Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 677–684.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 787–796.
Original Russian Text Copyright © 2002 by Ryabtsev, Tretyakov.

NUCLEI, PARTICLES, AND THEIR INTERACTION

The Break of L–S Coupling and the Double Stark Resonance in the Spectrum of 36*P* ⟶ 37*P* Two-Photon Transition in Rydberg Atoms of Sodium

I. I. Ryabtsev* and D. B. Tretyakov**

Institute of Semiconductor Physics, Siberian Division, Russian Academy of Sciences, Novosibirsk, 630090 Russia
**e-mail: ryabtsev@isp.nsc.ru*
***e-mail: dtret@isp.nsc.ru*
Received September 19, 2001

Abstract—New results are presented of an experimental investigation of the spectrum of 36*P* ⟶ 37*P* two-photon microwave transition in Rydberg atoms of sodium in a constant electric field. Depending on the conditions of excitation of the initial 36*P* state (the constant electric field is switched on either before or after the exciting laser pulse) and polarization of laser radiation, a strong variation is observed of the amplitudes of individual two-photon transitions between the fine-structure Stark components of the 36*P* and 37*P* states. This effect is an analog of the Paschen–Back effect in a strong magnetic field and is due to the break of L–S coupling and to the variation of the wave functions of Rydberg electrons in an electric field. It is also found that the break of L–S coupling affects considerably the shape of double Stark resonance arising upon intersection of the virtual intermediate level of two-photon transition with the real intermediate 37*S* level. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The double Stark resonance on two-photon

$$nP \longrightarrow (n+1)P$$

microwave transitions in Rydberg atoms of Na in a weak electric field was first observed and investigated in [1, 2]. It consists in that the detuning of the real intermediate level $(n+1)S$ from the virtual intermediate level of two-photon transition decreases rapidly when an electric field is switched on, and, with a certain value of the field, the two-photon resonance transforms to the exact double resonance,

$$nP \longrightarrow (n+1)S \longrightarrow (n+1)P.$$

The probability of transition increases by several orders of magnitude, which leads to the emergence, in the absorption spectrum, of a wide band instead of separate narrow peaks corresponding to transitions between the fine structure components of the *P* states.

It was demonstrated in [1, 2] that the double Stark resonance may be used for absolute calibration of the strength of a weak electric field in vacuum. For each transition component, the double resonance arises in a narrow range of electric field strengths (~0.1 V/cm) if the intensity of microwave radiation is less than the saturation intensity of intermediate single-photon transitions. The main advantages of this method are as follows: first, the experiments are performed with single Rydberg atoms, and, therefore, one can state that a noncontact method of measuring weak electric fields is found; second, the values of the critical fields of double Stark resonance may be calculated with high accuracy for any $nP \longrightarrow (n+1)P$ two-photon transition, and, in this manner, a set of reference points may be obtained in a wide range of field strengths.

Unfortunately, in [1, 2] some Stark components of two-photon transitions were identified incorrectly, and the dependence of the spectrum of double Stark resonance on the conditions of laser excitation was not studied; the approximate formula of quadratic Stark effect was used for the absolute calibration of the electric field. Therefore, we performed new investigations of the spectrum of double Stark resonance on the 36*P* ⟶ 37*P* two-photon transition in Rydberg atoms of Na (the transition frequency of about 72.6 GHz). It has been found that the spectrum depends substantially both on the method of excitation of the initial 36*P* state and on the polarization of exciting laser radiation. We also performed an exact numerical calculation of transition frequencies in the electric field and of the critical fields of double Stark resonance.

2. THEORY

Figure 1 gives a calculated Stark diagram of the energy levels of Rydberg atoms of Na in the vicinity of hydrogen-like sets of levels $n = 35, 36$. In a weak electric field $E \le 10$ V/cm, when the shifts of *P* levels are much less than the distances to the nearest *D* levels (the lower component of hydrogen-like set), the 36*P* and 37*P* states are characterized by the quadratic Stark effect, because the *P* states of Na possess a significant

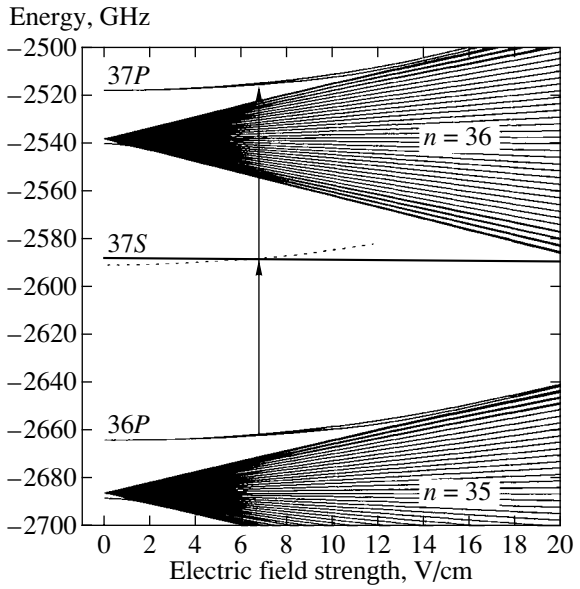

Fig. 1. A calculated Stark diagram of the energy levels of Rydberg atoms of Na in the vicinity of hydrogen-like sets of levels $n = 35, 36$ for sets with the projection of total angular momentum $|M| = 1/2$. The virtual intermediate level of $36P \longrightarrow 37P$ two-photon transition is indicated by the dotted curve. The double arrow corresponds to double Stark resonance.

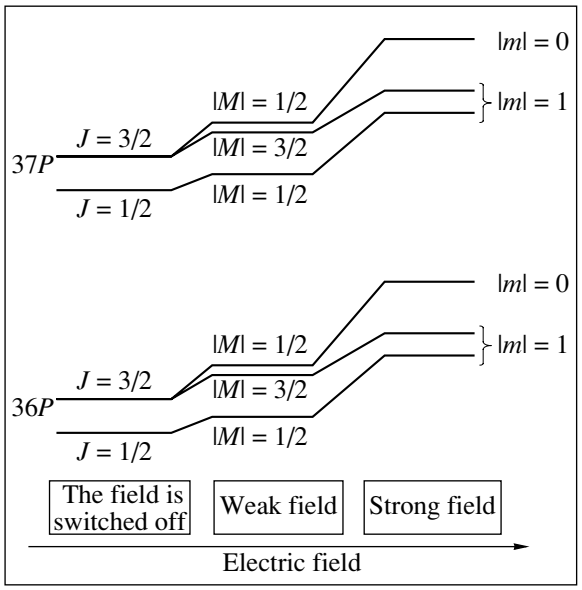

Fig. 2. The diagram of a break of L–S coupling for the 36*P* and 37*P* states in Na atoms in a constant electric field. In the absence of a field and in a weak field (<1 V/cm), the states are described by wave functions in the *nlJM* basis. In a strong field (>3 V/cm), the wave functions correspond to the *nlm* basis.

quantum defect ($\delta_P \approx 0.855$). The same is true of the intermediate level 37*S* ($\delta_S \approx 1.347$), which is located almost midway between the 36*P* and 37*P* levels (detuning from the virtual level of two-photon transition $\Omega \approx 2.6$ GHz; the virtual level is shown in Fig. 1 by the dotted curve).

The *nP* states are split by spin–orbit interaction [3],

$$\hat{H}_{LS} = \frac{\alpha^2}{2r^3}\mathbf{L}\cdot\mathbf{S} \qquad (1)$$

(α is the fine structure constant, r is the distance from electron to nucleus, **L** is the orbital angular momentum operator, and **S** is the electron spin operator), into two sublevels with the total angular momentum J equal to 1/2 and 3/2. The fine structure intervals are 124 and 114 MHz for 36*P* and 37*P*, respectively. The presence of *L–S* coupling in the case of *P* states complicates the Stark structure of transition spectra. The qualitative diagram of behavior of the 36*P* and 37*P* levels in the electric field is given in Fig. 2. In the absence of the field, the wave functions of individual fine-structure magnetic sublevels Ψ_{JM} of the *nP* state with the angular momentum J and its projection M are a linear superposition of undisturbed (disregarding the spin–orbit interaction) wave functions Φ_{lm} of a Rydberg electron with the orbital angular momentum $l = 1$ and its projection $m = 0, \pm 1$, multiplied by the respective spin wave functions φ_s with the spin projection $s = \pm 1/2$ [3],

$$\begin{aligned}
\Psi_{1/2\,-1/2} &= \sqrt{\frac{2}{3}}\Phi_{1\,-1}\varphi_{+1/2} - \sqrt{\frac{1}{3}}\Phi_{1\,0}\varphi_{-1/2},\\
\Psi_{1/2\,+1/2} &= \sqrt{\frac{1}{3}}\Phi_{1\,0}\varphi_{+1/2} - \sqrt{\frac{2}{3}}\Phi_{1\,+1}\varphi_{-1/2},\\
\Psi_{3/2\,-3/2} &= \Phi_{1\,-1}\varphi_{-1/2},\\
\Psi_{3/2\,-1/2} &= \sqrt{\frac{1}{3}}\Phi_{1\,-1}\varphi_{+1/2} + \sqrt{\frac{2}{3}}\Phi_{1\,0}\varphi_{-1/2},\\
\Psi_{3/2\,+1/2} &= \sqrt{\frac{2}{3}}\Phi_{1\,0}\varphi_{+1/2} + \sqrt{\frac{1}{3}}\Phi_{1\,+1}\varphi_{-1/2},\\
\Psi_{3/2\,+3/2} &= \Phi_{1\,+1}\varphi_{+1/2}.
\end{aligned} \qquad (2)$$

In a first approximation, the shift of the center of gravity and the splitting of these levels in a weak electric field are described by the formula [4]

$$\Delta W = -\frac{1}{2}\left\{\alpha_0 + \alpha_2\frac{3M^2 - J(J+1)}{J(2J-1)}\right\}E^2, \qquad (3)$$

where α_0 and α_2 denote the scalar and tensor polarizabilities, which increase rapidly with the principal quantum number of the Rydberg state in proportion to n^7. The values of polarizabilities for the 36*P* and 37*P* states were measured in [1, 2, 5]. Note that $\alpha_2 = 0$ for

the states with $J = 1/2$, i.e., these levels are not split, and the degeneracy of levels by the sign of M is not removed by the electric field.

In very weak fields (≤1 V/cm), formula (3) is fairly accurate (the error of <1%). However, in the region of double Stark resonance at $E = 6$–7 V/cm (see Fig. 1), it proves insufficient for a correct calculation of the critical fields of each of nine components of two-photon transition. This is associated with the fact that Eq. (3) is derived using perturbation theory and ignores the variation of the frequencies of transitions and wave functions of atomic states in an electric field. In [1, 2], the values of the critical fields were calculated using formula (3) and, therefore, need to be refined. More exact formulas may be borrowed from [6, 7]; however, they are rather complicated and call for measurements of additional parameters of Rydberg levels.

We used the numerical calculation of the eigenvalues of atomic energy in an electric field. The Stark diagram in Fig. 1 for the $|M| = 1/2$ states was obtained using the matrix diagonalization of the Hamiltonian of atom–field interaction [8],

$$\hat{H}_E = -\mathbf{d} \cdot \mathbf{E}, \quad (4)$$

where $\mathbf{d}$ is the operator of atomic dipole moment. The basis of unperturbed states was provided by fine-structure magnetic sublevels of the $nlJM$ states, $n = 34$–39, and the exact values of quantum defects were borrowed from [9]. This enabled us to calculate the eigenvalues of energies and obtain refined data on the critical field of double Stark resonance for the $36P \longrightarrow 37P$ transition, which are given in the table.

The calculation of wave functions and transition probabilities is a more complicated problem and calls for determination of the eigenvectors of the interaction operator matrix. For determining the wave functions in an arbitrary electric field, one must solve the problem of finding the eigenvectors of the matrix of the Hamiltonian $\hat{H}_{LS} + \hat{H}_E$ which includes both the spin–orbit interaction and the atom–electric field interaction. Therefore, we will restrict ourselves to qualitative treatment of the behavior of wave functions and transition probabilities in an electric field in accordance with Fig. 2.

In the absence of the field, the stationary wave functions of the $36P$ and $37P$ states are defined by formulas (2) and relate to the basis of $nlJM$ states, in which the Hamiltonian of spin–orbit interaction is diagonal. In a "weak" electric field (of the order of 1.5 V/cm), the atom–field interaction energy is compared with the spin–orbit interaction energy, which brings about a variation of the expansion coefficients in Eq. (1). Finally, in a "strong" field at $E \geq 3$–4 V/cm, an almost complete break of the L–S coupling occurs, and the behavior of Rydberg electrons may be described in the basis of nlm states using the wave functions Φ_{lm}. One should take into account the fact that a further field growth (>10 V/cm, see Fig. 1) is accompanied by the mixing of states of different parities, and the electron wave function will be a linear superposition of wave functions with different values of l.

The results of numerical calculation of the critical fields of double Stark resonance for two-photon transitions between the Stark sublevels of the $36P$ and $37P$ states in Na atoms. The resonance numbers correspond to their identification in the experimental records (Figs. 5 and 6)

No. of peak in Figs. 5 and 6	$36P_{J\|M\|} \longrightarrow 37P_{J'\|M'\|}$ transition	Critical field, V/cm
6	$36P_{3/2,\,1/2} \longrightarrow 37P_{3/2,\,1/2}$	6.37
8	$36P_{3/2,\,3/2} \longrightarrow 37P_{3/2,\,1/2}$	6.67
9	$36P_{1/2,\,1/2} \longrightarrow 37P_{3/2,\,1/2}$	6.73
2	$36P_{3/2,\,1/2} \longrightarrow 37P_{3/2,\,3/2}$	6.75
1	$36P_{3/2,\,1/2} \longrightarrow 37P_{1/2,\,1/2}$	6.79
4	$36P_{3/2,\,3/2} \longrightarrow 37P_{3/2,\,3/2}$	7.10
3	$36P_{3/2,\,3/2} \longrightarrow 37P_{1/2,\,1/2}$	7.15
7	$36P_{1/2,\,1/2} \longrightarrow 37P_{3/2,\,3/2}$	7.15
5	$36P_{1/2,\,1/2} \longrightarrow 37P_{1/2,\,1/2}$	7.20

The merit of $nP \longrightarrow (n+1)P$ two-photon microwave transitions in Rydberg atoms of Na consists in that they enable one to experimentally investigate the transition from weak to strong interaction of an atom with an electric field. Owing to close agreement between the scalar polarizabilities of the adjacent nP and $(n+1)P$ states, their great Stark shifts are partly compensated, and it turns out possible to determine the electric field dependences of the probabilities and frequencies of individual transitions by scanning the microwave generator frequency in a fairly narrow frequency range of the order of 1 GHz, rather than several gigahertz, as would be the case, for example, for single-photon $nP \longrightarrow n'S$ transitions [5]. The latter fact is important from the standpoint of ensuring the invariant intensity of microwave radiation in the region of interaction with Rydberg atoms.

3. EXPERIMENTAL SETUP

The experiments were performed with an effusion beam of Na atoms with a temperature of 500 K in a vacuum chamber at a residual gas pressure of 3×10^{-7} Torr (Fig. 3). The Rydberg states were excited according to the three-stage scheme of

$$3S_{1/2} \longrightarrow 3P_{3/2} \longrightarrow 4S_{1/2} \longrightarrow 36P_{J\,=\,1/2,\,3/2}$$

with the beam being transversely illuminated by the radiation of three synchronized tunable pulsed lasers with a high (5 kHz) pulse repetition rate. In the first and third stages, Rhodamine 6G and Oxazine 17 dye lasers were used, and in the second stage, a laser with F_2^- color centers in a LiF crystal.

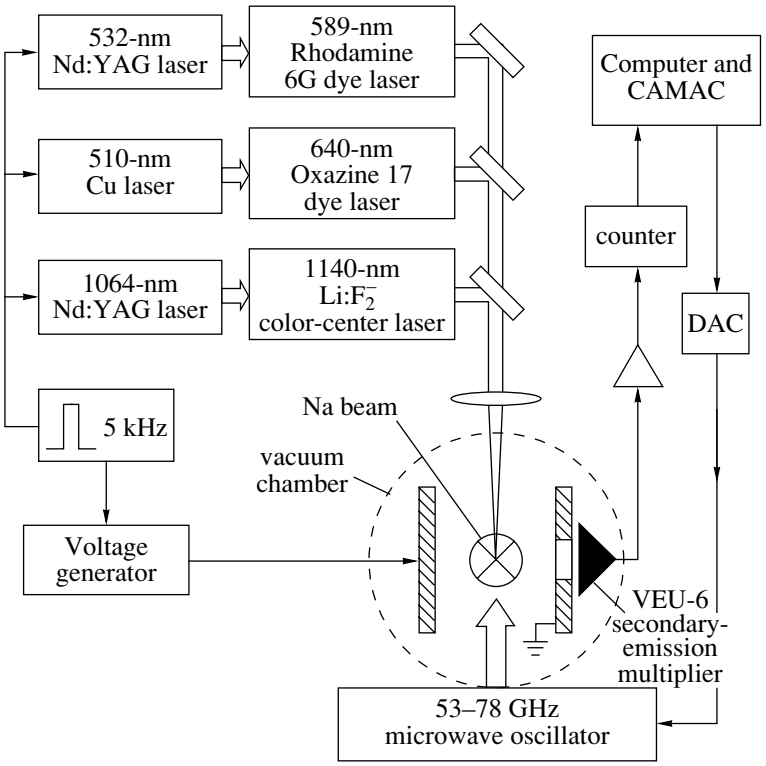

Fig. 3. The diagram of the experimental setup for microwave spectroscopy of Rydberg atoms of Na.

The atomic beam was then directed to the region of interaction with resonance microwave radiation at a frequency of 72–73 GHz, which was introduced via a waveguide into the space between two copper plates producing a constant electric field. The microwave field had the form of a standing wave of complex spatial configuration. Because the radiation wavelength (~4.2 mm) exceeded considerably the atomic beam diameter (1 mm), the point of interaction with the standing wave could be selected by varying the point of laser excitation of Rydberg states, as was demonstrated in [10].

The radiation source was provided by a G4-142 backward-wave oscillator with the tuning range of 53–78 GHz. The oscillator frequency was stabilized with the aid of an external heterodyne which, in turn, was locked to a quartz frequency synthesizer. The oscillator line width in the continuous mode was less than 20 kHz. The pulsed oscillation mode was used in the experiments in order to separate in time the moments of laser excitation, interaction with microwave radiation, and detection of populations of the Rydberg states. The effective radiation line width of 1 MHz corresponded to the microwave pulse duration of 1 μs.

In scanning the oscillator frequency, two-photon transitions between the fine-structure levels of the initial 36*P* state and the final 37*P* state were induced, which resulted in a variation of their populations. The populations were controlled by the method of selective field ionization in a pulsed electric field [11, 12]. The electrons formed as a result of ionization were detected by a VEU-6 vacuum channel secondary-emission multiplier, and the signal from the output of the latter was processed in the pulse counting mode in the CAMAC crate and computer. In order to reduce the effect of thermal background radiation, which causes unwanted transitions between adjacent Rydberg states and the reduction of their lifetimes [12], all elements of the detection system and input of microwave radiation were cooled down to the liquid nitrogen temperature of 77 K.

The time diagram of signals is given in Fig. 4. At the moment of time $t = 0$, a laser radiation pulse (Fig. 4a) excited both fine-structure sublevels of the 36*P* state. The electric field in the interaction region had a two-stage shape (Fig. 4b). The first, weak (0 to 10 V/cm), stage was switched on smoothly either before or immediately after the laser pulse and then reached a steady-state value prior to the moment of switching on of a microwave pulse with the duration of 1.2 μs (Fig. 4c). The second, strong, stage of the electric field increasing linearly to 220 V/cm was switched on after the termination of the microwave pulse and was used for selective field ionization of Rydberg atoms. Depending on the state of the atom, signals separated in time appeared at the VEU-6 output, which corresponded to the 36*P* and 37*P* states (Fig. 4d) ionized at different values of the electric field. In the pulse counting mode, the frequency

of each signal is proportional to the population of the given state. The signals were averaged over approximately 2000 laser pulses.

4. RESULTS

The spectrum of $36P \longrightarrow 37P$ two-photon transition was investigated for different intensities of the first ("weak") stage of the electric field and for different modes of switching this stage on, i.e., before the laser pulse (for brevity, mode A) or after the laser pulse (mode B). The polarization of pumping laser radiation was also varied. The unknown polarization of microwave field E was determined by the spectra of microwave transitions using the known angular parts of dipole moments of transitions between the S and P states [13]. It turned out to be linear, and the field contained both the component $\mathscr{E}_\sigma$ (orthogonal to the electric field) and the component $\mathscr{E}_\pi$ (collinear with the field), with $\mathscr{E}_\sigma/\mathscr{E}_\pi \approx 2$. The radiation intensity was evaluated by the power broadening of resonances using the calculated values of the radial parts of the dipole moments of transitions [14] and amounted to approximately 10^{-5} W/cm^2.

Figure 5 gives records of the spectrum obtained with linear σ polarization of pumping laser radiation. The left-hand column of the records relates to the A mode when the electric field E is switched on 0.8 μs before the laser pulse, and the right-hand column relates to the B mode when the field is switched on with a delay of 0.2 μs after the laser pulse. The dotted lines indicate the calculated positions of resonances, and the numbers adjacent to them correspond to the resonance numbers in the table.

In the absence of the field, the spectra coincide and contain four components of transitions between unperturbed fine-structure levels (see Fig. 2). In the electric field, the spectrum in the general case contains nine components as a result of Stark splitting of levels with $J = 3/2$. All nine components were observed in the B mode at $E = 4.73$ V/cm (Fig. 5). With the same field strength, the spectral components 1, 2, and 6 are absent in the A mode. We will dwell on this in more detail.

As was already noted, the field of 4.73 V/cm is a strong field in which the L–S coupling is broken, and the transition probabilities must be calculated in the basis of *nlm* states. Consequently, in such a field, the selection rules for dipole transitions will likewise be defined by the quantum number m rather than by M. In the spectra at $E = 4.73$ V/cm, one can see how three groups of resonances are formed, namely, {1–2}, {3–4–5–6–7}, and {8–9}, which correspond to the groups of transitions (see Fig. 2) $\{36P_{m=0} \longrightarrow 37P_{|m|=1}\}$, $\{36P_{|m|=1} \longrightarrow 37P_{|m|=1}, 36P_{m=0} \longrightarrow 37P_{m=0}\}$, and $\{36P_{|m|=1} \longrightarrow 37P_{m=0}\}$. In the electric field, the microwave radiation may induce all three groups of the above-identified transitions, because its polarization is such that the intermediate state $37S_{m=0}$ of two-photon transition is associated both with the sublevels $m = 0$ and with the sublevels $|m| = 1$ of the P states. However, it is also necessary to take into account the selection rules in the case of excitation of Rydberg states by polarized laser radiation.

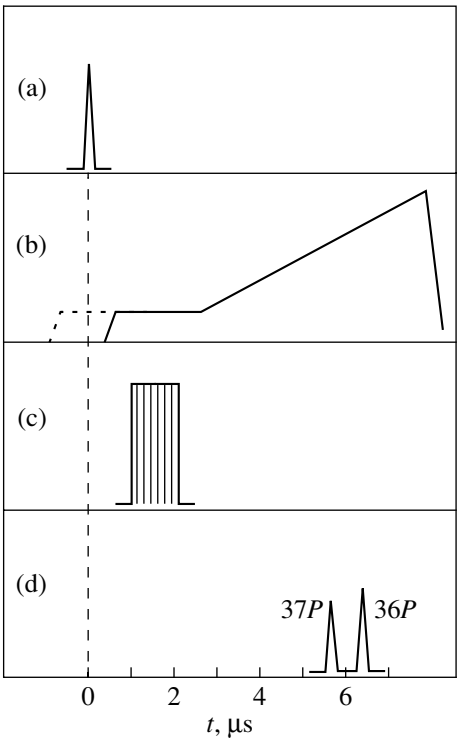

Fig. 4. The time diagram for signals. (a) The pulse of laser radiation exciting the initial $36P$ Rydberg state. (b) The electric field of two-stage shape. The first (weak) stage is switched on either before (A mode) or after (B mode) the laser pulse. The second (strong) linearly increasing stage is used for selective field ionization of Rydberg atoms. (c) The microwave radiation pulse inducing the $36P \longrightarrow 37P$ two-photon transition. (d) Signals at the output of VEU-6 channel secondary-emission multiplier corresponding to populations of Rydberg states.

