigma}$.
4. In expression (13) on p. 15 the denominator should appear raised to the power 1/2.
5. In Eq. (14) on p. 15 e in the numerator of the fraction should be replaced with e^3.
6. The eleventh and twelfth lines of Eq. (17) on p. 16 should read

$$+(q\varepsilon\kappa)\varepsilon_{\alpha\alpha'}(k''\varepsilon)_{\alpha''}]-\frac{1}{(m^2-\tilde{m}^2)^3}[(q\varepsilon\kappa)(\kappa_\alpha^{(1)}\kappa_{\alpha'}^{(2)}\kappa_{\alpha''}^{(3)}$$

$$+\kappa_\alpha^{(1)}(\kappa^{(2)}\varepsilon)_{\alpha'}(\kappa^3\varepsilon)_{\alpha''}+(\kappa^{(1)}\varepsilon)_\alpha\kappa_{\alpha'}^{(2)}(\kappa^{(3)}\varepsilon)_{\alpha''}+(\kappa^{(1)}\varepsilon)_\alpha(\kappa^{(2)}\varepsilon)_{\alpha'}\kappa_{\alpha''}^{(3)})-$$

7. In (17a) on p. 16 in the second term of the expression for $\tilde{m}^2$ the first factor should be $(q-k'')^2$.