If the excitation of the $4S \longrightarrow 36P$ transition in the case of σ polarization of laser radiation proceeds in the A mode, i.e., in the presence of an electric field, only the $36P_{|m|=1}$ sublevels will turn out to be excited, and the $36P_{m=0}$ sublevel is not populated, because the σ-polarized radiation induces transitions with $\Delta m = \pm 1$ from the initial state $4S_{m=0}$. It is this fact that results in the absence of the components 1, 2, and 6 in the A mode in the spectrum shown in Fig. 5 at $E = 4.73$ V/cm, because these components are associated with transitions from the $36P_{m=0}$ state.

If the excitation is effected in the B mode, the situation changes. Because, at the moment of a laser pulse, the electric field is absent, the selection rules with respect to M are valid; in this case, $\Delta M = \pm 1$ for σ-polarized radiation. As a result, all fine-structure magnetic sublevels of the $36P$ state are populated from the $4S_{J=1/2\,|M|=1/2}$ state. Then, after the adiabatic switching

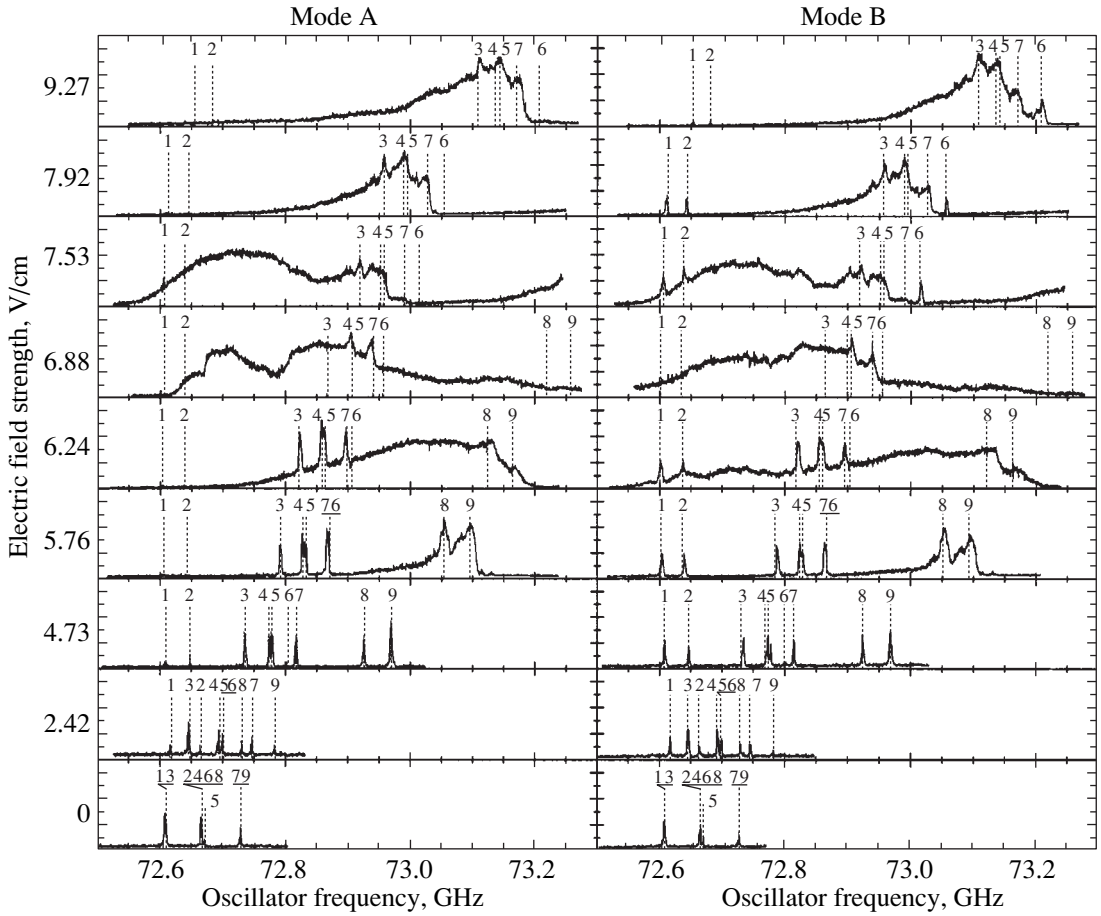

Fig. 5. The spectra of the $36P \longrightarrow 37P$ two-photon transition in Rydberg atoms of Na for different values of electric field strength and σ polarization of exciting laser radiation. The left-hand column of the records corresponds to the A mode when the electric field E is switched on 0.8 μs before the laser pulse, and the right-hand column relates to the B mode when the electric field is switched on 0.2 μs after the laser pulse.

on of the first stage of electric field, both the $36P_{|m|=1}$ sublevels and the $36P_{m=0}$ sublevel will turn out to be populated (see Fig. 2). Therefore, in Fig. 5, all nine components of two-photon transition are present in the B mode at $E = 4.73$ V/cm.

An analogous situation is observed in the case of $E = 2.42$ V/cm (Fig. 5), although no clear separation into three groups of resonances is observed so far. In the A mode, peaks 1 and 2 turn out to be several times lower than in the B mode. This indicates that even such a weak electric field is sufficient for the breaking of the L–S coupling in Rydberg states.

We will now turn to the field strength region $E = 6$–8 V/cm, where, according to the table, the double Stark resonance arises. The least critical strength is exhibited by peak 6. However, its amplitude is low compared with other peaks due to specific polarizations of laser and microwave radiation. Therefore, in the spectra recorded at $E = 5.76$ V/cm, the double Stark resonance is observed primarily at peaks 8 and 9, which shows up in their marked power broadening. The asymmetry of broadening is caused by a slight asymmetry in the microwave pulse spectrum due to the frequency deviation under conditions of pulse modulation of the G4-142 oscillator. This effect does not show up in the absence of double Stark resonance; however, in the region of double Stark resonance, the transition probabilities increase by several orders of magnitude. As a result, the intensity of even very weak spectral components of a microwave pulse is sufficient for the saturation of transition.

When the field strength increases to 6.88 V/cm, the double resonance includes other spectral components as well, as a result of which the spectrum assumes the form of a broad absorption band without clearly defined resonances. The formulas for evaluating the spectral width of double Stark resonance are given in [2]; for each of the resonances, this width turns out to be of the order of the Rabi frequency for exact resonance with single-photon transition (in our experiment, of the order of 100 MHz). In fields of 7.53 and 7.92 V/cm, a part of the transitions leave the double Stark resonance and recover their resonance pattern. Note the differences in the spectra of double Stark resonance for the A and B modes, which are associated only with the pres-

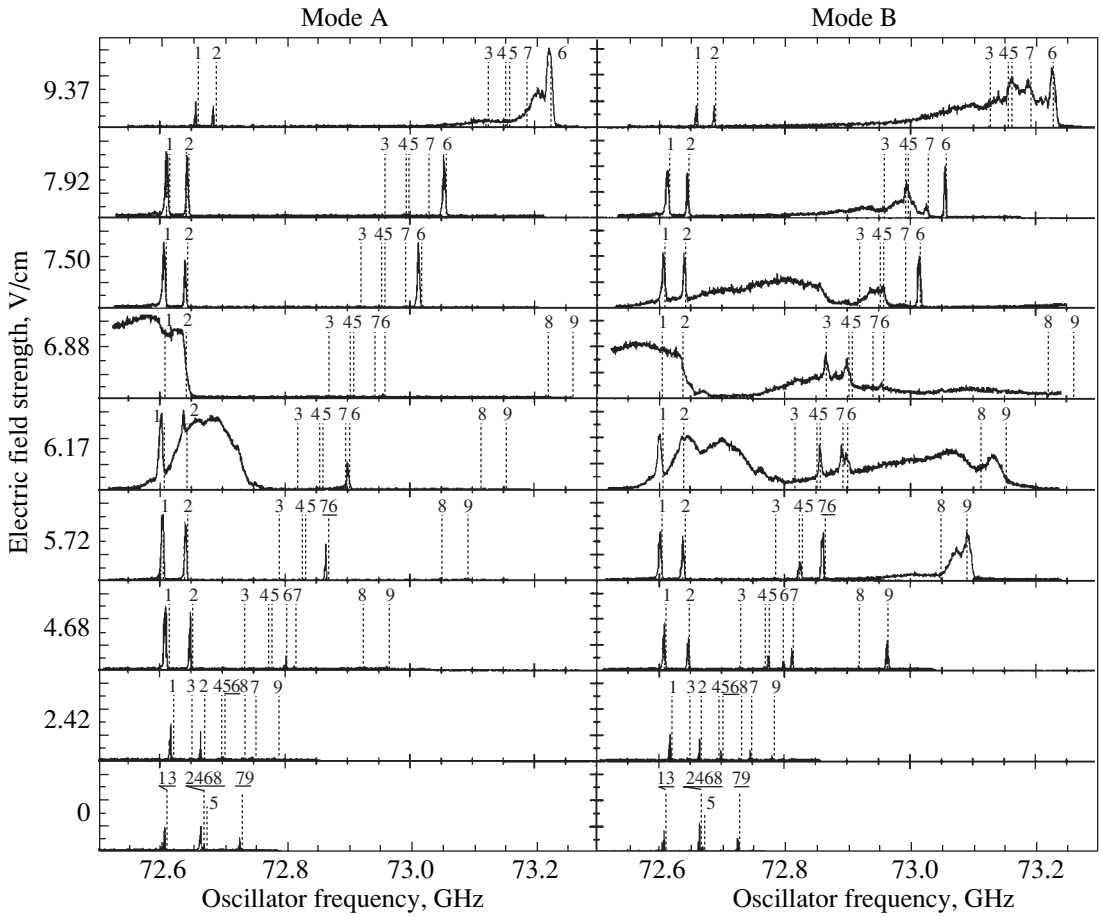

Fig. 6. Same as in Fig. 5, but for π polarization of exciting laser radiation.

ence or absence of an electric field at the moment of laser excitation.

A change in the polarization of laser radiation brings about an even more radical variation of the spectrum shape. Figure 6 gives spectral records obtained under conditions of π polarization. As before, the left-hand column corresponds to the A mode, and the right-hand column, to the B mode.

Now, in the A mode, the laser radiation excites only the $36P_{m=0}$ state due to the selection rule $\Delta m = 0$, and, in the B mode, the $36P_{J=3/2\,|M|=3/2}$ sublevels are not excited due to the selection rules $\Delta M = 0$. As a result, in Fig. 6 at $E = 4.68$ V/cm only three peaks are present in the A mode and six peaks in the B mode. Accordingly, in the region of double Stark resonance as well, the spectra appear much differently than in Fig. 5, especially, for the A mode. In particular, in the A mode at $E = 6.17$ and 6.88 V/cm, the broad absorption band disappeared from the center of the graph; however, the band in the left-hand part remained (it corresponds to double Stark resonance for peaks 1 and 2). For the A mode in the region of fields of 7.5–8 V/cm, no signal of double Stark resonance is observed at all; however, with a further increase in the strength to 9.37 V/cm, peak 6 starts broadening and entering into the double Stark resonance. The latter fact is unexpected, because the predicted value of the critical field for this peak is 6.37 V/cm.

5. DISCUSSION OF THE RESULTS

The results indicate that the observed spectrum of the $36P \longrightarrow 37P$ two-photon microwave transition in Rydberg atoms of Na depends significantly on the presence of the electric field at the moment of laser excitation. This effect is associated with the variation of the wave functions of Rydberg atoms in the electric field. In spite of the fact that, in an electric field, the spin–orbit interaction still shows up as the splitting of components of two-photon transition, it may be included as a perturbation for individual Stark sublevels [3], because the energy of interaction between an atom and electric field exceeds considerably the energy of L–S coupling. In this case, the transition probabilities are defined by the selection rules with respect to m rather than to M.

In this regard, a full analogy is observed with the transition from the Zeeman effect in a weak magnetic

field to the Paschen–Back effect in a strong magnetic field [3]. However, for the Paschen–Back effect to be observed, fields of the order of 10^5 Oe are required, which defines the complexity of experiments. At the same time, as was demonstrated by our experiment, the field strength of the order of just several V/cm is required for the breaking of the L–S coupling by an electric field in Rydberg atoms. This opens up new possibilities for investigation of the region of transition from weak interaction between an atom and electric field to strong interaction. Note that, for atoms in weakly excited states, the requisite fields reach values of hundreds of kV/cm or more.

The break of L–S coupling has a significant effect on the shape of double Stark resonance in Rydberg atoms. The disappearance of individual components of two-photon transition reduces the spectral width in the region of double resonance. This enables one to simplify the identification of transitions and, thereby, compare the experimental and theoretical values of the critical fields of double Stark resonance.

Note good agreement in all recordings in Figs. 5 and 6 between the observed frequencies of two-photon resonances and those calculated for the transitions which do not fit the exact double resonance and do not experience a strong power broadening. This points to the high accuracy of calculation of the level energies by the method of Zimmerman *et al.* [8]; therefore, the values of the critical fields of double Stark resonance given in the table must likewise correspond to the experimentally obtained values with a low intensity of microwave radiation, when the field broadening of resonances is minor. However, in our experiment, the single-photon Rabi frequency (of the order of 100 MHz) was comparable to the energy of L–S coupling. It is apparently this fact that brings about a marked difference between the experimentally and theoretically obtained values of the critical field for peak 6. Therefore, one must conclude that, in order to perform an exact calculation of the critical fields of double Stark resonance, it is necessary to simultaneously include both the spin–orbit interaction and the static and dynamic Stark effects.

ACKNOWLEDGMENTS

This study received financial support from the Russian Foundation for Basic Research (projects nos. 99-02-17131, 01-02-06507, 01-02-06508, and, in part, 00-02-17924).

REFERENCES

1. I. M. Beterov, G. L. Vasilenko, V. P. Kraĭnov, *et al.*, Pis'ma Zh. Tekh. Fiz. **17** (9), 44 (1991) [Sov. Tech. Phys. Lett. **17**, 331 (1991)].
2. I. M. Beterov, A. O. Vyrodov, I. I. Ryabtsev, and N. V. Fateev, Zh. Éksp. Teor. Fiz. **101**, 1154 (1992) [Sov. Phys. JETP **74**, 616 (1992)].
3. I. I. Sobelman, *Atomic Spectra and Radiative Transitions* (Nauka, Moscow, 1977; Springer-Verlag, Berlin, 1979).
4. V. A. Davydkin and B. A. Zon, Opt. Spektrosk. **52**, 600 (1982) [Opt. Spectrosc. **52**, 359 (1982)].
5. I. M. Beterov and I. I. Ryabtsev, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 853 (1998) [JETP Lett. **68**, 897 (1998)].
6. A. Khadjavi, A. Lurio, and A. Happer, Phys. Rev. **167**, 128 (1968).
7. M. S. O'Sullivan and B. P. Stoicheff, Phys. Rev. A **33**, 1640 (1986).
8. L. Zimmerman, M. G. Littman, M. M. Kash, *et al.*, Phys. Rev. A **20**, 2251 (1979).
9. S. F. Dyubko, M. N. Efimenko, V. A. Efremov, and S. V. Podnos, Kvantovaya Élektron. (Moscow) **22**, 946 (1995).
10. I. I. Ryabtsev and D. B. Tretyakov, Phys. Rev. A **64**, 033413 (2001).
11. R. A. Ambartsumyan, G. I. Bekov, V. S. Letokhov, and V. I. Mishin, Pis'ma Zh. Éksp. Teor. Fiz. **21**, 595 (1975) [JETP Lett. **21**, 279 (1975)].
12. *Rydberg States of Atoms and Molecules*, Ed. by R. F. Stebbings and F. B. Dunning (Cambridge Univ. Press, Cambridge, 1983; Mir, Moscow, 1985).
13. I. I. Ryabtsev and I. M. Beterov, Phys. Rev. A **61**, 063414 (2000).
14. S. P. Goreslavskiĭ, N. B. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **82**, 1789 (1982) [Sov. Phys. JETP **55**, 1032 (1982)].

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 685–693.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 797–806.
Original Russian Text Copyright © 2002 by Agafonov, Grigoryan, Znamenskiy, Manykin, Orlov, Petrenko, Sitnikov, Malakyan.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Effect of the Pump Absorption on Generation of Infrared Stimulated Electronic Raman Scattering in Alkali-Metal Vapors

A. I. Agafonov[a], G. G. Grigoryan[a], N. V. Znamenskiy[a], E. A. Manykin[a], Yu. V. Orlov[a], E. A. Petrenko[a], M. G. Sitnikov[a, *], and Yu. P. Malakyan[b]

[a]*Russian Research Centre Kurchatov Institute, Moscow, 123182 Russia*

**e-mail: zarja@isssph.kiae*

[b]*Institute of Physical Studies, Ashtarak-2, 378410 Armenia*

Received August 27, 2001

Abstract—The main properties of infrared stimulated electronic Raman scattering (SERS) at the $6\,^2S_{1/2}$–$7\,^2S_{1/2,\,3/2}$ transitions in cesium atoms are studied theoretically and experimentally as functions of the atomic concentration, which was varied from 10^{11} to 10^{16} cm^{-3}. It is found that the efficiency of generation of Stokes radiation strongly depends on one-photon absorption of the pump radiation tuned near frequencies of the $6^2S_{1/2}$–$7^2P_{1/2,\,3/2}$ transitions. By using the equation for the density matrix, which describes the evolution of a three-level system, the theory of resonance excitation of IR radiation upon one-photon absorption at an adjacent transition is developed. The theory describes well the main features of IR SERS in alkali-metal vapors. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Vapors of alkali metals, whose atoms have low-lying narrow energy levels coupled by strong dipole transitions, are very convenient media for exciting resonance stimulated electronic Raman scattering (SERS) of light, which allows one to convert comparatively simply and efficiently laser radiation in the visible range to the IR spectral region. For example, SERS in cesium vapors was used for conversion of the visible radiation from a dye laser to IR radiation tunable from 500 to 5000 cm^{-1} (2–20 μm) [1–3]. SERS in potassium vapors was used for obtaining IR radiation tunable from 2850 to 3500 cm^{-1} (2.85–3.5 μm), which was employed for studying the IR absorption spectrum of carbon dioxide [4]. We can also mention papers [5–9] in which picosecond laser pulses in the visible range were efficiently converted to the IR range using SERS in cesium and potassium vapors.

Theoretical estimates, which were confirmed by experimental results, show that the SERS efficiency should increase with increasing pump intensity I_L or upon approach of the pump frequency ω_L to the frequency of an allowed atomic transition. At the same time, the IR SERS under quasi-resonance conditions at high intensities I_L can be suppressed due to the development of other nonlinear-optical processes, for example, multiphoton ionization [10]. The quasi-resonance action of powerful electromagnetic radiation on an atomic system should also change the shape of an absorption line (see, for example, [11–14]). In this case, if the lifetime Γ^{-1} of one-photon or cascade radiative transition from an upper level to the ground state is small compared to the lifetime of stimulated emission of a Stokes photon of frequency ω_S, then one-photon absorption of pump radiation should dominate over SERS.

The perturbation of the absorption spectrum of a simplest two-level system in a strong quasi-resonance radiation field is well studied both theoretically and experimentally. The same can be said about the main properties of IR SERS excited near the resonance. However, the effect of absorption of pump radiation on the IR SERS under resonance conditions has not been studied in detail so far, although a decrease in the energy of Stokes radiation upon approach of ω_L to the resonance with frequencies of allowed atomic transitions was observed in some papers (see, for example, [2]), which was explained by absorption of pump radiation.

This paper is devoted to the experimental and theoretical study of the effect of absorption of pump radiation on the efficiency of IR SERS in alkali-metal vapors.

2. EXPERIMENTAL

We studied the IR SERS at the frequency ω_S by tuning the excitation frequency ω_L near frequencies of the $6^2S_{1/2}$–$7^2P_{1/2,\,3/2}$ transitions of a cesium atom (Fig. 1). Cesium vapors were produced in a special cell of length

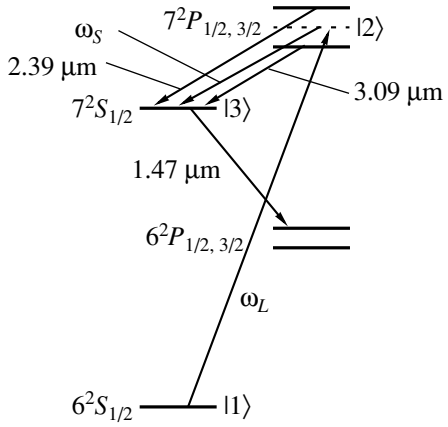

Fig. 1. Energy level diagram for a cesium atom.

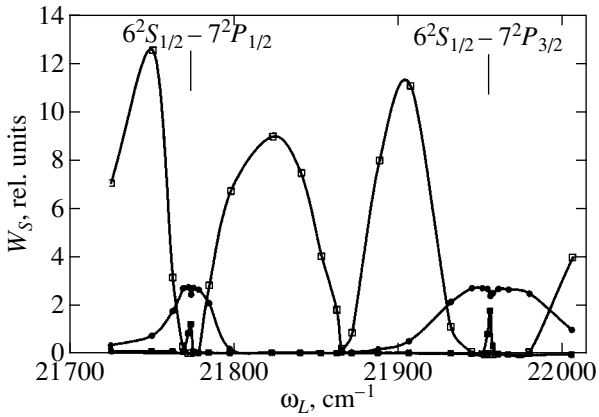

Fig. 2. Frequency dependences of the IR SERS energy at different concentrations of cesium atoms. (■) $N_1 = 8 \times 10^{11}$ cm^{-3}; (●) $N_2 = 4 \times 10^{13}$ cm^{-3}; (□) $N_3 = 5 \times 10^{15}$ cm^{-3}.

$l = 20$ cm, which was made of leucosapphire [15]. Several grams of cesium of purity 99.9% were introduced in a vacuum of 10^{-5} Torr into a cell placed inside a furnace. Due to a special furnace design, the temperature in the central part of the furnace was somewhat higher than that at its ends. This eliminated condensation on optical windows of the cell. The vapor pressure was varied from 10^{-5} to 1 Torr. This corresponded to a change in the concentration of cesium atoms in the range $N = 10^{11}$–10^{16} cm^{-3}.

Excitation was performed with a tunable dye laser, which was pumped by a pulsed excimer XeCl laser. The dye laser was continuously tuned in the spectral range $\omega_L = 21\,700$–$22\,050$ cm^{-1}, in which the frequencies of two $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transitions were located. The pulse duration of the dye laser was $\tau_L = 15$ ns, the peak pulse energy was $W_L = 5$ mJ, and the emission line width was $\Gamma_L/2\pi c = 1$ cm^{-1}.

The dye laser radiation was focused with a long-focus lens into a cell with cesium vapors so that the diameter of the laser beam was 5 mm at the cell entrance and 0.5 mm inside the cell at its center.

We studied simultaneously the dependences of the energy of visible pump radiation transmitted through cesium vapors and of the energy of the IR SERS signal on ω_L at different concentrations N of cesium atoms. The visible radiation coming from the cell was detected with a photodiode. The Stokes signal was focused with a spherical silver mirror on a liquid-nitrogen-cooled photoresistor, whose signal was amplified with a broadband amplifier mounted in the same housing.

Upon tuning ω_L near the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transition frequencies, along with IR SERS three other IR lines are excited at wavelengths 1.47, 2.39, and 3.09 μm, which correspond to the atomic transitions $6^2P_{1/2}$–$7^2S_{1/2}$ and $7^2S_{1/2}$–$7^2P_{1/2, 3/2}$ (Fig. 1). However, IR radiation at 2.39 and 3.09 μm was generated only when the frequency ω_L was coincident with an accuracy to the laser linewidth $\Gamma_L/2\pi c$ with the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transition frequencies, whereas the wavelength 1.47 μm lies outside the region of spectral photosensitivity of the photodetector. For this reason, in most experiments the Stokes radiation was directed to the photodetector without using an IR monochromator.

A system for processing output signals from the photodiode and photodetector also controls the dye laser radiation frequency ω_L. The system consists of CAMAC ADC and step-motor control modules and of a controller connecting the CAMAC line with a computer. The ADC operated in the regime of a peak detector. To avoid disturbances, the coincidence regime was used. A clock pulse from the excimer laser controlled the gate for transmitting a signal to the converter input only simultaneously with the laser pulse. The signal averaged over ten measurements was stored in the computer memory together with the current value of ω_L. Then, a command was sent for detuning the dye laser frequency, and the measurement cycle was repeated. The obtained results were processed and plotted using standard mathematical programs. The relative error of measurements did not exceed 3%.

3. EXPERIMENTAL RESULTS

Figure 2 shows the frequency dependences of the energy W_S of IR SERS in cesium vapors for three concentrations of cesium atoms: $N_1 = 8 \times 10^{11}$ cm^{-3}, $N_2 = 4 \times 10^{13}$ cm^{-3}, and $N_3 = 5 \times 10^{15}$ cm^{-3}. The laser radiation energy $W_L = 0.3$ mJ was constant. As expected, for $N_1 = 8 \times 10^{11}$ cm^{-3}, the IR signal energy reaches maximum values when ω_L is tuned to the resonance with frequencies of the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transitions. As the concentration of cesium atoms was increased, both the

value of the Stokes signal and the tuning region of ω_L where IR radiation was excited were increased. However, already at $N_2 = 4 \times 10^{13}$ cm^{-3}, the IR SERS energy decreased near the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transitions. A further increase in the concentration of cesium atoms up to $N_3 = 5 \times 10^{15}$ cm^{-3} resulted in a significant increase and broadening of holes near the resonances on the plot of the frequency dependence of the IR radiation energy.

Figure 3 shows the frequency dependences of the transmission coefficient of cesium vapors for laser radiation. The laser energy W_L and concentrations of cesium atoms N_1, N_2, and N_3 were as in the previous experiments. One can see that, at $N_1 = 8 \times 10^{11}$ cm^{-3}, ten percent of the pump radiation is absorbed only at exact resonances with the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transition frequencies. As the concentration of cesium atom increases up to 5×10^{15} cm^{-3}, absorption of laser radiation substantially increases, and the range of detunings $\Delta_0 = \omega_{21} - \omega_L$ (where ω_{21} is the frequency of the atomic transition $6^2S_{1/2}$–$7^2P_{1/2}$ or $6^2S_{1/2}$–$7^2P_{3/2}$) at which the pump radiation is strongly absorbed becomes broader. By comparing the plots shown in Figs. 2 and 3, we can conclude unambiguously that, when the pump radiation is strongly absorbed by cesium atoms, a hole appears in the plot of the frequency dependence of the IR SERS energy when ω_L approaches the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transition frequencies. The width and depth of the hole increase with increasing absorption of the pump radiation.

To optimize the excitation of IR SERS in cesium vapors near the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ transition frequencies, we studied the dependence of the Stokes emission energy on the vapor temperature at different pump energies. The results of these experiments for the case $\Delta_0 \approx 0$ obtained for W_L = 1, 0.3, and 0.04 mJ are presented in Fig. 4 (curves *1*, *2*, and *3*, respectively). One can see that I_S increases with increasing W_L. In this case, there exists an optimal value of the vapor temperature ($T \approx 130$°C, corresponding to the concentration of cesium atoms $N \approx 7 \times 10^{13}$ cm^{-3}) at which the efficiency of excitation of IR SERS was maximum. In our experiments, the excitation quantum efficiency could reach 40%.

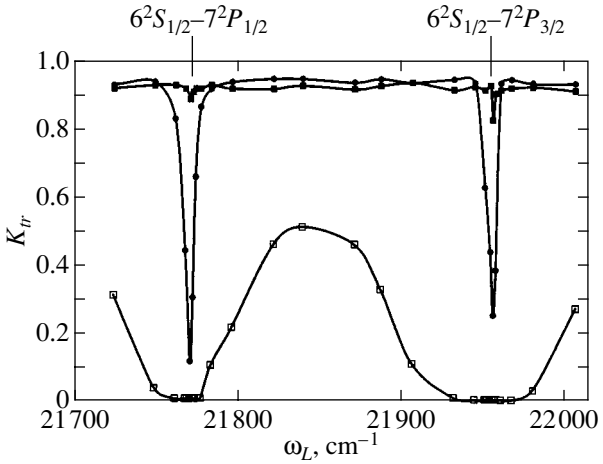

Fig. 3. Frequency dependences of the transmission coefficient of a cell with cesium atoms at different concentrations of cesium atoms: (■) $N_1 = 8 \times 10^{11}$ cm^{-3}; (●) $N_2 = 4 \times 10^{13}$ cm^{-3}; (□) $N_3 = 5 \times 10^{15}$ cm^{-3}.

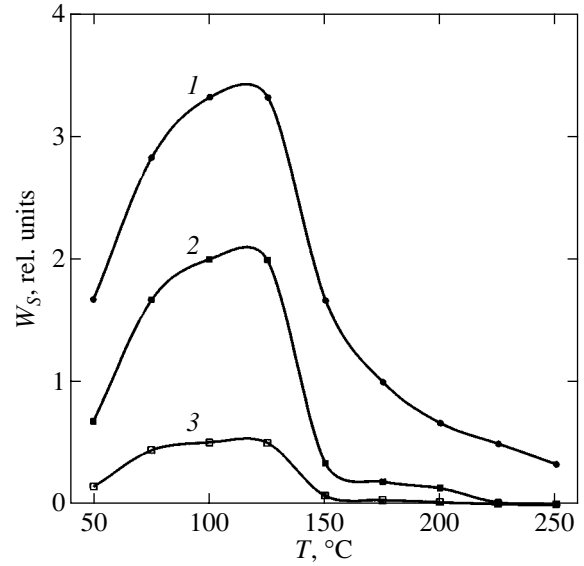

Fig. 4. Dependences of the IR SERS energy on the vapor temperature for $\Delta_0 \approx 0$ and pump energies W_L = 1 (*1*), 0.3 (*2*), and 0.04 mJ (*3*).

4. DISCUSSION OF RESULTS

We restrict the theoretical analysis of the results obtained to the case when the detuning Δ_0 of the pump frequency from the $6^2S_{1/2}$–$7^2P_{1/2, 3/2}$ atomic transition frequencies is small compared to the fine splitting of the 7P term. In this case, we can use the model of a three-level atom with the ground ($6^2S_{1/2}$), intermediate ($7^2P_{1/2}$ or $7^2P_{3/2}$), and final ($7^2S_{1/2}$) levels, which we denote below as $|1\rangle$, $|2\rangle$, and $|3\rangle$. This is the well-known Λ system, which was actively studied in recent years in connection with the atomic coherence effects (see, for example, [16]). However, these effects are manifested under the conditions that are opposite to those at which SERS takes place, and, therefore, we will not discuss them.

Consider the pump and IR radiation pulses with frequencies ω_L and ω_S propagating along the z axis in a medium consisting of three-level atoms under conditions when the pump field resonantly interacts with the atoms only at the $|1\rangle$–$|2\rangle$ transition and the IR radiation

field interacts with the atoms at the $|2\rangle$–$|3\rangle$ transition. We represent the fields in the form

$$\mathbf{E}_i = \mathbf{e}_i E_i(z, t)\exp(ik_i z - i\omega_i t), \quad (1)$$

where $E_i(z, t)$ are slowly varying complex amplitudes and k_i are the wave vectors of the fields, $i = L, S$.

In the resonance approximation, the interaction Hamiltonian of the system has the form

$$H_{\text{int}} = \hbar\Delta_0|2\rangle\langle 2| + \hbar(\Delta_0 - \Delta_1)|3\rangle\langle 3| + (\mu_{21}E_L|2\rangle\langle 1| + \mu_{23}E_S|2\rangle\langle 3| + \text{H.c.}), \quad (2)$$

where μ_{ik} is the dipole matrix element of the $i \longrightarrow k$ transition, $\Delta_0 = \omega_{21} - \omega_L$ and $\Delta_1 = \omega_{23} - \omega_S$ are the detunings of the fields from the corresponding atomic transitions, and ω_{ik} is the difference in frequencies of levels i and k.

In the general case, the IR radiation consists of two components: Stokes radiation, which is generated during SERS, and stimulated emission at the atomic $|2\rangle$–$|3\rangle$ transition, which is amplified due to the population inversion for levels $|2\rangle$ and $|3\rangle$. In the case of SERS, the frequency of IR radiation $\omega_S = \omega_L - \omega_{31}$ changes with changing the pump frequency, whereas the stimulated emission is generated at the fixed frequency ω_{23} because in this case the transition of an atom to the state $|3\rangle$ occurs from the real level 2, which is populated due to atomic collisions and due to the interaction of atoms with broadband pump radiation. It is obvious that, for $\Delta \le \Gamma_L$, where Γ_L is the pump line width, both radiations have the same frequency ω_{23} and are described by the same field E_S. In the case of large detunings ($\Delta \gg \Gamma_L$), they are already generated at different frequencies, and in principle we should take into account in Hamiltonian (3) the contributions from both fields simultaneously, with different amplitudes. However, as follows from experimental data [17], which were obtained under similar conditions in thallium vapors, the stimulated emission is weaker by more than an order of magnitude compared to SERS already for $\Delta > 3\Gamma_L$ and the pump intensity $I_L \ge 1$ MW/cm^2. For this reason, we will assume below that, when $\Delta > \Gamma_L$, only Stokes emission is generated at the frequency $\omega_S = \omega_L - \omega_{31}$, and we will set $\Delta_1 = \Delta_0 = \Delta$ in (3).

The time evolution of the system is described by the following equation for the density matrix ρ:

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + \Lambda\rho, \quad (3)$$

where Λ is the relaxation matrix. In our case, the longitudinal relaxation is determined by the spontaneous decay of levels $|2\rangle$ and $|3\rangle$ and two-photon resonance ionization. The width Γ_{ion} can be estimated from the expression

$$\Gamma_{\text{ion}} = n_{\text{ph}}\sigma_{\text{ion}},$$

where n_{ph} is the photon flux in the pump pulse and σ_{ion} is the cross section for one-photon ionization from the level $|3\rangle$. It follows from this that, for $n_{\text{ph}} \sim 10^{23}$–$10^{24}$ photon/(cm^2 s), which corresponds to energies 0.1–1 mJ of a focused laser beam with the cross-sectional area $S_0 \sim 10^{-3}$ cm^2, and for standard values of $\sigma_{\text{ion}} \approx 10^{-18}$–$10^{-17}$ cm^2, we have $\Gamma_{\text{ion}} \ll \tau_L$, where τ_L is the laser pulse duration. The radiative widths of the levels $7^2P_{1/2, 3/2}$ and $7^2S_{1/2}$ are also small compared to τ_L^{-1}, so that the longitudinal relaxation can be neglected below. The transverse relaxation is mainly determined by the resonance collision broadening with energy transfer caused by the intrinsic vapor pressure. As shown in [18, 19], the line width caused by one-photon resonance collisions is

$$\Gamma_{\text{col}} \approx 0.021N\lambda^3\gamma_{21}, \quad (4)$$

where λ is the wavelength and γ_{21} is the natural line width of the $|2\rangle$–$|1\rangle$ transition, and N is the vapor density. According to (4), even at high densities, when $N \sim 10^{16}$ cm^{-3}, the value of $\Gamma_{\text{col}}/2\pi c$ is only $\approx 10^{-2}$ cm^{-1}, which is far less than the pump line width $\Gamma_L/2\pi c \approx 1$ cm^{-1}. The Doppler broadening $\Gamma_D/2\pi c \approx 10^{-2}$ cm^{-1} is also small compared to Γ_L, which makes the averaging of final results over atomic velocities unnecessary. Therefore, only the phase modulation of the pump field should be taken into account in Eq. (3). By representing the field $E_L(z, t)$ in the form

$$E_L(z, t) = |E_L(z, t)|\exp[-i\varphi(t)],$$

we assume that the phase modulation $\varphi(t)$ is a process proceeding randomly in time t with the correlation function

$$\langle \exp[\varphi(t)]\exp[-i\varphi(t')]\rangle = \exp[-\Gamma_L(t - t')]. \quad (5)$$

It is well known [20, 21] that, in the region of a large gain, the phase of a Stokes wave follows the phase of the pump field, so that we can write

$$E_S(z, t) = |E_S(z, t)|\exp[-i\varphi(t)].$$

Then, after the changes

$$\rho_{21}(z, t) = \rho_{21}\exp[-i\varphi(t)]$$

$$\rho_{23}(z, t) = \rho_{23}\exp[-i\varphi(t)]$$

and the averaging of Eq. (3) over fluctuations of the phase $\varphi(t)$, taking (5) into account, we obtain the system of equations for the elements ρ_{ij} of the density matrix

$$\dot\rho_{11} = 2\Omega_L \text{Im}\rho_{21}, \quad (6)$$

$$\dot\rho_{33} = 2\Omega_S \text{Im}\rho_{23}, \quad (7)$$

$$\dot\rho_{31} = i\Omega_L\rho_{32} - i\Omega_S\rho_{21}, \quad (8)$$

$$\dot\rho_{21} = (-i\Delta + \Gamma_L)\rho_{21} - i\Omega_L(\rho_{11} - \rho_{22}) - i\Omega_S\rho_{31}, \quad (9)$$

$$\dot{\rho}_{23} = (-i\Delta + \Gamma_L)\rho_{23} - i\Omega_S(\rho_{33} - \rho_{22}) - i\Omega_L\rho_{13}, \quad (10)$$

$$\rho_{22} = 1 - \rho_{11} - \rho_{33}, \quad (11)$$

which are solved with the initial values $\rho_{ij} = \delta_{i1}\delta_{j1}$. Here, the real Rabi frequencies for fields $E_i(z, t)$ ($i = L, S$) are defined as $\Omega_L = \mu_{21}|E_L|/\hbar$ and $\Omega_S = \mu_{23}|E_S|/\hbar$.

Taking into account that $\Gamma_L \gg \tau_L^{-1}$, we can omit derivatives with respect to t in Eqs. (9) and (10) and find

$$\rho_{21} = -\frac{i}{\Delta - i\Gamma_L}[\Omega_L(\rho_{11} - \rho_{22}) + \Omega_S\rho_{31}], \quad (12)$$

$$\rho_{23} = -\frac{i}{\Delta - i\Gamma_L}[\Omega_S(\rho_{33} - \rho_{22}) + \Omega_L\rho_{13}]. \quad (13)$$

The propagation of the pump and IR emission fields is described by the Maxwell equations

$$\frac{d}{dz}|E_L(z, t)| = \frac{2\pi\omega_L}{c}N\mu_{21}\mathrm{Im}\rho_{21}(z, t), \quad (14)$$

$$\frac{d}{dz}|E_S(z, t)| = \frac{2\pi\omega_S}{c}N\mu_{23}\mathrm{Im}\rho_{23}(z, t). \quad (15)$$

It is convenient to pass from (14) and (15) to equations for $\Omega_i^2(z, t)$, which have the form

$$\frac{d}{dz}\Omega_L^2(z, t) = \alpha\Omega_L^2(z, t)\mathrm{Im}\left[\frac{\Gamma_L}{\Omega_L}\rho_{21}(z, t)\right], \quad (16)$$

$$\frac{d}{dz}\Omega_S^2(z, t) = \alpha g\Omega_S^2(z, t)\mathrm{Im}\left[\frac{\Gamma_L}{\Omega_S}\rho_{23}(z, t)\right], \quad (17)$$

where

$$\alpha = \frac{4\pi N\omega_L\mu_{21}^2}{\hbar c\Gamma_L}, \quad g = \frac{\omega_S\mu_{23}^2}{\omega_L\mu_{21}^2}.$$

We assume that IR emission is generated from the spontaneous noise, whose intensity is proportional to the pump intensity at the entrance to the medium. This intensity is written in the form

$$I_L(z = 0, t) = I_L f(t), \quad (18)$$

where the dimensionless function $f(t)$ determines the temporal shape of the pump pulse of duration τ_L. Without the loss of generality, we can assume that the pump and IR emission pulses have the same temporal shape over the entire region of interaction, i.e.,

$$\Omega_i^2(z, t) = \Omega_i^2(z) f(z, t), \quad i = L, S, \quad (19)$$

where $f(z, t) = f(z = 0, t)$. This approximation is justified by the fact that the time dependence of the field $E_S(z, t)$ in (15) is determined by the function $\rho_{23}(t)$, while this function according to (13) changes adiabatically as the pump field. In addition, this approximation allows us to significantly simplify calculations because, provided (19) is valid, Eqs. (6)–(8), (11) have an analytic solution if we introduce the new variable

$$\theta(z, t) = \int_{-\infty}^{t} f(z, \tau)d\tau.$$

Because these solutions are described by rather cumbersome expressions, we present them only for $\mathrm{Im}\rho_{21}$ and $\mathrm{Im}\rho_{23}$:

$$\mathrm{Im}\left[\frac{\Gamma_L}{\Omega_L}\rho_{21}\right] = -\frac{\Gamma_L^2}{V(\Delta^2 + \Gamma_L^2)} \times\left\{U_L\exp(-4V\theta) + U_S\exp(-V\theta)\right. \quad (20)$$
$$\left.\times\left[\cos\left(\frac{V\theta\Delta}{\Gamma_L}\right) + \frac{\Delta}{\Gamma_L}\sin\left(\frac{V\theta\Delta}{\Gamma_L}\right)\right]\right\},$$

$$\mathrm{Im}\left[\frac{\Gamma_L}{\Omega_S}\rho_{23}\right] = \frac{U_L\Gamma_L^2}{V(\Delta^2 + \Gamma_L^2)}\exp(-V\theta) \times\left[\cos\left(\frac{V\theta\Delta}{\Gamma_L}\right) + \frac{\Delta}{\Gamma_L}\sin\left(\frac{V\theta\Delta}{\Gamma_L}\right) - \exp(-3V\theta)\right], \quad (21)$$

where

$$U_i(z) = \frac{\Omega_i^2(z)}{\Delta^2 + \Gamma_L^2} \quad (i = L, S),$$

$$V(z) = U_L(z) + U_S(z).$$

We measured in our experiments the output radiation energy. Therefore, it is reasonable to pass from (20) and (21) to the equations for the quantities

$$W_i(z) = \frac{\Omega_i^2(z)\theta_\infty(z)}{\Gamma_L}, \quad i = L, S,$$
$$\theta_\infty(z) = \int_{-\infty}^{\infty} f(z, t)dt, \quad (22)$$

which, after multiplication by the cross-sectional area S_i ($i = L, S$), coincide with an accuracy to constants with the energies of the pump and IR radiation pulses. However, in the case of the pump pulse, care should be required because the quantity S_L becomes a function of the distance z due to the laser beam focusing, and it is obvious that even in the absence of active ions the pump intensity is inversely proportional to the beam cross section $S_L(z)$. One can easily verify that this dependence in the equation for $W_L(z)$ is described by the additional term $-(W_L(z)/S_L)dS_L/dz$. By substituting (20) and (21) into Eqs. (16) and (17) and integrating over

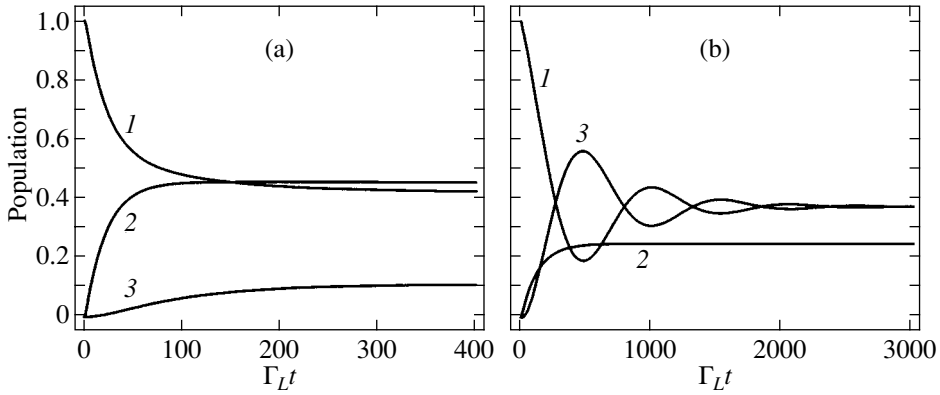

Fig. 5. Dependences of atomic level populations on time (normalized to $\Omega_L = 0.1\Gamma_L$) for (a) $\Delta = 0$, $\Omega_L = 0.1\Gamma_L$ (corresponding to the pump intensity of about 0.1 MW/cm^2), $\Omega_S = 0.01\Gamma_L$ and (b) $\Delta = 3\Gamma_L$, $\Omega_L = \Omega_S = 0.1\Gamma_L$, $\Gamma_L/2\pi c = 1$ cm^{-1}. The numbers at the curves denote the atomic levels.

time, we obtain finally the following equations for $W_i(z)$:

$$\frac{d}{dz}W_L(z) = -\alpha\frac{W_L}{4(W_L+W_S)^2}$$

$$\times\left\{W_L[1-\exp(-4I(z))]\right. \tag{23}$$

$$\left.+4W_S\left[1-\exp(-I(z))\cos\left(\frac{\Delta}{\Gamma_L}I(z)\right)\right]\right\}-\frac{W_L}{S_L}\frac{d}{dz}S_L,$$

$$\frac{d}{dz}W_S(z) = \alpha g\frac{W_L W_S}{4(W_L+W_S)^2}\left\{3+\exp(-4I(z))\right. \tag{24}$$

$$\left.-4\exp(-I(z))\cos\left(\frac{\Delta}{\Gamma_L}I(z)\right)\right\},$$

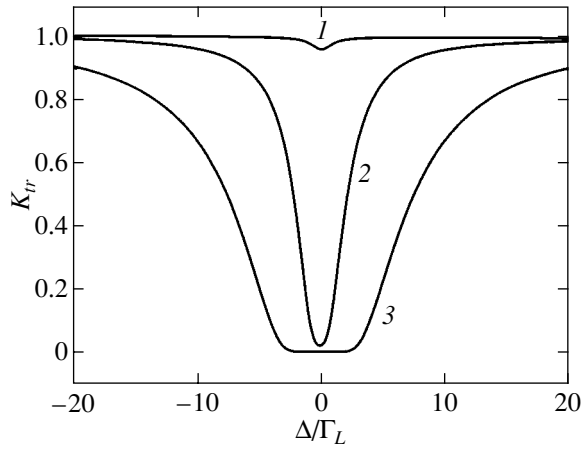

Fig. 6. Transmission coefficient of a cell with cesium vapors as a function of the detuning Δ at the concentration of cesium vapors $N = 10^{12}$ (*1*), 10^{14} (*2*), and 10^{15} cm^{-3} (*3*).

where

$$I(z) = \frac{[W_L(z)+W_S(z)]\Gamma_L^2}{\Delta^2+\Gamma_L^2}.$$

Equations (23) and (24) can be integrated numerically with the given initial values of $W_i(0)$ when the function $S_L(z)$ is known. The value of $W_L(0)$ can be readily found knowing the initial energy and the laser beam diameter. For IR emission, which is generated from spontaneous noise, $W_S(0)$ is written in the form

$$W_S(0) = \frac{C\Gamma_L^2}{\Delta^2+\Gamma_L^2}W_L(0), \tag{25}$$

where the coefficient C contains unknown parameters such as the cross-sectional area of the IR pulse and the solid angle of scattering. At the same time, there is no need to know the exact value of C because this coefficient determines only the relative value of the generated IR-emission energy. The latter is more sensitive to the ratio of matrix elements $g \sim \mu_{23}^2/\mu_{21}^2$, as follows from Eq. (24). Therefore, g can be used as a fitting parameter when comparing the theoretical results with experimental data. We will integrate numerically Eqs. (6)–(11) and (23), (24) using parameters corresponding to the $6^2S_{1/2}$–$7^2P_{1/2}$ transition in a cesium atom. In this case, the absorption coefficient for pump radiation for $N = 10^{12}$ cm^{-3} is $\alpha \approx 10^{-2}$ cm^{-1}. The cross-sectional area $S(z)$ of the pump pulse as a function of z is approximated taking into account that the laser beam is focused at the center of the cell of length $l = 20$ cm and has the diameter $d \approx 0.5$ cm at the entrance to the cell and $d \approx 0.05$ cm in the focal plane. For this reason, the pump intensity at small distances is so low that no IR emission is generated. The generation occurs only in the region of sharp focusing of the laser beam. We chose the length of this region in our calculations equal approximately to 2 cm.

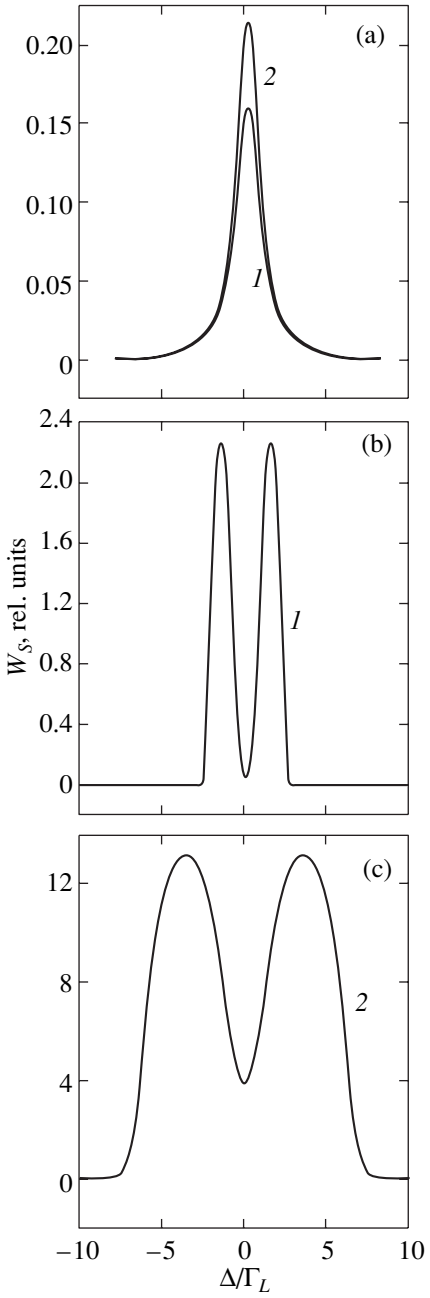

Fig. 7. Output IR emission energy as a function of the detuning Δ at $N = 10^{12}$ (a) and 10^{14} cm^{-3} (b, c) for the initial pump energy $W_L(0) = 0.04$ (curves *1*) and 0.3 mJ (curves *2*). The parameter g in Eq. (24) was chosen equal to 40.

Figure 5 shows the time dependence of the atomic level populations for different values of the field intensity and detuning Δ. One can see that, for the zero detuning $\Delta = 0$ and weak IR emission (Fig. 5a), the level populations rapidly become constant, the level $|3\rangle$ being populated weakly, while the levels $|1\rangle$ and $|2\rangle$ have approximately the same population. The reverse situation takes place for large detunings and strong amplification of the IR emission (Fig. 5b), when the

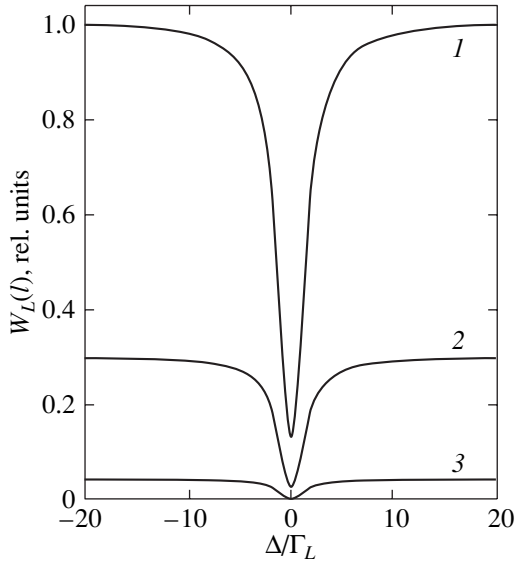

Fig. 8. Dependences of the pump energy on the detuning Δ for $z = l/2$ (where $l = 20$ cm is the cell length) at $N = 10^{14}$ cm^{-3} for the initial pump energy $W_L(0) = 1$ (*1*), 0.3 (*2*), and 0.1 mJ (*3*).

Rabi frequency Ω_S becomes equal to Ω_L (here, one should keep in mind that $\mu(7^2S_{1/2}–7^2P_{1/2}) \gg \mu(6^2S_{1/2}–7^2P_{1/2})$. In this case, the time oscillations of the populations of state $|1\rangle$ and $|3\rangle$ are observed, which represent the Rabi oscillations upon the two-photon $|1\rangle$–$|3\rangle$ transition. It is obvious that these oscillations should also be manifested in the frequency dependence of the IR emission (see Fig. 10 below). Note the Rabi oscillations are absent when $\Delta \le \Gamma_L$ because of a strong dephasing of atomic polarizations at the $|1\rangle$–$|2\rangle$ and $|2\rangle$–$|3\rangle$ transitions caused by the incoherence of the pump field. Recall that, under our experimental conditions, $\Gamma_L\tau_L \approx 4000$.

Figure 6 shows the spectrum of transmitted pump radiation for the input energy of 300 μJ and different vapor densities. In the region of an exact resonance, absorption and the width of the spectrum strongly increase with increasing N, in accordance with the experimental data. The dependence of the output IR-emission energy on the pump frequency is shown in Fig. 7. When $N = 10^{12}$ cm^{-3} (Fig. 7a), a sharp peak was observed at $\Delta = 0$ with the width of approximately Γ_L, its amplitude increasing with the pump energy. When the vapor density is increased up to $N = 10^{14}$ cm^{-3} (Figs. 7b, 7c), a hole appears at the center of the emission line, whose depth decreases with increasing input pump energy. The hole appears because at a high vapor density and small values of Δ the pump radiation is strongly absorbed already at small distances. For this reason, the generation of IR emission is strongly suppressed for small Δ and it is observed only for the values of Δ at which absorption of the pump radiation is weak. As the pump energy increases, it is still absorbed at small Δ. However, as one can see from Fig. 8, this

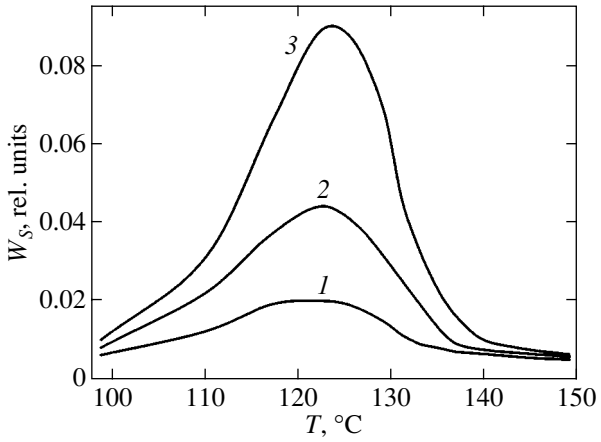

Fig. 9. Dependences of the IR emission energy on the vapor temperature for $\Delta = 0$ and the initial pump energy $W_L(0) =$ 0.1 (*1*), 0.3 (*2*), and 1 mJ (*3*); $g = 40$.

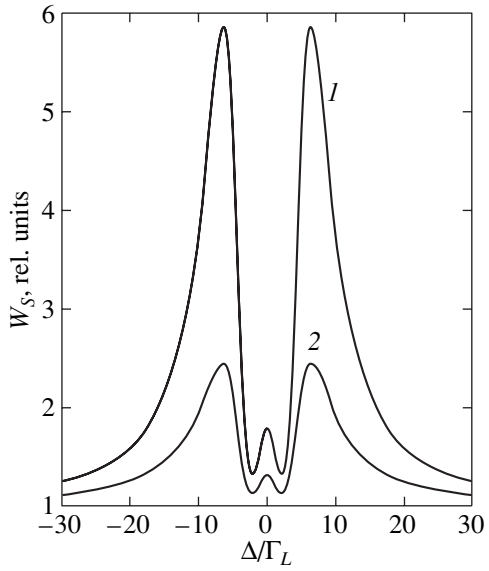

Fig. 10. Dependences of the output IR emission energy on the detuning Δ at $N = 10^{15}$ cm^{-3} for the initial pump energy $W_L(0) = 1$ (*1*) and 0.3 mJ (*2*) and $g = 40$. A weak maximum is observed at $\Delta = 0$.

absorption is no longer complete, which results in a decrease in the depth of the hole in the IR emission line. Figure 9 shows the dependences of the IR emission energy on the vapor temperature for $\Delta = 0$ and different pump energies. These results are also in good qualitative agreement with experimental data.

Finally, we discuss the possibility of observing the IR emission intensity oscillations, which we mentioned above. These oscillations are determined by the last term on the right-hand side of Eq. (24), which is proportional to $\cos(I(z)\Delta/\Gamma_L)$. Unfortunately, this effect is rather weak due to the smallness of the factor $\exp[-I(z)]$. In particular, when $N = 10^{15}$ cm^{-3}, a weak maximum of the IR emission appears at $\Delta = 0$ (Fig. 10), which is difficult to observe experimentally.

5. CONCLUSIONS

We have studied experimentally and theoretically the effect of one-photon resonance absorption of pump radiation on the IR SERS in cesium vapors. This effect strongly increases with increasing density of cesium atoms, resulting in a virtually complete suppression of the Stokes signal near the resonance. We have determined the optimal temperature of vapors at which the quantum efficiency of excitation of IR SERS is maximal and can be as high as 40%.

Based on the three-level atom model, we have developed the theory of generation of resonance IR SERS in alkali-metal vapors taking into account the absorption of pump radiation, which describes well the experimental properties of excitation of IR emission.

The results obtained can be used for developing compact, highly efficient laser sources based on resonance SERS in alkali-metal vapors, which generate powerful stimulated IR emission tunable in the near- and mid-IR ranges.

REFERENCES

1. D. Cotter, D. C. Hanna, and R. Wyatt, Opt. Commun. **16**, 256 (1976).
2. D. Cotter and D. C. Hanna, Opt. Quantum Electron. **9**, 509 (1977).
3. A. L. Harris and N. S. Levinson, Appl. Opt. **26**, 3996 (1987).
4. D. Cotter, D. C. Hanna, P. A. Karkainen, and R. Wyatt, Opt. Commun. **15**, 143 (1975).
5. D. G. Sarkisyan, A. A. Badalyan, S. O. Sapondzhyan, *et al.*, Kvantovaya Élektron. (Moscow) **13**, 872 (1986).
6. D. G. Sarkisyan, Kvantovaya Élektron. (Moscow) **15**, 2358 (1988).
7. R. Wyatt and D. Cotter, Opt. Commun. **32**, 481 (1980).
8. M. Berg and A. L. Harris, Opt. Lett. **9**, 50 (1984).
9. H. Ohde, S. Lin, A. Minoh, *et al.*, Appl. Phys. B **62**, 15 (1996).
10. D. Popescu, C. B. Collins, B. W. Johnson, and I. Popescu, Phys. Rev. A **9**, 1182 (1974).
11. B. R. Mollow, Phys. Rev. A **5**, 1522 (1972).
12. N. N. Kostin, M. P. Sokolova, V. A. Khodovoĭ, and V. V. Khromov, Zh. Éksp. Teor. Fiz. **62**, 475 (1972) [Sov. Phys. JETP **35**, 253 (1972)].
13. A. M. Bonch-Bruevich, S. G. Przhibel'skiĭ, V. A. Khodovoĭ, and N. A. Chigir', Zh. Éksp. Teor. Fiz. **70**, 445 (1976) [Sov. Phys. JETP **43**, 230 (1976)].
14. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979).

15. D. H. Sarkisyan, A. S. Sarkisyan, and A. K. Yalanusyan, Appl. Phys. B **66**, 241 (1998).

16. S. E. Harris, G. Y. Yin, M. Jain, *et al.*, Philos. Trans. R. Soc. London, Ser. A **355**, 2291 (1997).

17. M. G. Raymer and J. L. Carlsten, Phys. Rev. Lett. **39**, 1326 (1977).

18. A. P. Kazantsev, Zh. Éksp. Teor. Fiz. **51**, 1751 (1966) [Sov. Phys. JETP **24**, 1183 (1967)].

19. P. R. Berman and W. E. Lamb, Phys. Rev. **187**, 221 (1969).

20. S. A. Akhmanov, K. N. Drabovich, A. P. Sukhorukov, and A. S. Chirkin, Zh. Éksp. Teor. Fiz. **59**, 485 (1970) [Sov. Phys. JETP **32**, 266 (1971)].

21. R. L. Carman, F. Shimizu, C. S. Wang, and N. Bloembergen, Phys. Rev. A **2**, 60 (1970).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 694–703.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 807–818.
Original Russian Text Copyright © 2002 by Babin, Podivilov, Potapov, Churkin, Shapiro.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Nonlinear Resonance Induced by the Higher Spatial Coherence Harmonics

S. A. Babin*, E. V. Podivilov, V. V. Potapov, D. V. Churkin, and D. A. Shapiro

Institute of Automatics and Electrometry, Siberian Division, Russian Academy of Sciences, Universitetskiĭ pr. 1, Novosibirsk, 630090 Russia

**e-mail: babin@iae.nsk.su*

Received October 1, 2001

Abstract—The spectrum of a test field in a three-level ArII Λ scheme in the presence of a strong standing wave on the adjacent transition was measured. The known light-induced transparency peak was observed; the peak shifted as the detuning of the strong field was varied. In addition, a new resonance in the line center arose whose position was independent of the strong-field frequency. The resonance was caused by the higher spatial coherence harmonics on the test transition. Perturbation theory for a low and numerical calculations for a high standing wave intensity give qualitative agreement with experiment and substantiate the nature of the central transition. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Studies of nonlinear spectral resonances in the field of a strong standing wave have been under way since the 1960s. First, the resonance in the spectrum of spontaneous emission in a standing wave field was calculated [1]. The shape of the spectrum was substantially different from the simple sum of resonance structures induced by counterrunning waves that form a standing wave. This difference was caused by spatial standing wave field inhomogeneity responsible for the appearance of the higher spatial harmonics in both atomic level populations and medium polarization.

Calculations for a two-level system interacting with a standing wave [2, 3] showed that the contour of the Bennett dip in the velocity distribution of the difference of level populations was modulated by the higher spatial harmonics. After averaging over velocities and calculations of the Lamb dip shape in the spectrum, the modulation disappeared, and only a small change in the depth of the dip was observed. The effect was maximum near the center of the line, where the field experienced resonance interactions with particles most sensitive to spatial field inhomogeneity in the region of zero velocities. Atoms that occurred in standing wave antinodes effectively interacted with the field, and atoms in wave nodes virtually did not interact with it.

A consideration of the interaction between a strong running and a comparatively weak counterrunning wave and a two-level system with detuning between the frequency Ω and the resonance showed that Bennett structures arose in the velocity distribution of the difference of populations if

$$kv = \frac{\Omega}{2n+1}, \quad n = 0, 1, \ldots.$$

These Bennett structures were interpreted as multiphoton transitions with odd numbers of quanta $N = 2n + 1$ [4]. Multiphoton transitions with even numbers of quanta, $N = 2n$, result in the arising of a Bennett structure at zero velocities ($v = 0$).

When we pass from velocity distributions to observed test field spectra, the shape of resonances becomes more complex because of the influence of coherent processes. Spectral manifestations of multiphoton transitions were for the first time experimentally observed in [5]. Structures of the type of inverted Lamb dips were recorded in a CO_2 laser with an absorbing cell when the frequency of an external radiofrequency field was scanned. Alongside the principal peak, additional peaks of a complex shape appeared. These peaks were related to multiphoton processes. In [6], multiphoton resonances were also recorded in the test field spectrum of a two-level system which interacted with a strong standing wave. As in the preceding work, an absorbing cell with a molecular gas (CH_3F) was placed into the resonator cavity of a CO_2 laser, and the source of the test field was an additional tunable laser. Alongside the principal Bennett structures at $v = \pm\Omega/k$, where Ω is the detuning of the strong-field frequency, the test field spectrum contained odd subharmonics at velocities $v = \pm\Omega/3k$ and an even resonance at zero velocity $v = 0$. However, a comparison of the theory [7–9] with the experimental data [6] was only performed at a qualitative level because of the complex structure of experimental molecular spectra, on the one hand, and the impossibility of separating populational and coherent processes in the theory of two-level systems, on the other.

Satisfactory quantitative agreement between theory and experiment was attained in [10, 11] in studying an

atomic system (Cd^{113}) that interacted with two counter-running waves of different frequencies and amplitudes. When the frequency of one of the waves was scanned, a subradiation structure was recorded in the absorption spectrum. This structure contained up to five lines that converged to the line center. The structure of the spectra, especially their central part, was very sensitive to the ratio between the saturating and test wave amplitudes.

New possibilities for studying spectral resonances induced by a strong standing wave appear in three-level systems. Numerical calculations of the test field spectrum on the adjacent transition show [12] that higher spatial harmonic effects appear not only in the velocity distribution of level populations but also in the test field spectrum. The analysis performed in [12] also shows that the resonance of field splitting in the spectrum is observed for both Stokes and anti-Stokes test waves, as opposed to strong running field effects. What is more, we can separate populational and coherent processes in a three-level system, which simplifies an analysis of experimental spectra [8, 9] and offers possibilities for designing new experiments. In particular, it was suggested in [9] to suppress the contribution of coherent processes by dephasing collisions (or strong field phase fluctuations), which would allow us to observe purely populational resonances (multiphoton transitions and a peak of "slow" atoms). So far as we know, such measurements have not been performed yet.

In recent years, a three-level system with a strong standing wave tuned in resonance with one of the transitions was also studied to increase the efficiency of generation by molecular lasers with optical pumping in the far IR region. An increase in output power in the presence of a standing pumping wave compared with a running wave was predicted theoretically [13] and proved experimentally [14]. The effect of electromagnetically induced transparency in a closed three-level system with a standing saturating wave was also studied theoretically [15]. Under certain conditions, the use of a standing wave for medium bleaching is more effective than the use of running waves. In addition, transparency oscillations near the center of the transparency window appear.

In this work, we experimentally studied the spectrum of a test field in the presence of a strong standing wave on the adjacent transition in the three-level argon ion Λ scheme (Fig. 1). The conditions for separating out the field splitting effect, which is one of the fundamental coherent effects, were provided by selecting resonance medium parameters $k_\mu < k$, $N_l \gg N_m$, and N_n and $\Gamma_l \leq \Gamma_m \ll \Gamma_n$, where N_n, N_m, and N_l are the populations and Γ_n, Γ_m, and Γ_l are the relaxation constants of levels n, m, and l, respectively. The resonance structures and populations under the conditions of a gas discharge plasma were additionally broadened and suppressed by Coulomb collisions [16]. It was found experimentally that an additional resonance formed in the center of the test field spectrum line against the background of the usual split resonance as the standing wave intensity increased. This resonance looked like "splitting of field splitting" when the standing wave was tuned to the center of the line and transformed into an isolated peak in the center of the line when the standing wave was detuned from the resonance. We constructed an analytic model based on expanding the density matrix elements up to second-order terms with respect to intensity. At low intensities, this model coincided with the exact solution in the form of a continued fraction. We suggested a qualitative interpretation of the formation of the resonance as a result of spatial coherence modulation. We believe this result to be the first observation of the effect of the higher spatial coherence harmonics.

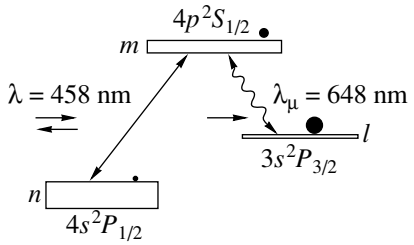

Fig. 1. Scheme of energy levels. Relative level widths are shown by rectangle heights, and level populations, by circle diameters.

This work is organized as follows. Section 2 describes the experimental unit and the procedure for measurements; Section 3 contains the most important experimental results; in Section 4, perturbation theory equations are obtained to qualitatively explain the experimental data; and, in Section 5, the shape of the resonance is calculated for arbitrary standing wave intensities. The experimental and theoretical results are compared and discussed in Section 6.

2. EXPERIMENTAL UNIT

The experimental unit for studying higher spatial harmonic effects in a standing wave field is shown in Fig. 2. The discharge tube of argon laser *1* was placed into a cavity with input *2* and output *3* mirrors opaque to generated radiation but transparent to test radiation. It follows that two waves were present in the resonator cavity, namely, the standing linearly polarized generated wave and the running test wave, also linearly polarized. Etalon *5* ensured the selection of one longitudinal mode and smooth tuning of generation frequency. Diaphragm *6* was used to select the TEM_{00} mode. Cavity mirrors selected the 457 nm line; for this line, transmission loss in the cavity was about 0.3%, which ensured high intensity within the cavity. Output argon laser radiation was directed by a mirror to grating *8*, from which one order was introduced into Fabry–Perot scanning interferometer *9* used to control the mode composition of radiation and determine detuning of the

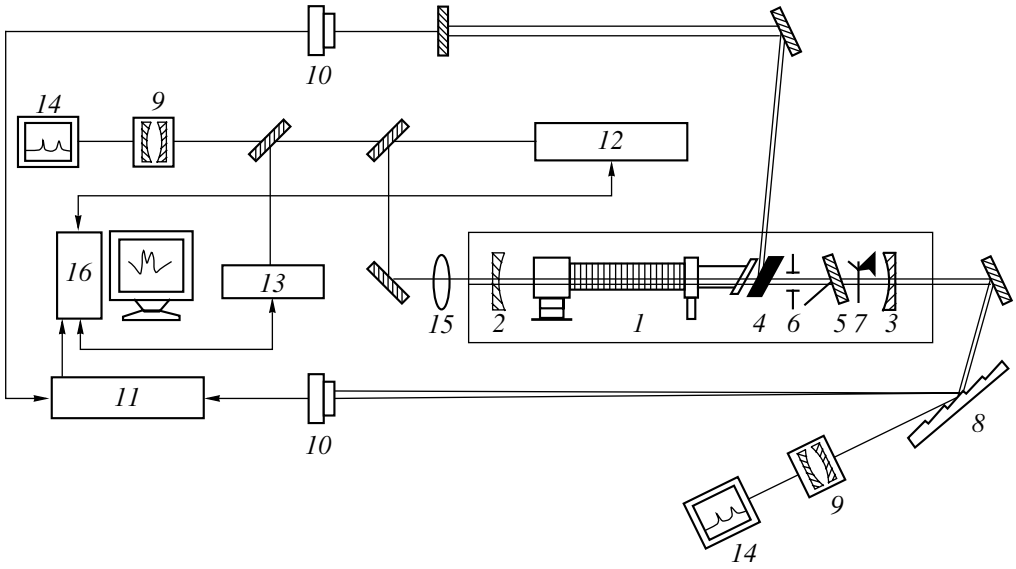

Fig. 2. Scheme of experimental unit: (*1*) discharge tube; (*2*, *3*) mirrors; (*4*) Brewster plate; (*5*) etalon; (*6*) diaphragm; (*7*) chopper; (*8*) grating; (*9*) scanning interferometer; (*10*) photodetector; (*11*) synchronous detector; (*12*) tunable dye laser; (*13*) wavelength meter; (*14*) oscilloscope; (*15*) lens; and (*16*) computer.

strong field from the resonance frequency. Another order was directed to photodetector *10* whose signal was the reference signal for synchronous detector *11*.

Measurements were performed by the test field method in the level scheme shown in Fig. 1. The parameters of the scheme taken from [17–20] are listed in the table. The main characteristic of measurements was a large difference of both the lifetimes of levels (the $\Gamma_l \lesssim \Gamma_m \ll \Gamma_n$ inequality was satisfied) and the level populations ($N_l \gg N_m \gtrsim N_n$, that is, the N_l unperturbed level was strongly populated).

The source of the test field was a tunable DCM 12 dye laser (see Fig. 2), whose generation wavelength was recorded by wavelength meter *13*. The automated system for frequency tuning and retuning [21] enabled us to smoothly vary the test field frequency in the frequency range up to 4.5 GHz in steps as small as 20 MHz; the step value was close to the radiation line width (about 10 MHz). The spectrum of the dye laser was recorded by scanning interferometer *9* with a 5-GHz free-dispersion region. The interferometer was connected to an oscilloscope used to control the mode composition of radiation.

Relaxation constants and level populations

Γ_n	Γ_m	Γ_l	A_{mn}	A_{ml}	units
300	15	8	9	1	10^7 s^{-1}

N_n	N_m	N_l	units
≈1	≈5	≈100	10^9 cm^{-3}

The test field that was introduced into the discharge tube was preliminarily focused by lens *15* to ensure maximum field uniformity in the cavity. After passing the discharge tube, a test field beam was reflected from additional plate *4* and introduced by mirrors to photodetector *10* connected to synchronous detector *11*. The angle between test field and generated radiation beams was about 10^{-3} radians, which prevented feedback.

We used synchronous detection, which allowed the Doppler backing to be automatically subtracted. For this purpose, the strong field was remodulated at a frequency of about 1 kHz with the use of light chopper *7*. Synchronous detection at the modulation frequency allowed us to identify nonlinear strong-field-induced additions. Experimental runs and synchronous data collection and recording were controlled by PC *16*, to which all measuring instruments were connected through an ADP.

3. RESULTS

A series of plots obtained for a 135 A discharge current is shown in Fig. 3. These plots illustrate the dependence of the test field spectrum on the strong-field intensity tuned to the center of the line. The first plot contains a characteristic structure caused by the field splitting effect at low intensities, namely, a split absorption contour with an about 0.5 GHz splitting. Under exact resonance conditions, we observe a transmission peak. Note that the negative values in the plot correspond to a strong-field-induced increase in test field absorption, and positive values correspond to a decrease in absorption, which is equivalent to light-induced transparency. The splitting increases as the

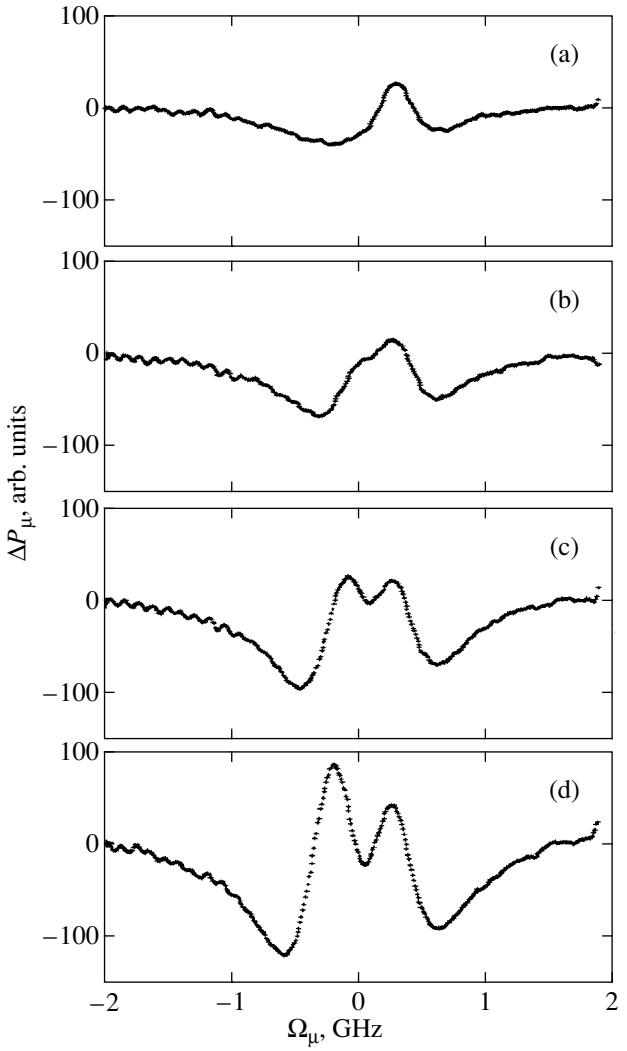

Fig. 3. Dependences of experimental spectra on strong-field intensity at $\Omega = 0$: $G \approx 25$ (a), 50 (b), 75 (c), and 100 MHz (d).

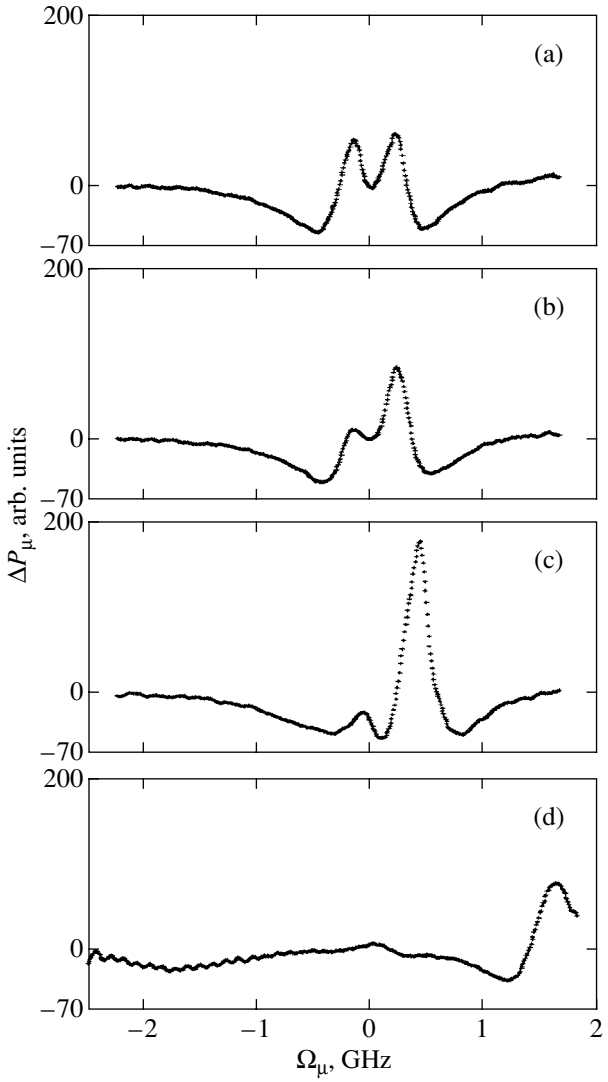

Fig. 4. Dependences of experimental spectra on strong-field detuning at $G = 100$ MHz: $\Omega \approx 0$ (a), 0.2 (b), 0.7 (c), and 2 GHz (d).

strong-field amplitude becomes larger; the transmission peak broadens; and, most interestingly, additional peak splitting, that is, "splitting of field splitting," appears. No such additional splitting is observed in the presence of a strong running wave. The split contour is slightly asymmetric. This asymmetry depends on intensity and is related to the nonlinear lens effect (e.g., see [22]); it can be compensated by slightly detuning the strong field from the resonance frequency.

A series of plots in Fig. 4 illustrate the dependence of the form of the nonlinear addition to the test field spectrum on strong field detuning (the intracavity strong field intensity was about 20 W/cm^2, which corresponded to $|G| \sim 100$ MHz). An increase in detuning to $\Omega \gtrsim 1$GHz causes frequency separation of the structures. At $\Omega \approx 2$ GHz (Fig. 4d), we only observe a broad populational dip at a frequency of $\Omega_\mu \approx -k_\mu\Omega/k \approx$ –1.5 GHz for negative detunings and a field-splitting structure with a transparency peak at $\Omega_\mu \approx k_\mu\Omega/k \approx$ 1.5 GHz for positive detunings. The characteristic special feature of these spectra is a field-splitting resonance shift observed synchronously as strong-field detuning increases. The resonance, however, remains in the center of the test field line. Its amplitude gradually damps, and it transforms from a dip into a peak at large strong-field detunings. The peak in the line center is quite discernible up to detunings of $|\Omega| \approx 2$ GHz $\sim kv_T/2$. When the sign of strong-field detuning changes, the picture of the test field spectrum (Fig. 4) changes to its reflection in the axis of ordinates. Note also that the structure in the line center is substantially (more than two times) narrower than the field-splitting resonance transparency peak.

4. PERTURBATION THEORY

Consider the interaction of atoms with an electromagnetic field in the three-level Λ scheme shown in Fig. 1. The field of light wave running along axis x,

$$\mathbf{E}_\mu(x, t) = \mathbf{E}_\mu \exp(ik_\mu x - i\omega_\mu t),$$

with detuning

$$\Omega_\mu = \omega_\mu - \omega_{ml}$$

from the Bohr frequency of the test transition tests the three-level system, which is under the action of the strong standing wave

$$\mathbf{E}(x, t) = 2\mathbf{E}\cos(kx)\exp(-i\omega t)$$

with an $\omega = \Omega + \omega_{mn}$ frequency close to the ω_{mn} frequency of the transition under consideration.

The absorption spectrum of the test field in an optically thin medium is expressed through the off-diagonal element of the density matrix $\rho_{ml} = \rho_\mu \exp(ik_\mu x - i\omega_\mu t)$,

$$P_\mu(\Omega_\mu) = 2\hbar\omega_\mu \mathrm{Re}(-iG_\mu^*)\int dv \int_0^L \frac{dx}{L}\rho_\mu, \qquad (1)$$

where

$$G_\mu = \frac{\mathbf{E}_\mu \cdot \mathbf{d}_{ml}}{2\hbar}$$

is the Rabi frequency, $\mathbf{d}_{ml}$ is the dipole moment of the transition, and L is the medium thickness in the direction of light propagation.

The special feature of our scheme of levels is a large N_l population of the level unperturbed by the strong field. This population is more than one order of magnitude higher than the populations at the working transition. We can therefore ignore all effects induced by the strong standing wave and related to changes in both populations and coherence at the working transition. The main nonlinear spectroscopic effect in such a system is related to mixing of coherences on both test ρ_μ and forbidden ρ_ν transitions by the standing wave field. This effect is described by the equations

$$\begin{aligned}
&(\Gamma_{ml} - i(\Omega_\mu - k_\mu v) + v\partial_x)\rho_v \\
&+ i(G_+ e^{ikx} + G_- e^{-ikx})\rho_v^* = iG_\mu N_l(v), \\
&(\Gamma_{nl} + i(\Omega - \Omega_\mu + k_\mu v) + v\partial_x)\rho_v^* \\
&+ i(G_+^* e^{-ikx} + G_-^* e^{ikx})\rho_\mu = 0.
\end{aligned} \qquad (2)$$

Here,

$$\rho_{ln} = \rho_\nu \exp(-ik_\mu x - i(\omega - \omega_\mu)t),$$

$$G_\pm = G = \frac{\mathbf{E}\cdot\mathbf{d}_{mn}}{2\hbar}$$

is the Rabi frequency of standing wave field components parallel and antiparallel, respectively, to the direction of test field propagation; Γ_{ml} and Γ_{nl} are the relaxation constants of test $(m - l)$ and forbidden $(n - l)$ transition coherences;

$$N_l(v) = \frac{N_l \exp(-v^2/v_T^2)}{\sqrt{\pi}v_T}$$

is the Maxwell velocity distribution of unperturbed level l particles; and ∂_x is the operator of the differentiation with respect to variable x.

System (2) has no finite analytic solution if the Rabi frequency of the standing wave is much larger than the relaxation constants. For this reason, the main tool of study is numerical calculations. However, at moderate fields, $G \le \Gamma_{ml}, \Gamma_{nl}$, the problem can be analyzed using perturbation theory. The main goal of such an analysis is to determine spectral resonances in test field absorption caused by spatial coherence modulation $\rho_\mu(x)$. Importantly, such a modulation can only arise when the three-level system simultaneously interacts with the standing wave component running in the same direction G_+ as the test field and the component running in the opposite direction G_-. Standing wave effects are therefore absent in first-order perturbation theory and can only appear starting with its second order with respect to standing wave intensities ($I^2 \propto |G_+|^2|G_-|^2$). Indeed, in first-order perturbation theory, the contribution to spectrum $P_\mu^{(1)}$ of interactions between the system and the G_- standing wave component vanishes in the Doppler limit. The nonlinear correction[1] for the G_+ component only appears in the Stokes case ($k_\mu < k$) as a field-splitting resonance [12] with a peak of light-induced transparency of width

$$\Gamma_p = \frac{k_\mu\Gamma_{nl} + (k - k_\mu)\Gamma_{ml}}{k}$$

at an $\Omega_\mu = k_\mu\Omega/k$ test field frequency,

$$\begin{aligned}
\Delta P_\mu^{(1)} &= P_\mu^{(1)}(0) - P_\mu^{(1)}(G) \\
&= 2\hbar\omega_h|G_\mu|^2 \frac{\sqrt{\pi}N_l \exp(-\Omega_\mu^2/k_\mu^2 v_T^2)}{kv_T} \\
&\times \mathrm{Re}\frac{2|G_+|^2(k - k_\mu)}{k(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2}.
\end{aligned} \qquad (3)$$

After solving (2) accurate to terms of fourth order in $G_\pm$ and integrating the result over velocities in the Doppler limit with the use of the theory of residues, we obtain the equation for the correction to the spectrum of

[1] The difference of the test field absorption coefficients in the presence and absence of the strong field, which was measured experimentally.

absorbed power in second-order perturbation theory [23]

$$\Delta P_\mu^{(2)} = -2\hbar\omega_\mu |G_\mu|^2 \frac{\sqrt{\pi} N_l \exp(-\Omega_\mu^2/k_\mu^2 v_T^2)}{k v_T}$$
$$\times \mathrm{Re} \frac{2|G_+|^2 (k-k_\mu)}{k(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2}$$
$$\times \left[\frac{3|G_+|^2 k_\mu (k-k_\mu)}{k^2(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2} \right.$$
$$+ \frac{|G_-|^2 (k-k_\mu)^2}{k^2(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))(\Gamma_{nl} - i(\Omega_\mu - \Omega))}$$
$$+ \frac{|G_-|^2 k_\mu^2 (2k-k_\mu)}{4k^3(\Gamma_{ml} - i\Omega_\mu)^2}$$
$$\left. + \frac{|G_-|^2 k_\mu^2 (k-k_\mu)}{k^3(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k)(\Gamma_{ml} - i\Omega_\mu))} \right]. \quad (4)$$

The first term in square brackets is the second term of the expansion in $|G_+|^2$ and describes an increase in field splitting proportional to $|G_+|$, that is, peak widths at $\Omega_\mu = k_\mu\Omega/k$ [see exact equation (5) given below]. The second term appears because of spatial coherence modulation on the ρ_ν forbidden transition and results in the arising in the spectrum of a nonlinear structure with forbidden transition width Γ_{nl} at the two-photon resonance frequency ($\Omega_\mu = \Omega$). Lastly, the two last terms owe their existence to spatial modulation of test transition coherence and describe the nonlinear structure of width Γ_{ml} strictly at the line center,[2] $\Omega_\mu = 0$.

In the Stokes case, both spatial modulation resonances arise against the background of a stronger field-splitting resonance. For this reason, they can only be well resolved under the condition $\Gamma_{nl} \ll \Gamma_p$ or $\Gamma_{ml} \ll \Gamma_p$. In the system that we study, the second inequality is fulfilled, $\Gamma_{ml} \ll \Gamma_p, \Gamma_{nl}$. The most important spatial modulation effect is therefore a narrow resonance in the line center described by the third term in square brackets in (4). At a zero standing wave field detuning ($\Omega = 0$), this resonance has the form of splitting Γ_{ml} of the field-splitting resonance of width Γ_p. The corresponding nonlinear test field power spectrum [the sum of Eqs. (3) and (4)] is shown in Fig. 5 for several Rabi standing wave frequencies. As field intensity increases, the dip in the center grows quadratically ($\propto I^2$), whereas the peak of width Γ_p that surrounds this dip increases linearly ($\propto I$).

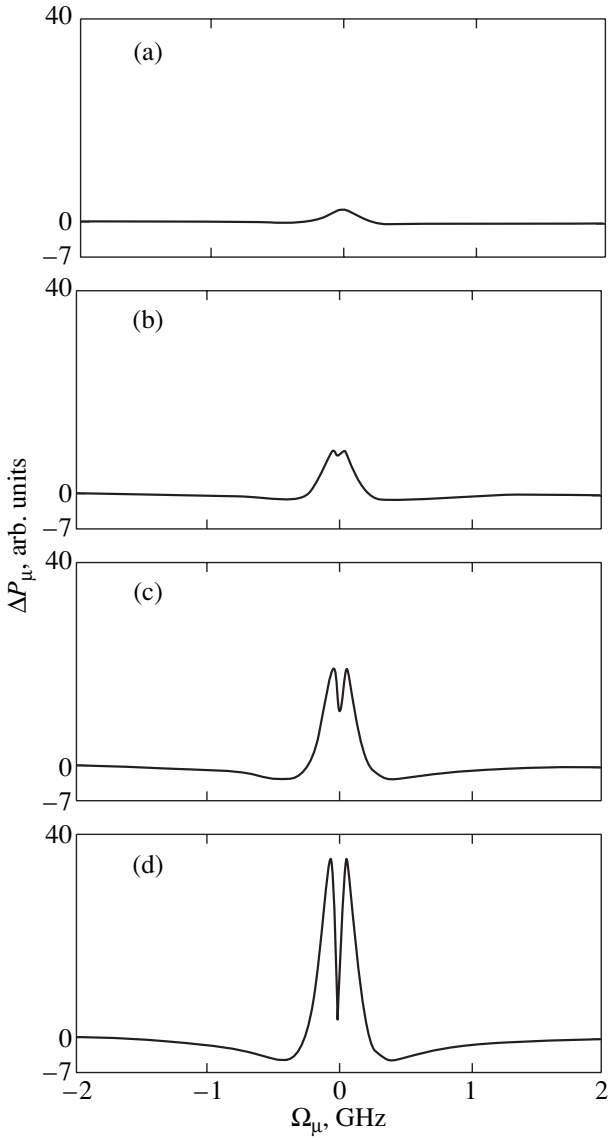

Fig. 5. Nonlinear resonance $\Delta P_\mu^{(1)} + \Delta P_\mu^{(2)}$ in test field spectrum as a function of strong-field amplitude: $G \approx 25$(a), 50 (b), 75 (c), and 100 MHz (d). Calculations were performed by (3) and (4) for $\Omega = 0$, $\Gamma_{nl} = 300$ MHz, and $\Gamma_{ml} = 50$ MHz.

The test field spectrum at large detunings, $|\Omega| \gg |G|$, is shown in Fig. 6 ($G \approx 100$ MHz). Narrow coherence resonance ρ_μ in the line center is quite discernible, and its amplitude decreases proportionally to $1/\Omega^2$ as detuning increases. As far as the main field-splitting resonance at the $\Omega_\mu = k_\mu\Omega/k$ frequency is concerned, this resonance is virtually insensitive to spatial modulations at large detunings. The nonlinear test field power spec-

[2] The possibility of the formation of a higher order resonance of width Γ_{ml} at the line center was not discussed in [23], because the opposite situation with $\Gamma_{nl} \ll \Gamma_{ml}$, when the effect was exceedingly small, was considered there.

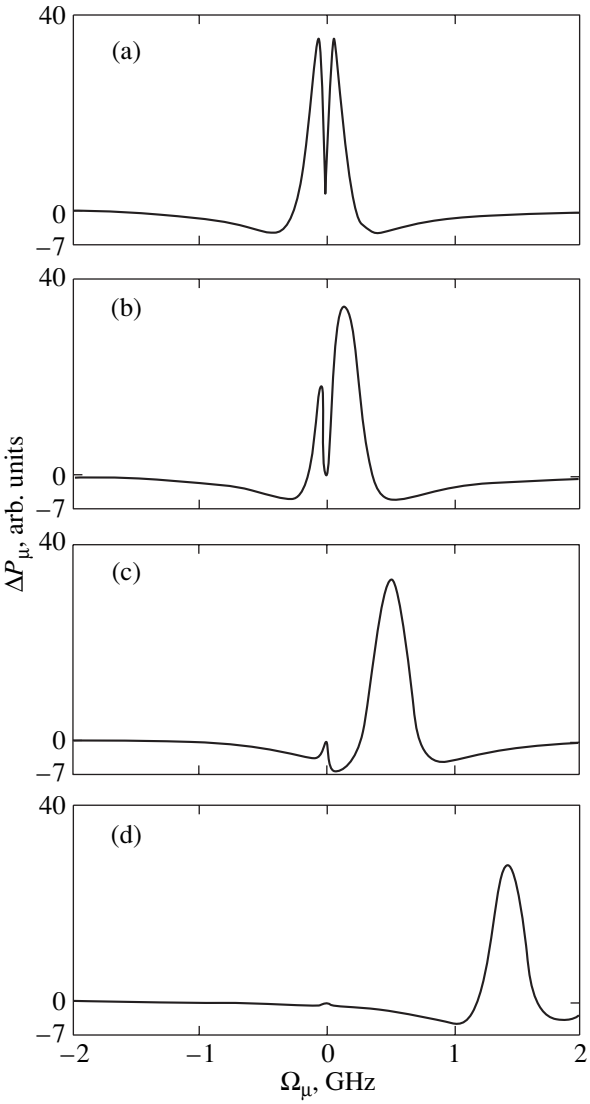

Fig. 6. Nonlinear resonance $\Delta P_\mu^{(1)} + \Delta P_\mu^{(2)}$ in test field spectrum as a function of strong field detuning: $\Omega \approx$ (a) 0, (b) 0.2, (c) 0.7, and (d) 2 GHz. Calculations were performed by (3) and (4) for G = 100 MHz, Γ_{nl} = 300 MHz, and Γ_{ml} = 50 MHz.

trum is then ($|\Omega_\mu|$, $|\Omega| \gg |G|$) determined by the equation obtained in [12] for a strong running wave (that is, for $G_- = 0$) in the Doppler limit,

$$\Delta P_\mu = 2\hbar\omega_\mu |G_\mu|^2 \frac{\sqrt{\pi} N_l \exp(-\Omega_\mu^2/k_\mu^2 v_T^2)}{k_\mu v_T} \times \left(1 - \mathrm{Re}\frac{\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k)}{\sqrt{(\Gamma_p - i(\Omega_\mu - k_\mu\Omega/k))^2 + \frac{4k_\mu(k-k_\mu)|G|^2}{k^2}}}\right). \quad (5)$$

According to this equation, the transparency peak width should increase as field amplitude $|G|$ becomes larger.

5. CALCULATIONS FOR ARBITRARY INTENSITY

An equation for the work of the test field with a strong standing wave on the adjacent transition was obtained by Feldman and Feld [12]. We are interested in the situation when only finite level l is populated in the Λ scheme. In this situation, the work can be written as

$$P_\mu \propto N_l |G_\mu|^2 \mathrm{Im}\langle [L_0 - |G|^2(u_+ + u_-)]^{-1}\rangle, \quad (6)$$

$$L_{\pm 1}u_\pm = \cfrac{1}{1 - \cfrac{|G|^2}{L_{\pm1}L_{\pm2}\left(1 - \cfrac{|G|^2}{L_{\pm2}L_{\pm3}\left(1 - \cfrac{|G|^2}{L_{\pm3}L_{\pm4}(1-\dots)}\right)}\right)}}, \quad (7)$$

$$L_{\pm n} = \begin{cases} \Omega_\mu - \Omega - (k_\mu \pm nk)v + i\Gamma_{nl}, & n = 1, 3, 5, \dots, \\ \Omega_\mu - (k_\mu \pm nk)v + i\Gamma_{ml}, & n = 0, 2, 4, \dots. \end{cases} \quad (8)$$

Here, angle brackets denote averaging over velocities with a Maxwell distribution. At $v \gg \Omega/k$, Γ/k, the fraction rapidly converges, and it suffices to retain $N \sim G/kv$ terms.

To prove convergence of the fraction

$$\frac{P_n}{Q_n} = \cfrac{1}{1 + \cfrac{p_1}{1 + \cfrac{p_2}{1 + \cfrac{p_3}{1 + \dots}}}}, \quad (9)$$

as $n \longrightarrow \infty$, note that the numerators and denominators of the appropriate fractions satisfy the recurrent relations [24]

$$P_n = P_{n-1} + p_n P_{n-2}, \quad Q_n = Q_{n-1} + p_n Q_{n-2} \quad (10)$$

with the initial conditions

$$P_0 = P_1 = 1, \quad Q_0 = 1, \quad Q_1 = 1 + p_1.$$

We therefore have

$$P_n Q_{n-1} - P_{n-1}Q_n = -p_n(P_{n-1}Q_{n-2} - P_{n-2}Q_{n-1}),$$

and the difference of neighboring appropriate fractions equals

$$\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}} = \frac{(-1)^n}{Q_n Q_{n-1}} \prod_{i=1}^{n} p_i. \qquad (11)$$

A comparison of (9) and (7) for $n \longrightarrow \infty$ shows that

$$p_n \sim (G/kvn)^2.$$

In (10), the main term is the first one, and Q_n becomes independent of n. Difference (11) therefore decreases as $(n!)^{-2}$. To prove convergence, it remains to rewrite fraction (9) in the form of the sum of the differences

$$\frac{P_n}{Q_n} = \left(\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}}\right) + \ldots + \left(\frac{P_1}{Q_1} - \frac{P_0}{Q_0}\right) + \frac{P_0}{Q_0}.$$

The series converges uniformly as the series of the $J_0(2G/kv)$ Bessel function, similarly to the analytic solution obtained for equal relaxation constants and a zero detuning [1]. The real and imaginary parts of transition nm polarization as functions of velocity become strongly oscillating functions as $v \longrightarrow 0$.

The continued fraction was calculated numerically, and the integration over velocities was performed by the Simpson's rule in the interval $|v| \le 5v_T$. For convenience of comparison with experiment, field work at $G = 0$ was subtracted from its value given by (6). The calculation results for the $\Omega = 0$ resonance are given in Fig. 7. A comparison with perturbation theory formulas (5) shows that the depth of the dip between the split central peak components increases as the saturation parameter grows at a rate which is, because of saturation, lower than that observed in a comparatively weak field. Under nonresonance conditions (Fig. 8), the test field spectrum shifts and becomes asymmetric. The dimensionless saturation parameter for the spectrum shown in Fig. 8 is

$$\kappa = 8|G|^2/\Gamma_m\Gamma_{mn} \approx 13.$$

The transparency peak shifts as the detuning between the strong field and resonance increases, and an alternating contour is formed from the dip in the center of the line. This contour transforms into a peak as the detuning increases further. The amplitude of the peak decreases as the strong-field detuning becomes larger, but not so fast as in perturbation theory calculations (Fig. 6).

Note also that similar calculations performed by Feldman and Feld [12] for a three-level system with $k = 2k_\mu$ and equal relaxation constants of the higher and lower levels, $\Gamma_m = \Gamma_n$ and $\Gamma_l \ll \Gamma_{n,m}$, did not predict transparency peak splitting in the line center, that is, at $\Omega = 0$. At a large strong standing wave frequency detuning from the resonance, the transparency peak was shifted virtually without distortions, and no singularity was formed in the peak center. In [12], the opposite sign was selected, and positive values on the axis of ordinates corresponded to an increase in absorption in the strong field.

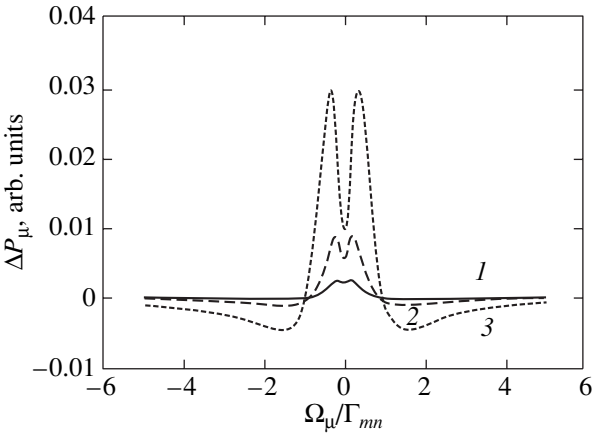

Fig. 7. Test field spectrum under resonance conditions, $\Omega = 0$, $kv_T = 20\Gamma_{mn}$, $\Gamma_{ml}/\Gamma_{nl} = 1/6$, $\Gamma_{ml} \approx \Gamma_{nl}$: strong field, $G/\Gamma_{mn} = (1)$ 0.1, (2) 0.2, and (3) 0.4.

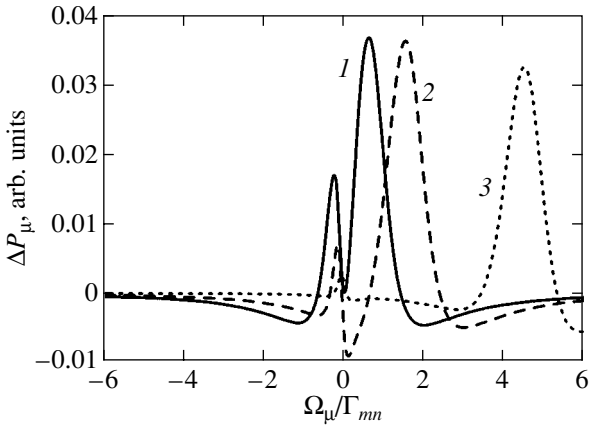

Fig. 8. Test field spectrum under nonresonance conditions, $G/\Gamma_{mn} = 0.4$, $kv_T = 20\Gamma_{mn}$, $\Gamma_{ml}/\Gamma_{nl} = 1/6$, $\Gamma_{ml} \approx \Gamma_{nl}$: strong-field detuning $\Omega/\Gamma_{mn} = (1)$ 0.7, (2) 2.2, and (3) 6.6.

We also performed calculations for the parameters used in [12]. Indeed, at $\Omega = 0$, no dip (transparency peak splitting) was formed, because, at the selected ratio between level widths, the relaxation constants corresponding to the test Γ_{ml} and forbidden Γ_{nl} transitions were equal, and all resonances in (3) and (4) had equal widths and did not separate at $\Omega = 0$. Frequency separation of resonances, however, became possible when strong-field detuning satisfied the inequality $|\Omega| > \Gamma_{\mu,\nu}, G$. This was substantiated by numerical calculations, according to which an additional resonance in the line center $\Omega_\mu = 0$ was formed against the background of the usual field-splitting resonance with a transparency peak at $\Omega_\mu = k_\mu\Omega/k$; the width of the additional resonance equaled that of field splitting. The peak in the line center decreased in amplitude as strong-field detuning

increased but remained quite resolvable up to detuning of $\Omega = 15\Gamma_{mn}$, for which the calculations in [12] were performed.

6. DISCUSSION

A comparison of exact calculation results with those obtained using perturbation theory and their comparison with both theory and experiment shows that, qualitatively, the observed resonance behaviors depending on the (G, Ω) strong-field parameters are nearly identical. Namely, under exact resonance conditions, the transparency peak width increases as field amplitude G grows, and a new resonance (a dip) is formed at $G/\Gamma_{mn} \sim 0.1$ characteristic values. This results in transparency peak splitting, whose depth increases with G. The transparency peak in the test field spectrum synchronously shifts as strong-field detuning Ω from resonance increases, $\Omega_\mu = \pm k_\mu \Omega/k$, and the observed resonance remains in the line center and changes sign while smoothly decreasing in amplitude.

A quantitative comparison of the shapes of the theoretical and experimental spectra reveals differences in resonance widths, however, not fundamental in character. Consider this point in more detail. In first-order perturbation theory, the width of the transparency peak at half height is

$$\Delta_{EIT} = \Gamma_p = \Gamma_{nl} k_\mu/k + (1 - k_\mu/k)\Gamma_{ml} \approx 230 \text{ MHz}.$$

This closely agrees with the experimental data obtained at low fields with $\Omega = 0$ (Fig. 3a). Although perturbation theory describes transparency peak splitting caused by an increase in the field amplitude, the experimental dip width equal to 150–200 MHz is much larger than that predicted theoretically ($2\Gamma_{ml} \lesssim 100$ MHz) even in comparatively low fields $G \sim 50$ MHz $\sim 0.1\Gamma_{mn}$. An increase in the field amplitude to $G \approx 100$ MHz (note that the standing wave amplitude then amounts to $2G \approx 200$ MHz) causes dip and main peak broadening by a factor of 1.5–2, the ratio between their widths remaining unchanged.

An increase in detuning to $|\Omega| > \Gamma_{nl}$, $|G|$ causes complete separation between the resonances from two running components of the standing wave and the resonance of higher harmonics; the centers of these resonances correspond to frequencies of $\Omega_\mu = k_\mu \Omega/k$ and $\Omega_\mu = 0$. The shape of the transparency peak centered at the $\Omega_\mu = k_\mu \Omega/k$ frequency can then be described by exact formula (5) obtained for a strong running wave. Calculations show that, at characteristic fields, the peak broadens insignificantly. The $\Delta_{EIT} \approx 250$ MHz peak width at $G \approx 100$ MHz closely agrees with the results of both numerical and experimental calculations (Figs. 4c, 4d). The theoretical width of the central peak is, however, 1.5–2 times smaller compared with the experimental value. The radiation width on the test transition is

$$\Gamma^0_{ml} = \frac{1}{2}(\Gamma_m + \Gamma_l) \approx 20 \text{ MHz},$$

and, if Stark broadening is taken into account,

$$\Gamma_{ml} = \Gamma^0_{ml} + \Delta\Gamma_{st} \approx 50 \text{ MHz}$$

(see [16–18]). In addition, we must take into account phase fluctuations involved in Coulomb scattering of ions (see [16]), which cannot be considered a mere addition to the coherence relaxation constant. This circumstance may be the reason for quantitative discrepancies between resonance width values. The experimental amplitudes of the transparency peak at $\Omega_\mu = k_\mu \Omega/k$ and the higher harmonic peak at $\Omega_\mu = 0$ closely agree with numerical calculations; the ratio between these amplitudes approximately equals 7 at $\Omega \approx 2$GHz and $G \approx 100$ MHz. At such fields, perturbation theory is already inapplicable and gives inaccurate amplitude values.

7. CONCLUSION

To summarize, we observed a new resonance in the test field spectrum of a three-level system in a strong standing light wave field. This resonance is observed in the line center independent of strong-field detuning even in comparatively weak fields $G \gtrsim \Gamma_{ml} \sim 0.1\Gamma_{mn}$. Note that, at $\Gamma_{ml} \ll \Gamma_{mn}$, the effect may be discernible even earlier than light-induced transparency peak broadening, which becomes noticeable at $G \gtrsim \Gamma_{mn}$. Under exact strong-field resonance conditions, the effect manifests itself as light-induced transparency peak splitting. Under nonresonance conditions, it takes the form of an additional isolated transparency peak in the line center. Perturbation theory including terms of second order in intensity was constructed to qualitatively describe this effect. This theory allows the new resonance to be interpreted as a manifestation of the higher coherence spatial harmonics on the transition with which the test field is in resonance. As spatial coherence modulation arises under simultaneous actions of counterrunning standing wave components, the effect is maximum for particles with zero velocities independent of the detuning of the standing wave frequency. At a qualitative level, this means that particles with zero velocities are most sensitive to spatial harmonics. Atoms that occur in standing wave nodes are insensitive to the strong field. Accordingly, strong-field-induced transparency (or absorption) disappears at zero detunings. Numerical calculations with the use of a formula in which polarization is described by a continued fraction allowed us to more accurately describe the experimental results. Some discrepancy (by a factor of 2 or less) between the theoretical and experimental central peak widths in nonresonance strong fields may be caused by Coulomb collisions of

ions, which result in Doppler coherence dephasing; this point requires a more detailed study. In this respect, measurements in a similar Λ scheme for uncharged particles would be of use. Note that the effect of the higher spatial coherence harmonics can be observed in both Stokes and anti-Stokes spectra in the interval $k \le k_\mu \le 2k$. The anti-Stokes scheme is of special interest, because, at $k_\mu > k$, the effect of the higher spatial harmonics may become predominant.

ACKNOWLEDGMENTS

The authors thank S.G. Rautian and A.M. Shalagin for useful discussions and S.I. Kablukov for help with assembling the unit for measurements. This work was financially supported by the Russian Foundation for Basic Research (projects nos. 00-02-17973 and 00-15-96808) and the Ministry of Science and Technology, the "Physics of Quantum and Wave Processes" program.

REFERENCES

1. S. G. Rautian and I. I. Sobel'man, Zh. Éksp. Teor. Fiz. **44**, 934 (1963) [Sov. Phys. JETP **17**, 635 (1963)].
2. S. Stenholm and W. E. Lamb, Phys. Rev. **181**, 618 (1969).
3. B. J. Feldman and M. S. Feld, Phys. Rev. A **1**, 1375 (1970).
4. S. Haroche and F. Hartmann, Phys. Rev. A **6**, 1280 (1972).
5. S. M. Freund, M. Römheld, and T. Oka, Phys. Rev. Lett. **35**, 1497 (1975).
6. J. Reid and T. Oka, Phys. Rev. Lett. **38**, 67 (1977).
7. E. Kyrölä and S. Stenholm, Opt. Commun. **22**, 123 (1977).
8. R. Corbalan, G. Orriols, L. Roso, *et al.*, Opt. Commun. **38**, 113 (1981).
9. R. Corbalan, G. Orriols, L. Roso, *et al.*, Opt. Commun. **40**, 29 (1981).
10. A. M. Bonch-Bruevich, T. A. Vartanyan, and N. A. Chigir', Zh. Éksp. Teor. Fiz. **77**, 1899 (1979) [Sov. Phys. JETP **50**, 901 (1979)].
11. An. M. Mak, S. G. Przhibel'skiĭ, and N. A. Chigir', Izv. Akad. Nauk SSSR, Ser. Fiz. **47**, 1976 (1983).
12. B. J. Feldman and M. S. Feld, Phys. Rev. A **5**, 899 (1972).
13. F. Silva, R. Corbalan, and R. Vilaseca, Opt. Commun. **114**, 519 (1995).
14. R. Corbalan, A. N. Pisarchik, V. N. Chizhevsky, and R. Vilaseca, Opt. Commun. **133**, 225 (1997).
15. F. Silva, J. Mompart, V. Ahufinger, and R. Corbalan, Phys. Rev. A **64**, 033802 (2001).
16. S. A. Babin and D. A. Shapiro, Phys. Rep. **241**, 119 (1994).
17. C. C. Davis and T. A. King, in *Advances in Quantum Electronics* (Academic, New York, 1975), Vol. 3, p. 169.
18. M. H. Dunn and J. N. Ross, Prog. Quantum Electron. **4**, 233 (1976).
19. B. F. G. Luyken, Physica (Amsterdam) **60**, 432 (1972).
20. S. A. Babin, S. I. Kablukov, and S. M. Kobtsev, Opt. Spektrosk. **84**, 915 (1998) [Opt. Spectrosc. **84**, 828 (1998)].
21. B. V. Bondarev, S. M. Kobtsev, A. V. Karablev, and V. M. Lunin, Opt. Atmos. **2**, 1319 (1989).
22. G. Stephan and M. Trümper, Phys. Rev. A **30**, 1925 (1984).
23. A. K. Popov, *An Introduction to the Nonlinear Spectroscopy* (Nauka, Novosibirsk, 1983).
24. A. Ya. Khinchin, *Chain Fractions* (Nauka, Moscow, 1978).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 704–719.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 819–837.
Original Russian Text Copyright © 2002 by Korol, Lyalin, Obolensky, Solovyov, Solovjev.

NUCLEI, PARTICLES, AND THEIR INTERACTION

A Relativistic Description of the Polarization Mechanism of Elastic Bremsstrahlung

A. V. Korol[a, *], A. G. Lyalin[b], O. I. Obolensky[c], A. V. Solovyov[c], and I. A. Solovjev[a]

[a]*St. Petersburg State Marine Technical University, ul. Lotsmanskaya 3, St. Petersburg, 198262 Russia*
[b]*Institute of Physics, St. Petersburg State University, St. Petersburg, 198504 Russia*
[c]*Ioffe Physicotechnical Institute, Russian Academy of Sciences, ul. Politekhnicheskaya 26, St. Petersburg, 194021 Russia*
**e-mail: korol@rpro.ioffe.rssi.ru*

Received October 16, 2001

Abstract—A completely relativistic mechanism for describing polarization bremsstrahlung caused by an elastic collision of a charged particle with a many-electron target was suggested. Multipole expansions for the amplitude and cross section of the process taking into account radiation lag effects were obtained. Including higher order multipoles was shown to result in substantial asymmetry of the angular distribution of emitted photons compared with the dipole case and in a noticeable change in the spectral characteristics of polarization radiation. The cross section of polarization bremsstrahlung was found to increase logarithmically as the energy of incident particles grew. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

When charged particles collide with a target with an internal electronic structure, bremsstrahlung arises as a result of the action of the so-called usual and polarization mechanisms. In the first mechanism, radiation is caused by the deceleration of incident particles in the static field of the target (e.g., see [1, 2]). The second mechanism is caused by the dynamic polarizability of the target under the action of the electric field of incident particles. Photon emission occurs as a result of virtual excitation (polarization) of target electrons by the electric field of an incident particle (e.g., see [3]). The existence of this radiation mechanism was discovered comparatively recently [4–7]. The purpose of this work was to study the frequency and angular dependences of the spectrum of polarization bremsstrahlung which arises in relativistic collisions between a charged particle and a many-electron target.

Polarization bremsstrahlung is a more complex process than the thoroughly studied usual bremsstrahlung, because it should be described taking into account the dynamic response of an atom target to the action of two fields created by an incident particle and an emitted photon. As a consequence, the problem acquires an essentially many-particle character (except when collisions with hydrogen atoms or hydrogen-like ions occur and simpler methods can be used [7–11]).

Polarization bremsstrahlung can be accompanied by target excitation or ionization in the final state. In what follows, we use the terms elastic and inelastic with reference to polarization bremsstrahlung without and with target excitation in the final state, respectively. In a wide range of photon frequencies, the elastic polarization bremsstrahlung channel predominates over the inelastic channel in scattering of both heavy [12–18] and light [19] particles on a many-electron atom, because [13, 20], in elastic polarization bremsstrahlung, atomic electrons radiate coherently (as is characteristic of Rayleigh light scattering). Conversely, in inelastic polarization bremsstrahlung, the contributions of separate target electrons to the total emission spectrum are incoherent (as in Raman scattering) and process cross sections rather than amplitudes should therefore be summed. As a result, the cross section of inelastic polarization bremsstrahlung is parametrically small for many-electron ($Z \gg 1$) targets. The contributions of the two process channels are approximately equal for targets with small Z values [13]. Note that this conclusion remains valid when the scattered (incident) particle has an internal structure (atom–atom, ion–atom, and ion–ion collisions). In this work, we restrict our consideration to elastic polarization bremsstrahlung.

A comprehensive review of the theoretical approaches to and the available experimental data on polarization bremsstrahlung can be found in [3, 21–24] and the references cited therein. We will only briefly mention the methods that were used to describe polarization bremsstrahlung arising in collisions between charged particles and many-electron atoms. These methods can conventionally be separated into two groups, namely, methods for analyzing the scattering process and approximations used to describe dynamic atomic response.

Two main approximations are extensively used to describe the scattering of both light (positron and electron) and heavy (proton, ion, or atom) incident particles

on many-electron atoms. These are the Born approximation (for both relativistic and nonrelativistic collisions) and the approximation based on the use of the nonrelativistic formalism of the method of distorted partial waves (DPWA) [25–30]. The Born approximation gives good results for scattering of heavy particles. The use of the DPWA method, which generalizes the Born approximation, allows us to determine the boundaries of the applicability of the latter and calculate polarization bremsstrahlung cross sections for comparatively low collision energies to which the Born approximation is inapplicable.

The dynamic response of a target to combined action of incident particle and emitted electron fields is described in terms of the nonrelativistic dipole approximation applicable to targets with small atomic numbers Z and to emission of photons with wavelengths ($\sim k^{-1}$) much larger than the size of the target in all works except the three mentioned below. At higher Z values (when atomic electron motions become essentially relativistic in character) and $k^{-1} \ll R_{at}$ (R_{at} is the target radius), this approximation cannot be considered satisfactory.

In such a description, the amplitude of polarization bremsstrahlung is expressed through the dipole polarizabilities of the target, which depend on energy ω and absolute momentum $\mathbf{q}$ of the virtual photon but are independent of momentum $\mathbf{k}$ of the emitted photon. If the incident particle is nonrelativistic (lag effects are ignored), the amplitude of polarization bremsstrahlung is proportional to scalar dipole polarizability $\alpha_d(\omega, q)$, which takes into account virtual excitation of the atom target caused by interactions with the Coulomb field of the incident particle [3, 21–24]. If a relativistic particle is scattered (lag effects are taken into account), the polarization bremsstrahlung amplitude contains the second polarizability, $\beta_d(\omega, q)$, in addition to $\alpha_d(\omega, q)$. This polarizability describes the dynamic response of an atom to the vector part of the virtual photon field [19]. The two polarizabilities coincide in the $q \longrightarrow 0$ limit. So far, numerical calculations have only been performed for the $\alpha_d(\omega, q)$ polarizability.

Outside the framework of the dipole approximation, polarization bremsstrahlung was considered for collisions of a heavy particle with nonrelativistic many-electron atoms [31–33] in terms of various approximations based on the nonrelativistic formalism of the many-body theory [3, 21–24] and for collisions with hydrogen atoms, when analytic methods can be applied directly [7–11]. In addition, atomic response in polarization bremsstrahlung processes was described using the nonrelativistic unscreened atom approximation [34]. Recently [35, 36], polarization bremsstrahlung arising in collisions between fast electrons (positrons) and many-electron atoms was considered in the local electron density and static Thomas–Fermi model approximation.

The theory of polarization bremsstrahlung resulting from collisions between relativistic particles has been developed to a much lesser degree. General equations for the amplitude and cross section of polarization bremsstrahlung were obtained in [19] within the framework of the relativistic Born approximation for emission of soft photons (for which the condition of the dipole character of radiation, $kR_{at} \ll 1$, is satisfied). It was also predicted in [19] for relativistic collisions that the contribution of the polarization bremsstrahlung mechanism to the total bremsstrahlung spectrum should exceed the corresponding nonrelativistic contribution and should logarithmically increase with increasing energy of incident particles because of the lag of interaction between the incident particle and the atom.[1]

Recently [37], a completely relativistic formalism was suggested to describe the arising of bremsstrahlung in elastic collisions between charged particles and many-electron targets whose states could be described in terms of the relativistic one-particle self-consistent field approximation. Both principal bremsstrahlung mechanisms, usual and polarization, and their interference were considered. The motion of the incident particle was described by the relativistic formalism of the DPWA method. This approach generalizes simpler approximations applied earlier and allowed the spectral and spectral–angular characteristics of bremsstrahlung to be calculated for various collision conditions (from nonrelativistic to ultrarelativistic) with the participation of nonrelativistic (light) and relativistic (heavy) atoms and ions.

In this work, we restrict ourselves to consideration of collisions of heavy incident particles with many-electron targets within the framework of the approach developed in [37]. It will be assumed that polarization bremsstrahlung arises in a collision between an incident particle whose motion can be described in the relativistic Born approximation and a spherically symmetrical target (atom or ion) whose initial and final states coincide. According to [37], the characteristics of polarization bremsstrahlung are then expressed through generalized multipole target polarizabilities of three types corresponding to the allowed combinations of the types (longitudinal, electric, and magnetic) of the virtual and emitted photons in the amplitude of polarization bremsstrahlung.

[1] In considering the polarization bremsstrahlung mechanism, we must distinguish between the effects that are due to interaction lag and the emission of high-multipolarity photons. Interaction lag implies that the incident relativistic particle interacts with the target not only via the Coulomb field but also (in the ultrarelativistic case, mainly) via the field of transverse virtual photons (see [1]). In the latter case, the effective interaction radius increases as the velocity of the incident particle grows. The multipole character of radiation is solely determined by the kR_{at} parameter value. If $kR_{at} \ll 1$, radiation is of an essentially dipolar character no matter what the velocity of the incident particle.

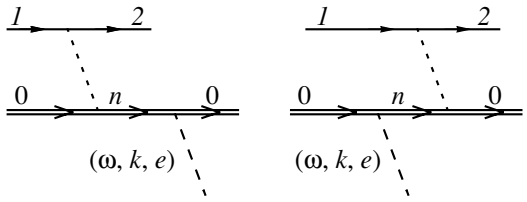

Fig. 1. Feynman diagrams for elastic polarization bremsstrahlung of a structureless charged particle which experiences scattering on a many-electron atom. Solid lines are the wave functions of the incident particle satisfying the Dirac equation (Farry representation). The initial (index "*1*") and final (index "*2*") incident particle states are characterized by momenta $\mathbf{p}_{1,2}$ and polarizations $\mu_{1,2}$. Double lines depict target states, indices "0" and "*n*" correspond to the initial (final) and intermediate (virtual) target states. Dashed lines correspond to the wave function of the emitted photon with energy ω, momentum **k**, and polarization vector **e**, and dotted lines denote Green's function of the virtual photon with energy ω and momentum **q**.

The formalism described in this work allows multipole expansions to be obtained for the amplitude and cross section of polarization bremsstrahlung; these expansions take into account radiation lag effects and emission of high-multipolarity photons and can be used in numerical calculations of the spectral and angular characteristics of polarization bremsstrahlung. Most attention is given to polarization bremsstrahlung of target inner shell electrons. Such a study is expedient for two reasons. First, inner shell electrons make the major contribution to the formation of the complete polarization bremsstrahlung spectrum in a very wide frequency range [3]. Secondly, the special features of the complete polarization bremsstrahlung spectrum that are due to inner-shell electrons can be described using simple approximations (such as the hydrogen-like and the Hartree–Fock–Dirac approximations) without taking into account many-electron correlations. The use of such approximations allows the observed effects to be given simple physical interpretation and substantially decreases laboriousness of calculations.

So far, no detailed numerical analysis of the spectral and angular distributions has been performed for polarization bremsstrahlung of inner shell electrons (except several first numerical results published in [37]). We therefore believe the problem tackled in this work to be fairly topical. The calculation results described below show that taking into account higher order multipoles results in a substantial asymmetry of the angular distribution of emitted photons compared with the dipole approximation and in noticeable changes in the spectral characteristics of polarization radiation. In addition, the important special feature of the cross section of relativistic polarization bremsstrahlung, namely, its logarithmic increase with increasing the energy of the incident particle, is clearly demonstrated.

Note that such calculations are of special importance for exact comparison with the recent experimental data on bremsstrahlung caused by collisions between electrons with 10–100 keV energies and various solid-state, thin-film, and gaseous targets such as Al, Cu, Ni, Ag, and Xe [38–40]. The experimental results reported in these works are indicative of essential polarization bremsstrahlung mechanism contributions to the complete radiation spectrum in the energy range in which a relativistic description of polarization bremsstrahlung is necessary. In addition, of considerable interest are calculations related to relativistic collisions involving heavy ions in view of the attempts made recently in this direction [41]. In [41], the spectrum of bremsstrahlung formed in collisions between helium-like uranium ions with a 223 MeV/nucleon energy and N_2 and Ar gas targets was measured.

2. POLARIZATION BREMSSTRAHLUNG AMPLITUDE AND CROSS SECTION

Consider polarization bremsstrahlung (Fig. 1) caused by an elastic collision between an incident particle with charge $Z_p e$ and mass m_p and a spherically symmetrical target (atom or ion) accompanied by the transition of the particle from initial state $(\mathbf{p}_1, \mu_1)$ (with energy $\varepsilon_1 = \sqrt{\mathbf{p}_1^2 + m_p^2}$) to final state $(\mathbf{p}_2, \mu_2)$ (with energy $\varepsilon_2 = \sqrt{\mathbf{p}_2^2 + m_p^2}$). As the initial and final states of the target coincide, the law of the conservation of energy has the form[2]

$$\varepsilon_1 = \varepsilon_2 + \omega. \tag{1}$$

The spectral–angular radiation distribution $d^2\sigma/d\omega d\Omega_{\mathbf{k}}$, which is obtained from the differential cross section by the integration along the scattered particle momentum $\mathbf{p}_2$ direction, the summation over the polarization of the emitted photon (λ_p) and the scattered particle (μ_2), and averaging over incident particle polarizations μ_1, is written as

$$\frac{d^2\sigma}{d\omega d\Omega_{\mathbf{k}}} = \frac{p_2}{8p_1}\frac{\omega}{(2\pi)^4}\sum_{\lambda_p}\sum_{\mu_1,\mu_2}\int_{(4\pi)} d\Omega_{\mathbf{p}_2}|\mathcal{M}|^2, \tag{2}$$

where $d\Omega_{\mathbf{k}}$ and $\Omega_{\mathbf{p}_2}$ are the solid angle elements of the emitted photon and scattered particle, respectively. The amplitude of the process ($\mathcal{M}$) is described by two Feynman diagrams (Fig. 1), which correspond to composite matrix elements of target transitions from initial state "0" to excited state "*n*" and back to final state "0" under the action of emitted photon $\mathbf{A}_{(\gamma)}$ field and the field of

[2] We use the relativistic system of units ($\hbar = c = 1$). For four-dimensional values, the (+−−−) metric signature is used.

the virtual photon characterized by the 4-potential $A^\nu = (\Phi, \mathbf{A})$. The amplitude has the form

$$\mathcal{M} = e^2 \sum_n \left\{ \frac{\langle 0|\mathbf{A}_{(\gamma)}\boldsymbol{\gamma}|n\rangle\langle n|\gamma^\nu A_\nu|0\rangle}{\varepsilon_n(1-i0)-\varepsilon_0-\omega} + \frac{\langle 0|\gamma^\nu A_\nu|n\rangle\langle n|\mathbf{A}_{(\gamma)}\boldsymbol{\gamma}|0\rangle}{\varepsilon_n(1-i0)-\varepsilon_0+\omega} \right\}. \quad (3)$$

Here, $\gamma^\nu = (\gamma^0, \boldsymbol{\gamma})$ are the Dirac matrices. The summation is over the quantum numbers of the complete spectrum of target excited states and includes the contributions of intermediate states with both positive ($\varepsilon_n > 0$) and negative ($\varepsilon_n < 0$) energies. The potential of the emitted photon field has the form[3]

$$\mathbf{A}_{(\gamma)} = \sum_{a=1}^{N} \mathbf{e} \exp(-i\mathbf{k}\cdot\mathbf{r}_a), \quad (4)$$

where $\mathbf{e}$ is the polarization vector (further, we use $\mathbf{e}^* = \mathbf{e}$ on the assumption of a linear photon polarization). The summation is over target electrons, $\mathbf{r}_a$ is the coordinate of the ath electron, and N is the total number of electrons.

In the relativistic Born approximation, the initial and final states of an incident particle are described by the wave functions

$$\Psi^{(\pm)}_{\mathbf{p}\mu}(\mathbf{r}) = u_\mu(\varepsilon, \mathbf{p}) \exp(i\mathbf{p}\cdot\mathbf{r}) \quad (5)$$

with bispinor amplitudes $u_\mu(\varepsilon, \mathbf{p})$ defined as [2]

$$u_\mu(\varepsilon, \mathbf{p}) = \begin{pmatrix} \sqrt{\varepsilon + m}\,\chi_\mu(\mathbf{n_p}) \\ \sqrt{\varepsilon - m}\,(\boldsymbol{\sigma}\cdot\mathbf{n_p})\chi_\mu(\mathbf{n_p}) \end{pmatrix}. \quad (6)$$

Here and throughout, $\mathbf{n_a}$ is the unit vector along direction $\mathbf{a}$, and $\boldsymbol{\sigma}$ are the Pauli matrices. Two-component spinors $\chi_\mu(\mathbf{n_p})$ satisfy the normalization condition $\chi^\dagger_\mu(\mathbf{n_p})\chi_\mu(\mathbf{n_p}) = 1$ (symbol † denotes Hermitian conjugation).

The four-dimensional potential of the field created by the incident particle is written as

$$A^\nu = Z_p e \sum_{a=1}^{N} \int d\mathbf{r} \Psi^{(-)\dagger}_{\mathbf{p}_2\mu_2}(\mathbf{r}) \gamma^\mu D_{\mu\nu}(\omega, \mathbf{r}-\mathbf{r}_a) \Psi^{(+)}_{\mathbf{p}_1\mu_1}(\mathbf{r}) = Z_p e \sum_{a=1}^{N} \int \frac{d\mathbf{q}}{(2\pi)^3} \exp(i\mathbf{q}\cdot\mathbf{r}_a) D_{\mu\nu}(\omega, \mathbf{q}) B^\mu_{21}(\mathbf{q}). \quad (7)$$

Here, ω and $\mathbf{q}$ are the energy and the momentum of the virtual photon, $D_{\mu\nu}(\omega, \mathbf{r}-\mathbf{r}_a)$ is Green's function of the photon,[4] and the $B^\mu_{21}(\mathbf{q})$ 4-vector is the Fourier transform of the incident particle current,

$$B^\mu_{21}(\mathbf{q}) = \int d\mathbf{r} \Psi^{(-)\dagger}_{\mathbf{p}_2\mu_2}(\mathbf{r}) \gamma^0\gamma^\mu \exp(-i\mathbf{q}\cdot\mathbf{r}) \Psi^{(+)}_{\mathbf{p}_1\mu_1}(\mathbf{r}) = (2\pi)^3 b^\mu \delta(\mathbf{p}_1 - \mathbf{p}_2 - \mathbf{q}), \quad (8)$$

where

$$b^\mu = u^\dagger_{\mu_2}(\varepsilon_2, \mathbf{p}_2)\gamma^0\gamma^\mu u_{\mu_1}(\varepsilon_1, \mathbf{p}_1). \quad (9)$$

The δ function present in (8) allows the integration over $d\mathbf{q}$ in the last integral in (7) to be easily performed, and the virtual photon momentum proves to be equal to the difference of the initial and final momenta of the incident particle,

$$\mathbf{q} = \mathbf{p}_1 - \mathbf{p}_2. \quad (10)$$

At a fixed photon energy, the minimum and maximum transferred momentum values are $q_{\min} = p_1 - p_2$ and $q_{\max} = p_1 + p_2$.

By analogy with the approach used in the Hartree–Fock–Dirac approximation (e.g., see [42]), we assume that target states $|0\rangle$, $|n\rangle$ can be described by one-electron wave functions corresponding to states with definite relativistic energy ε, total angular momentum j, orbital angular momentum $l = j \pm 1/2$, and total angular momentum projection m values. Substituting (4)–(7) into (3) can then be performed with replacing the summation over target electrons by the summation over quantum numbers ε_i, j_i, l_i, and m_i (i takes on the values 0 and n) of target subshells in the ground and excited states. One-electron wave functions $\Psi_{\varepsilon_i j_i l_i m_i}(\mathbf{r})$ then have the following bispinor structure:

$$\Psi_{\varepsilon_i j_i l_i m_i}(\mathbf{r}) = \frac{1}{r}\begin{pmatrix} g(r)\Omega_{j_i l_i m_i}(\mathbf{n}) \\ -i f(r)(\boldsymbol{\sigma}\cdot\mathbf{n})\Omega_{j_i l_i m_i}(\mathbf{n}) \end{pmatrix}. \quad (11)$$

Here, $g(r) \equiv g_{\varepsilon_i j_i l_i}(r)$ and $f(r) \equiv f_{\varepsilon_i j_i l_i}(r)$ are the large and small relativistic wave function components, respectively, which can be obtained by solving the system of Hartree–Fock–Dirac radial equations (e.g., see [43]), and $\Omega_{jlm}(\mathbf{n})$ are the spherical spinors defined according to [44].

It follows that, for a many-electron target, the polarization bremsstrahlung amplitude takes the form

$$\mathcal{M} = Z_p e b^\mu D_{\mu\nu}(\omega, \mathbf{q}) M^\nu(\omega, \mathbf{k}, \mathbf{e}; \mathbf{q}). \quad (12)$$

[3] The $\sqrt{2\pi/\omega}$ normalization factor is not included in the definition of $\mathbf{A}_{(\gamma)}$.

[4] We use the Coulomb calibration of the photon propagator.

The dynamic response of the target is described by the equation

$$M^{\nu}(\omega, \mathbf{k}, \mathbf{e}; \mathbf{q}) = e^2 \sum_{\{0,n\}} \left\{ \frac{\langle 0|\mathbf{e}\cdot\boldsymbol{\gamma} e^{-i\mathbf{k}\cdot\mathbf{r}}|n\rangle\langle n|\gamma^{\nu} e^{i\mathbf{q}\cdot\mathbf{r}}|0\rangle}{\omega_{n0}-\omega} + \frac{\langle 0|\gamma^{\nu} e^{i\mathbf{q}\cdot\mathbf{r}}|n\rangle\langle n|\mathbf{e}\cdot\boldsymbol{\gamma}^{-i\mathbf{k}\cdot\mathbf{r}}|0\rangle}{\omega_{n0}+\omega} \right\}. \quad (13)$$

Here, $\omega_{n0} = \varepsilon_n(1 - i0) - \varepsilon_0$ is the transition energy, and the $\sum_{\{0\}}$ summation is over the ε_0, j_0, l_0, and m_0 quantum numbers of target subshells in the initial (final) state. All target subshells are assumed to be occupied. We then have $m_0 = -l_0, \ldots, l_0$. The $\sum_{\{n\}}$ sum is taken over the ε_n, j_n, l_n, and m_n transition state quantum numbers and contains the contributions of states with both positive ($\varepsilon_n > 0$) and negative ($\varepsilon_n < 0$) energies.

After separating the contributions of the scalar and vector parts of the A^{ν} potential [see (7)], the amplitude is written as

$$\mathcal{M} = (4\pi)^2 Z_p e(\mathbf{e}\cdot\mathbf{F}) = (4\pi)^2 Z_p e\mathbf{e}\cdot(\mathbf{F}^{(s)} + \mathbf{F}^{(e)} + \mathbf{F}^{(m)}), \quad (14)$$

where the $\mathbf{F}^{(s,e,m)}$terms [the meaning of indices (s, e, m) is explained below] have the form

$$\mathbf{e}\cdot\mathbf{F}^{(s)} = -\frac{1}{4\pi}\frac{e^2}{q^2} b^0 \times \sum_{\{0,n\}} \left[\frac{\mathcal{F}_{0n}(\mathbf{e}, -\mathbf{k}) F_{n0}(\mathbf{q})}{\omega_{n0}-\omega} + \frac{F_{0n}(\mathbf{q})\mathcal{F}_{n0}(\mathbf{e}, -\mathbf{k})}{\omega_{n0}+\omega} \right], \quad (15)$$

$$\mathbf{e}\cdot(\mathbf{F}^{(e)} + \mathbf{F}^{(m)}) = -\frac{1}{4\pi}\frac{e^2}{\omega^2 - q^2} \times \sum_{\{0,n\}} \left[\frac{\mathcal{F}_{0n}(\mathbf{e}, -\mathbf{k})\mathcal{F}_{0n}(\mathbf{R}, \mathbf{q})}{\omega_{n0}-\omega} + \frac{\mathcal{F}_{0n}(\mathbf{R}, \mathbf{q})\mathcal{F}_{n0}(\mathbf{e}, -\mathbf{k})}{\omega_{n0}+\omega} \right]. \quad (16)$$

The **R** vector, which is orthogonal to **q**, is defined as

$$\mathbf{R} = \mathbf{b} - \frac{(\mathbf{b}\cdot\mathbf{q})\mathbf{q}}{q^2}. \quad (17)$$

The $\mathcal{F}_{n0}(\mathbf{e}, -\mathbf{k})$ matrix element describes target transition from ground state 0 to excited (virtual) state n under the action of the $\mathbf{e}\cdot\boldsymbol{\gamma}\exp(-i\mathbf{k}\cdot\mathbf{r})$ operator,

$$\mathcal{F}_{0n}(\mathbf{e}, -\mathbf{k}) = \int d\mathbf{r}\Psi^{\dagger}_{\varepsilon_n j_n l_n m_n}(\mathbf{r}) \times \gamma^0 \mathbf{e}\cdot\boldsymbol{\gamma}\exp(-i\mathbf{k}\cdot\mathbf{r})\Psi_{\varepsilon_0 j_0 l_0 m_0}(\mathbf{r}). \quad (18)$$

The $F_{n0}(\mathbf{q})$ and $\mathcal{F}_{n0}(\mathbf{R}, \mathbf{q})$ matrix elements describe the $0 \longrightarrow n$ transition caused by interaction between a target electron and the scalar [proportional to $\gamma^0\exp(i\mathbf{q}\cdot\mathbf{r})$] and vector [proportional to $\boldsymbol{\gamma}\exp(i\mathbf{q}\cdot\mathbf{r})$] parts of the virtual photon field created by the incident particle,

$$F_{n0}(\mathbf{q}) = \int d\mathbf{r}\Psi^{\dagger}_{\varepsilon_n j_n l_n m_n}(\mathbf{r})\exp(i\mathbf{q}\cdot\mathbf{r})\Psi_{\varepsilon_0 j_0 l_0 m_0}(\mathbf{r}), \quad (19)$$

$$\mathcal{F}_{n0}(\mathbf{R}, \mathbf{q}) = \int d\mathbf{r}\Psi^{\dagger}_{\varepsilon_n j_n l_n m_n}(\mathbf{r})\gamma^0\mathbf{R}\cdot\boldsymbol{\gamma} \times \exp(i\mathbf{q}\cdot\mathbf{r})\Psi_{\varepsilon_0 j_0 l_0 m_0}(\mathbf{r}). \quad (20)$$

The $\mathcal{F}_{0n}(\mathbf{e}, -\mathbf{k})$, $F_{0n}(\mathbf{q})$, and $\mathcal{F}_{0n}(\mathbf{R}, \mathbf{q})$ matrix elements corresponding to the $n \longrightarrow 0$ virtual transition can be obtained from (18)–(20) using the substitution $n \longleftrightarrow 0$.

Substituting the multipole expansions of the $\mathbf{e}\cdot\boldsymbol{\gamma}\exp(-i\mathbf{k}\cdot\mathbf{r})$ and $\exp(-i\mathbf{q}\cdot\mathbf{r})$ operators in spherical vector $\mathbf{Y}^{(0,1)}_{lm}(\mathbf{n})$ and spherical $Y_{lm}(\mathbf{n}) = \mathbf{n}\cdot\mathbf{Y}^{(-1)}_{lm}(\mathbf{n})$ harmonics defined as in [44] and performing the necessary transformations by the formulas from [44, §§5, 7], we can obtain the multipole expansions for the $\mathbf{F}^{(s,e,m)}$ vectors given below. The selection rules for multipole transitions of the magnetic [index (m), $\lambda = 0$], electric [index (e), $\lambda = 1$], and longitudinal [index (s), $\lambda = -1$] types are different,

$$l_n + l_0 + l = \begin{cases} \text{odd}, & \lambda = 0, \\ \text{even}, & \lambda = \pm 1. \end{cases} \quad (21)$$

The expression for $\mathbf{F}^{(s)}$ therefore only contains components proportional to the electric multipoles of the emitted photon field, whereas the $\mathbf{F}^{(e)}$ and $\mathbf{F}^{(m)}$ terms are diagonal with respect to the λ indices of the virtual and real photons.

It follows that, for a spherically symmetrical target, the amplitude of polarization bremsstrahlung is expressed through partial generalized polarizabilities of three different types corresponding to the following combinations of virtual and real photon types: longitudinal–electric, electric–electric, and magnetic–magnetic. Each of these polarizabilities depends on photon energy ω and orbital momentum l (these values are equal for both photons) and on the absolute values of momenta q and k of the virtual and real photons.

Such a form of representing the polarization bremsstrahlung amplitude has considerable advantages of both fundamental and computational character. The possibility of expressing $\mathcal{M}$ in terms of the polarizability types specified above, which contain all necessary information on the dynamic structure of the many-electron target, simplifies the problem and allows computation time for calculating the characteristics of polarization bremsstrahlung to be substantially reduced.

The analytic expressions for $\mathbf{F}^{(s,e,m)}$ have the form

$$\mathbf{F}^{(s)} = -\frac{\mathbf{b}\cdot\mathbf{q}}{q}\sum_{lm}\frac{\sqrt{l(l+1)}}{2l+1} \times Y^{*}_{lm}(\mathbf{n}_{\mathbf{q}})\mathbf{Y}^{(1)}_{lm}(\mathbf{n}_{\mathbf{k}})\alpha_l(\omega, q, k), \quad (22)$$

$$\mathbf{F}^{(e)} = \frac{1}{\omega^2 - q^2}\sum_{lm}\frac{l(l+1)}{2l+1} \times (\mathbf{b}\cdot\mathbf{Y}^{(1)*}_{lm}(\mathbf{n_q}))\mathbf{Y}^{(1)}_{lm}(\mathbf{n_k})\beta^{(1)}_l(\omega, q, k), \quad (23)$$

$$\mathbf{F}^{(m)} = \frac{1}{\omega^2 - q^2}\sum_{lm}\frac{l(l+1)}{2l+1} \times (\mathbf{b}\cdot\mathbf{Y}^{(0)}_{lm}(\mathbf{n_q}))\mathbf{Y}^{(0)}_{lm}(\mathbf{n_k})\beta^{(0)}_l(\omega, q, k). \quad (24)$$

In (22), it is taken into account that $b^0\omega = \mathbf{b}\cdot\mathbf{q}$; this relation follows from the law of conservation of the 4-vector of incident particle current. The summation over l is from $l = 1$ to infinity. The partial generalized polarizabilities in the right-hand sides of (22)–(24) are given by

$$\alpha_l(\omega, q, k) = \frac{2l+1}{q}e^2 \times \sum_{\substack{\varepsilon_0 j_0 l_0 \\ \varepsilon_n j_n l_n}} \frac{2C^{(-1)}_{0n}(l) f^{(1)}_{0n}(k; l) f^{(-1)}_{n0}(q; l)}{\omega^2_{n0} - \omega^2}, \quad (25)$$

$$\beta^{(\lambda)}_l(\omega, q, k) = (2l+1)e^2 \times \sum_{\substack{\varepsilon_0 j_0 l_0 \\ \varepsilon_n j_n l_n}} \frac{2\omega_{n0}C^{(\lambda)}_{0n}(l) f^{(\lambda)}_{0n}(k; l) f^{(\lambda)}_{n0}(q; l)}{\omega^2_{n0} - \omega^2}, \quad (26)$$

$$\lambda = 0, 1.$$

The $f^{(\lambda)}_{ba}(s; l)$ [$\lambda = -1, 0, 1$; $(b, a) = (0, n)$; $x = (k, q)$] radial integrals and the $C^{(\lambda)}_{0n}(l)$ coefficients have the form

$$f^{(-1)}_{ba}(x; l) = \int_0^\infty dr(g^*_b(r)g_a(r) + f^*_b(r)f_a(r))j_l(xr), \quad (27)$$

$$f^{(0)}_{ba}(x; l) = \int_0^\infty dr(g^*_b(r)f_a(r) + f^*_b(r)g_a(r))j_l(xr), \quad (28)$$

$$f^{(1)}_{ba}(x; l) = \int_0^\infty dr\left\{(g^*_b(r)f_a(r) - f^*_b(r)g_a(r))\frac{j_l(xr)}{xr} - \frac{\kappa_b - \kappa_a}{l(l+1)}(g^*_b(r)f_a(r) + f^*_b(r)g_a(r)) \times\left[\frac{1}{x}\frac{dj_l(xr)}{dr} + \frac{j_l(xr)}{xr}\right]\right\}, \quad (29)$$

$$C^{(\lambda)}_{0n}(l) = \xi(l_n l_0 l\lambda 1)\Pi^2_{j_0 j_n}\begin{pmatrix} j_0 & j_n & l \\ \frac{1}{2} & -\frac{1}{2} & 0\end{pmatrix}^2 \times\begin{cases}1, & \lambda = \pm 1,\\ -\dfrac{(\kappa_0 + \kappa_n)^2}{l^2(l+1)^2}, & \lambda = 0.\end{cases} \quad (30)$$

Here, $j_n(z)$ is a spherical Bessel function,

$$\kappa = l(l+1) - j(j+1) - \frac{1}{4}$$

is the relativistic quantum number, the $3j$ Wigner symbol whose square appears in the expression for $C^{(\lambda)}_{0n}(l)$ [the coefficient in front of braces in (30)] is defined according to [44], and

$$\Pi_{l_\alpha l_b \ldots} = \sqrt{(2l_a+1)(2l_b+1)\ldots},$$

$$\xi(l_a l_b l_c l_d l_e) = \frac{1 + (-1)^{l_a + l_b + l_c + l_d + l_e}}{2}.$$

The right-hand sides of (25) and (26) can be transformed to a form more convenient for computations by introducing the relativistic one-particle Green's function to describe excited electronic states of the target. The corresponding equations are given in the Appendix.

To determine the spectral–angular distribution of polarization bremsstrahlung, (14) should be substituted into (2). Summing over emitted photon polarizations (taking into account that

$$\mathbf{k}\cdot\mathbf{F}^{(s, e, m)} \propto \mathbf{k}\cdot\mathbf{Y}^{(0, 1)}_{lm}(\mathbf{n_k}) = 0),$$

and representing the $d\Omega_{\mathbf{p}_2}$ differential in the form $d\phi_{\mathbf{q}}qdq/(p_1p_2)$, we obtain

$$\frac{d^2\sigma}{d\omega d\Omega_{\mathbf{k}}} = Z^\alpha_p\alpha\frac{2\omega}{p^2_1}\sum_{\mu_1, \mu_2}\int_0^{2\pi} d\phi_{\mathbf{q}}\int_{q_{\min}}^{q_{\max}} qdq\mathbf{F}\cdot\mathbf{F}^\dagger. \quad (31)$$

Here, $\alpha = e^2/\hbar c \approx 1/137$ is the fine structure constant.

The further transformations of (31) include the summation (averaging) over incident particle polarizations, the multipole expansion of the $\mathbf{Y}^{(l)}_{lm}(\mathbf{n_k})\cdot\mathbf{Y}^{(\lambda')*}_{l'm'}(\mathbf{n_k})$ quadratic form, and the integration over the $d\phi_{\mathbf{q}}$ angle.

Omitting intermediate calculations, let us write the final result for cross section (31), which determines the spectral–angular polarization bremsstrahlung distribution,

$$\frac{d^2\sigma}{d\omega d\Omega_{\mathbf{k}}} = \sum_{l_k = 0}^\infty (2l_k+1)P_{l_k}(\cos\theta_{\mathbf{k}})\frac{d\sigma_{l_k}}{d\omega}, \quad (32)$$

$$\frac{d\sigma_{l_k}}{d\omega} = Z_p^2\alpha\frac{\omega}{8\pi p_1^2}\int\limits_{q_{\min}}^{q_{\max}} dq^2[A_{l_k} + B_{l_k} + C_{l_k} + D_{l_k}], \quad (33)$$

$$A_{l_k} = \left(\frac{q^2-\omega^2}{4} + (\mathbf{p}_1\cdot\mathbf{n}_\mathbf{q})(\mathbf{p}_2\cdot\mathbf{n}_\mathbf{q})\right)P_{l_k}(y)$$
$$\times\sum_{l,l'=1}^{\infty}(l(l+1)+l'(l'+1)-l_k(l_k+1)) \quad (34)$$
$$\times\begin{pmatrix} l' & l & l_k \\ 0 & 0 & 0\end{pmatrix}^2 \mathrm{Re}[\alpha_l(\omega,q,k)\alpha_{l'}^*(\omega,q,k)],$$

$$B_{l_k} = \frac{1}{(q^2-\omega^2)^2}\left(\frac{q^2-\omega^2}{2} + p_1^2\sin^2\theta_\mathbf{q}\right)P_{l_k}(y)$$
$$\times\sum_{l,l'=1}^{\infty}\sum_{\lambda,\lambda'=0}^{1}\xi(l_k ll'\lambda\lambda')ll'(l+1)(l'+1) \quad (35)$$
$$\times\begin{pmatrix} l' & l & l_k \\ 1 & -1 & 0\end{pmatrix}^2 \mathrm{Re}[\beta_l^{(\lambda)}(\omega,q,k)\beta_{l'}^{(\lambda')*}(\omega,q,k)],$$

$$C_{l_k} = \frac{p_1^2\sin^2\theta_\mathbf{q}}{(q^2-\omega^2)^2}\frac{2yP'_{l_k}(y) - l_k(l_k+1)P_{l_k}(y)}{\sqrt{(l_k-1)l_k(l_k+1)(l_k+2)}}$$
$$\times\sum_{l,l'=1}^{\infty}\sum_{\lambda,\lambda'=0}^{1}(-1)^{\lambda\lambda'}\xi(l_k ll'\lambda\lambda')ll'(l+1)(l'+1) \quad (36)$$
$$\times\begin{pmatrix} l' & l & l_k \\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} l & l_k & l' \\ 1 & -2 & 1\end{pmatrix}$$
$$\times\mathrm{Re}[\beta_l^{(\lambda)}(\omega,q,k)\beta_{l'}^{(\lambda')*}(\omega,q,k)],$$

$$D_{l_k} = \frac{4p_1(\mathbf{p}_2\cdot\mathbf{n}_\mathbf{q})\sin^2\theta_\mathbf{q}}{q^2-\omega^2}\frac{P'_{l_k}(y)}{\sqrt{l_k(l_k+1)}}$$
$$\times\sum_{l,l'=1}^{\infty}\sum_{\lambda=0}^{1}\xi(l_k ll'\lambda 1)l(l+1)\sqrt{l'(l'+1)} \quad (37)$$
$$\times\begin{pmatrix} l' & l & l_k \\ 1 & -1 & 0\end{pmatrix}\begin{pmatrix} l' & l_k & l \\ 0 & -1 & 1\end{pmatrix}$$
$$\times\mathrm{Re}[\beta_l^{(\lambda)}(\omega,q,k)\alpha_{l'}^*(\omega,q,k)].$$

In these equations, $P'_{l_k}(y)$ is the derivative of the Legendre polynomial with respect to its argument. Using the law of conservation of energy [Eq. (1)] and (10), we can express all kinematic coefficients in (34)–(37) through variables ω, q, $q_{\min}$, and $q_{\max}$ as

$$\sin^2\theta_\mathbf{q} = \frac{p_1^2q^2 - (\mathbf{p}_1\cdot\mathbf{q})^2}{p_1^2q^2} = \frac{(q^2-q_{\min}^2)(q_{\max}^2-q^2)}{4p_1^2q^2}, \quad (38)$$

$$2\mathbf{p}_1\cdot\mathbf{q} = q_{\min}q_{\max} + q^2,$$
$$2\mathbf{p}_2\cdot\mathbf{q} = q_{\min}q_{\max} - q^2, \quad (39)$$

$$y = \cos\theta_\mathbf{q} = \frac{\mathbf{p}_1\cdot\mathbf{q}}{p_1q} = \frac{q_{\min}q_{\max}+q^2}{2p_1q}, \quad (40)$$

$$\frac{q^2-\omega^2}{4} + (\mathbf{p}_1\cdot\mathbf{n}_\mathbf{q})(\mathbf{p}_2\cdot\mathbf{n}_\mathbf{q}) = \frac{q_{\min}^2q_{\max}^2 - q^2\omega^2}{4q^2}. \quad (41)$$

It is easy to see that the $C_{0,1}$ and D_0 terms in (36) and (37) identically equal zero because the corresponding combinations of the Legendre polynomials and their derivatives vanish. The summation over l_k in (32) therefore starts with $l_k = 0$ for A_{l_k} and B_{l_k}, $l_k = 2$ for C_{l_k}, and $l_k = 1$ for D_{l_k}.

The spectral distribution of radiation is obtained from (32) by the integration over solid angle $d\Omega_\mathbf{k}$,

$$\frac{d\sigma}{d\omega} = Z_p^2\alpha\frac{\omega}{2p_1^2}\sum_{l=1}^{\infty}\frac{l(l+1)}{2l+1}$$
$$\times\int\limits_{q_{\min}}^{q_{\max}}dq^2\left[\left(\frac{q^2-\omega^2}{2} + 2(\mathbf{p}_1\cdot\mathbf{n}_\mathbf{q})(\mathbf{p}_2\cdot\mathbf{n}_\mathbf{q})\right)|\alpha_l(\omega,q,k)|^2\right. \quad (42)$$
$$+\frac{q^2-\omega^2+2p_1^2\sin^2\theta_\mathbf{q}}{2(q^2-\omega^2)^2}l(l+1)$$
$$\left.\times\sum_{\lambda=0,1}|\beta_l^{(\lambda)}(\omega,q,k)|^2\right].$$

An important characteristic of the relativistic polarization bremsstrahlung cross section is its logarithmic growth as the energy of the incident particle increases [3]. Qualitatively, the reason for this is as follows. Unlike a nonrelativistic particle, a relativistic one interacts with the target not only by its Coulomb field but also (in the ultrarelativistic case, predominantly) by the field of transverse virtual photons (e.g., see [12]). The effective radius of this field increases as the energy of the incident particle grows, almost to infinity in the ultrarelativistic case. As a result, the distances at which an incident particle can effectively polarize a target increase; accordingly, the polarization bremsstrahlung cross section increases. We will analyze this conclusion below.

We assume that the initial and final energies of the incident particle and the emitted photon energy satisfy the relation $\varepsilon_1 \approx \varepsilon_2 \gg \omega$. It can then be taken that

$$q_{\min} \approx \frac{\omega}{v_1}, \quad q_{\max} \approx 2p_1,$$

where v_1 is the initial velocity of the incident particle.

As an incident relativistic particle effectively polarizes the target even if passes at a large distance from it, the characteristic transferred momenta are small, $q \sim q_{\min} \sim R_{at}^{-1}$. Formally, the $f_{n0}^{(-1, 0, 1)}(q; l)$ matrix elements in (25) and (26) contain a rapidly oscillating function at $qR_{at} \gg 1$ and are therefore close to zero.

For brevity, let $\mathcal{H}_{l_k}^{A, B, C, D}(\omega, q, k)$ denote the terms to the right of the kinematic coefficients in (34)–(37), which are smooth functions of q. The contribution of the $q \sim q_{\min}$ region to integral (33) can then be written in the form

$$\left[\frac{d\sigma_{l_k}}{d\omega}\right]_{q \sim q_{\min}} = Z_p^2 \alpha \frac{\omega}{8\pi p_1^2} \int_{q_{\min}}^{q_0} dq^2$$

$$\times \left[\left(\frac{q^2 - \omega^2}{4} + (\mathbf{p}_1 \cdot \mathbf{n}_\mathbf{q})(\mathbf{p}_2 \cdot \mathbf{n}_\mathbf{q})\right)\mathcal{H}_{l_k}^A(\omega, q, k)\right.$$

$$+ \frac{q^2 - \omega^2 + 2p_1^2 \sin^2\theta_\mathbf{q}}{2(q^2 - \omega^2)^2}\mathcal{H}_{l_k}^B(\omega, q, k) \qquad (43)$$

$$+ \frac{p_1^2 \sin^2\theta_\mathbf{q}}{(q^2 - \omega^2)^2}\mathcal{H}_{l_k}^C(\omega, q, k)$$

$$\left. + \frac{4p_1(\mathbf{p}_2 \cdot \mathbf{n}_\mathbf{q})\sin^2\theta_\mathbf{q}}{q^2 - \omega^2}\mathcal{H}_{l_k}^D(\omega, q, k)\right].$$

Here, q_0 is the cutoff parameter, which satisfies the inequality $q_{\max} \gg q_0 \gg q_{\min}$ and is close to $q_0 R_{at} \sim 1$ in order of magnitude.

Let us integrate each of the four terms in (43) with respect to q assuming that $\mathcal{H}_{l_k}^{A, B, C, D}(\omega, q, k) \approx \mathcal{H}_{l_k}^{A, B, C, D}(\omega, q_{\min}, k)$ and rejecting small terms of the order of $\omega/\varepsilon_{1, 2} \ll 1$. This gives

$$\left[\frac{d\sigma_{l_k}}{d\omega}\right]_{q \sim q_{\min}} \approx Z_p^2 \alpha \frac{\omega^3}{4\pi v_1^2}\left\{\mathcal{H}_{l_k}^A(\omega, q_{\min}, k)\ln\left(\frac{q_0}{q_{\min}}\right)\right.$$

$$+ \frac{1}{2\varepsilon_1^2\omega^2}\mathcal{H}_{l_k}^B(\omega, q_{\min}, k)$$

$$\times \left[\ln\left(\frac{q_0\gamma}{q_{\min}}\right) + \frac{p_1^2}{\omega^2}\left(\frac{2}{v_1^2}\ln\gamma - 1\right)\right] \qquad (44)$$

$$+ \frac{1}{\omega^4}\mathcal{H}_{l_k}^C(\omega, q_{\min}, k)\left[\ln\gamma - \frac{v_1^2}{2}\right]$$

$$\left. + \frac{4}{\omega^2}\mathcal{H}_{l_k}^D(\omega, q_{\min}, k)\left[1 - \frac{1}{2v_1\gamma^2}\ln\left(\frac{1 + v_1}{1 - v_1}\right)\right]\right\}.$$

Here, the $\gamma = \varepsilon_1/m_p c^2$ relativistic parameter is introduced. The spectral radiation distribution takes the form

$$\left[\frac{d\sigma}{d\omega}\right]_{q \sim q_{\min}} \approx Z_p^2 \alpha \frac{\omega^3}{v_1^2}\left\{\mathcal{H}_0^A(\omega, q_{\min}, k)\ln\left(\frac{q_0}{q_{\min}}\right)\right.$$

$$+ \frac{1}{2\varepsilon_1^2\omega^2}\mathcal{H}_0^B(\omega, q_{\min}, k) \qquad (45)$$

$$\left. \times \left[\ln\left(\frac{q_0\gamma}{q_{\min}}\right) + \frac{p_1^2}{\omega^2}\left(\frac{2}{v_1^2}\ln\gamma - 1\right)\right]\right\},$$

where the $\mathcal{H}_0^A(\omega, q, k)$ and $\mathcal{H}_0^B(\omega, q, k)$ values can be written in the form

$$\mathcal{H}_0^A(\omega, q, k) = \sum_{l=1}^{\infty} \frac{2l(l+1)}{2l+1}|\alpha_l(\omega, q, k)|^2,$$

$$\mathcal{H}_0^B(\omega, q, k) = \sum_{l=1}^{\infty} \frac{l^2(l+1)^2}{2l+1} \sum_{\lambda = 0, 1} |\beta_l^{(\lambda)}(\omega, q, k)|^2.$$

The behavior of the polarization bremsstrahlung cross section at $\varepsilon_{1, 2} \gg \omega$ is therefore determined by the terms proportional to $\ln(q_{\min}^{-1} R_{at}^{-1})$ and $\ln\gamma$ which correspond to the contributions of the longitudinal (Coulomb) and transverse (vector) parts of the electromagnetic interaction between a relativistic incident particle and target electrons. The contribution of the terms with $\ln\gamma$ increases as the energy of the incident particle grows and becomes predominant at $\gamma \gg 1$, which causes an increase in the polarization bremsstrahlung cross section. This was for the first time noted in [19], where polarization bremsstrahlung caused by a collision between a relativistic particle and a target was treated at the level of the dipole approximation.

3. CALCULATION DATA ON POLARIZATION BREMSSTRAHLUNG CROSS SECTIONS

The results described in this section (except those shown in Fig. 4) refer to proton collisions with the Al^{+12}, Ag^{+46}, and Au^{+78} hydrogen-like ions. Because of

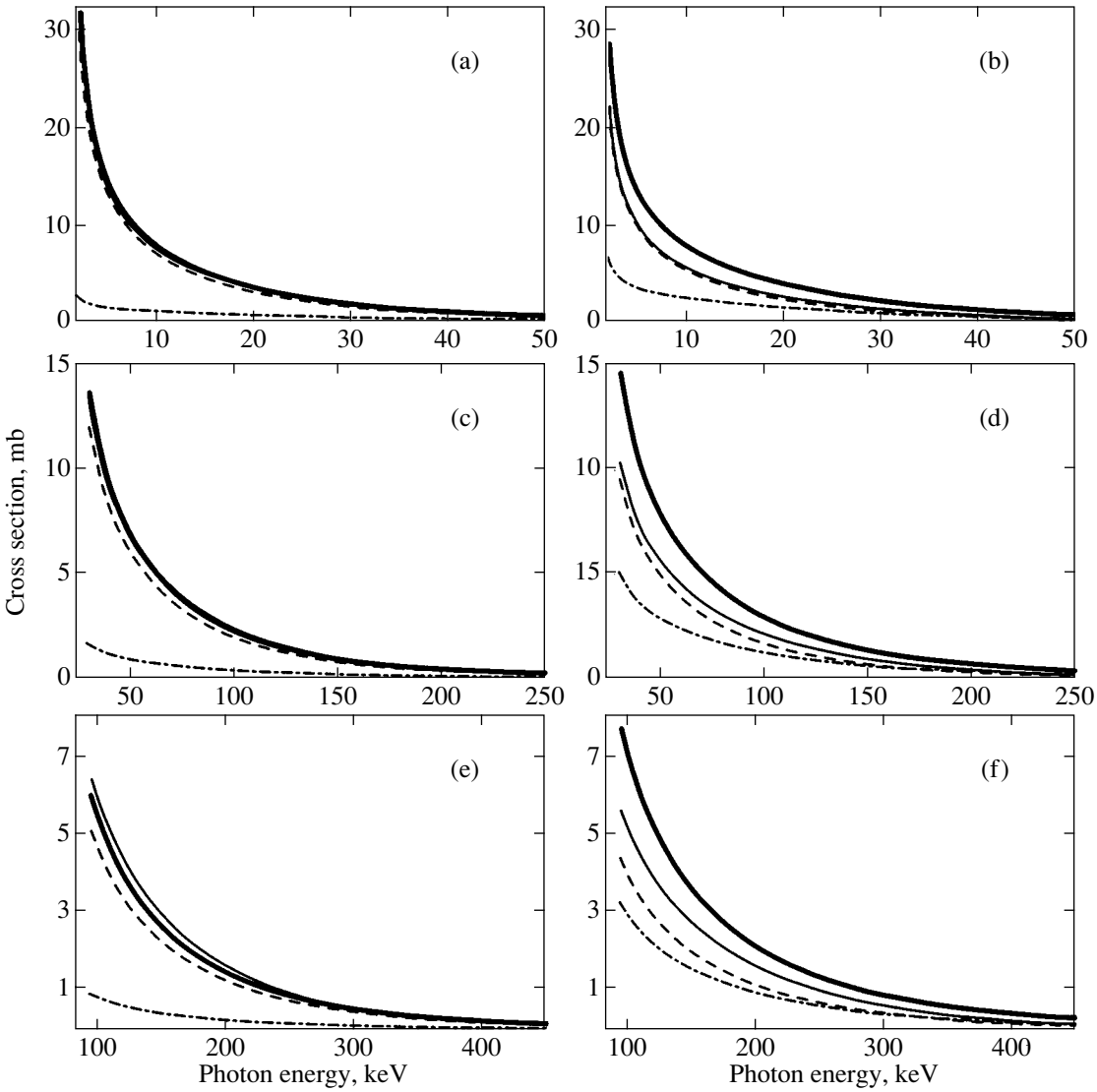

Fig. 2. Spectral dependences of $\omega d\sigma/d\omega$ for polarization bremsstrahlung arising in collisions of protons with (a, c, and e) 1.5 GeV and (b, d, and f) 3 GeV energies with (a and b) Al^{+12}, (c and d) Ag^{+46}, and (e and f) Au^{+78} ions. The thick solid line describes the behavior of relativistic cross section (42), and the thin solid line corresponds to the nonrelativistic dipole case. The contributions of the terms proportional to the squares of the $\alpha_l(\omega, q, k)$ and $\beta_l^{(1)}(\omega, q, k)$ polarizabilities [see (42)] are shown by dashed and dot-and-dash lines, respectively. The contribution of the term proportional to the square of the modulus of the $\beta_l^{(0)}(\omega, q, k)$ polarizability is small and is not shown.

the large mass of the proton, the mechanism of usual bremsstrahlung is suppressed by the $m_p^{-2} \sim 10^{-6}$ factor and can therefore be excluded from consideration in a wide range of emitted photon energies [21–24]. In addition, in the range of emitted photon energies that we are considering, which are close to the ionization potentials of inner target shells, polarization bremsstrahlung predominates over the other radiation mechanisms such as secondary electron bremsstrahlung [45], radiation ionization [46, 13, 20], and molecular orbital radiation [47].

The polarization bremsstrahlung cross sections were calculated at the relativistic Born approximation level by (42) and (32) for spectral and spectral-angular radiation distributions. The contributions of the first five multipoles were taken into account. The dynamic response of targets was described by a method based on representing partial polarizabilities with the use of a relativistic Coulomb Green's function [37]. The curves that corresponded to the nonrelativistic dipole case were obtained as described in [28].

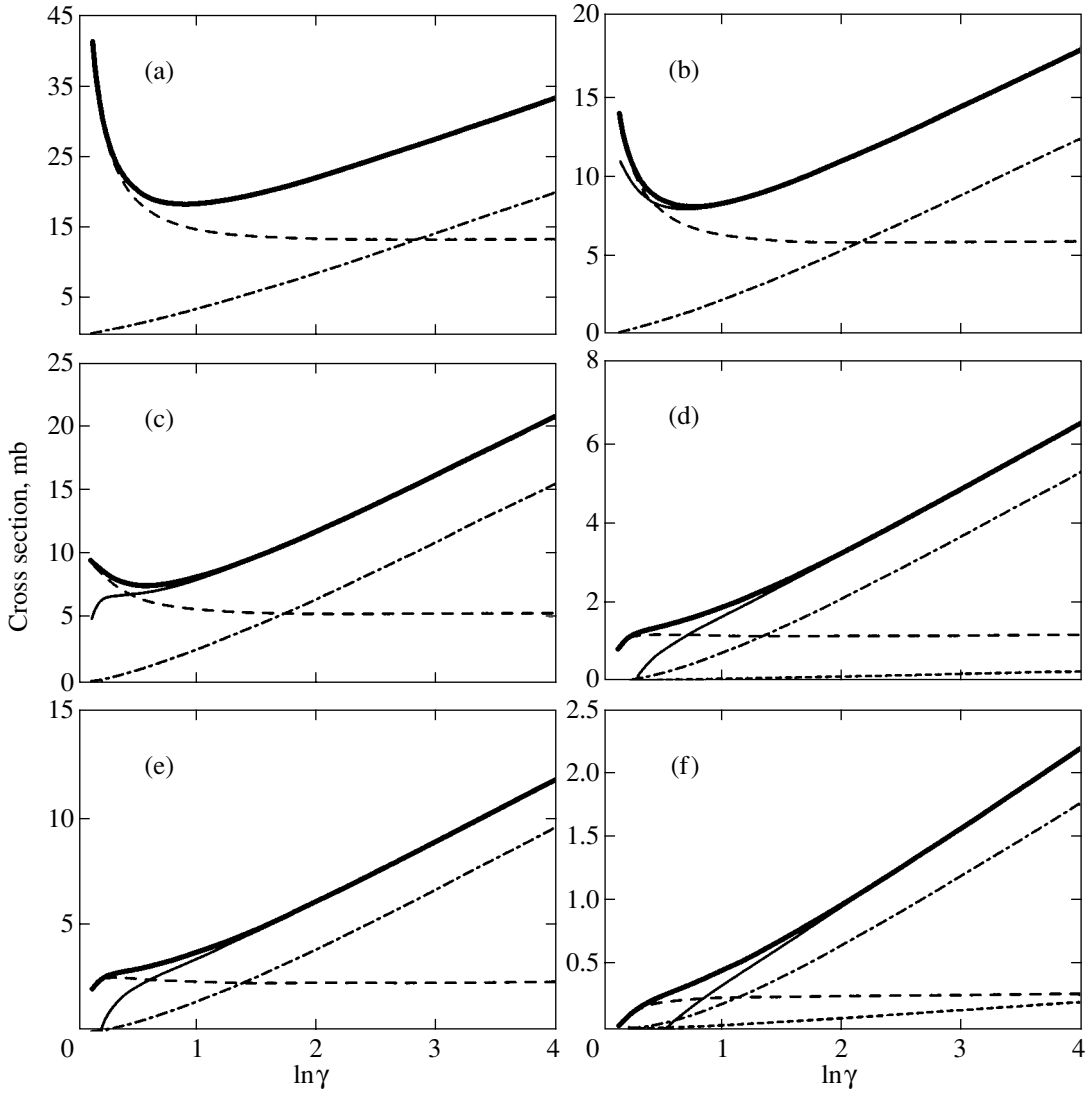

Fig. 3. Dependence of polarization bremsstrahlung cross section $\omega d\sigma/d\omega$ on incident proton relativistic γ factor in collisions with (a and b) Al^{+12}, (c and d) Ag^{+46}, and (e and f) Au^{+78} for emitted photon energies ω = (a, c, and e) $1.5I$ and (b, d, and f) $4I$ (here, I is the ionization potential of the 1s target subshell, which approximately equals 2.3, 31, and 93.5 keV for Al^{+12}, Ag^{+46}, and Au^{+78}, respectively). The thick solid line depicts relativistic cross section (42). The behavior of the terms proportional to the squares of the $\alpha_l(\omega, q, k)$, $\beta_l^{(0)}(\omega, q, k)$, and $\beta_l^{(1)}(\omega, q, k)$ polarizabilities [see (42)] is shown by dashed, dotted, and dot-and-dash lines, respectively.

Target inner shell electrons make the major contribution to the formation of the complete polarization bremsstrahlung spectrum in the region of photon frequencies higher than the ionization potentials of the corresponding shells (e.g., see [3]). The results given below can therefore easily be generalized to neutral Al, Ag, and Au atoms by multiplying the cross sections obtained for the hydrogen-like ions by a factor of 4, which takes into account a twofold increase in the polarizabilities of filled atomic K shells.

The spectral dependences of $\omega d\sigma/d\omega$ calculated for two incident proton energies, ε_1 = 1.5 and 3 GeV, are shown in Fig. 2. Note two special features of these plots. First, the contribution of the cross section part proportional to the square of polarizability $\beta_l^{(1)}(\omega, q, k)$ modulus [see (42)] becomes more noticeable as the energy of the incident particle increases because of a relative increase in the contribution of the mechanism of exchange of transverse virtual photons between target and incident particle electrons as the energy of the latter grows, whereas the contribution of the Coulomb interaction part is only determined by the particle velocity [see (45)] and is virtually independent of ε_1 at $v_1 \approx c$.

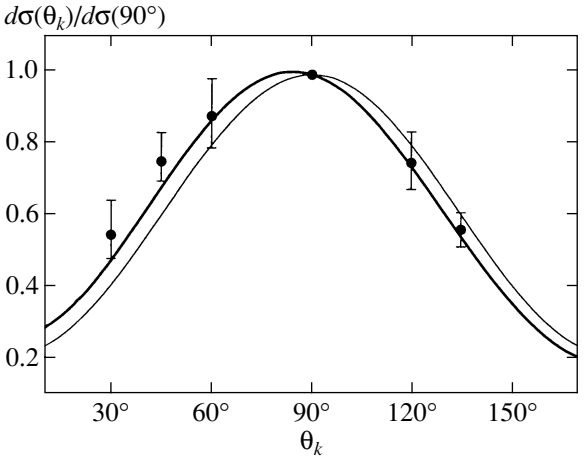

Fig. 4. Angular dependence of polarization bremsstrahlung cross section (normalized with respect to the cross section value at $\theta_{\mathbf{k}} = 90°$) for scattering of a proton with a 1.5 MeV kinetic energy on an aluminum atom and emitted photon energies in the range 5.18–5.67 keV. The thick solid line corresponds to cross section (32), and the thin solid line was obtained using the nonrelativistic dipole approximation. Solid circles are the experimental data (see [15]).

The second point to be mentioned is the ratio between the relativistic and nonrelativistic curves. At intermediate energies $\varepsilon_1 \sim m_p c^2$, the $\omega d\sigma/d\omega$ cross section calculated by the nonrelativistic dipole approximation exceeds the relativistic polarization bremsstrahlung cross section, because, at a given photon energy, nonrelativistic dipole polarizability $\alpha_d(\omega, q)$ exceeds the corresponding relativistic component $\alpha_1(\omega, q, k)$, which makes the largest contribution to the spectral dependence of polarization bremsstrahlung at medium incident particle energies (a more detailed analysis of the contributions of various multipoles to the polarization bremsstrahlung cross section is given below in this section). The contribution of the term proportional to $\left|\beta_l^{(1)}(\omega, q, k)\right|^2$ becomes more significant as the incident particle energy increases. The relativistic cross section therefore begins to increase and exceeds the nonrelativistic cross section. Such a behavior follows from the analysis of the integrand in (33) given in the preceding section.

The $\omega d\sigma/d\omega$ cross section and the contributions of separate cross section parts proportional to the squares of the moduli of the corresponding polarizabilities are shown in Fig. 3 as functions of the $\gamma = \varepsilon_1/m_p c^2$ relativistic factor of the incident proton for two emitted photon energies specified above and for each of the three ions under consideration, Al^{+12}, Ag^{+46}, and Au^{+78}. These results visually demonstrate a logarithmic growth of the polarization bremsstrahlung cross section with increasing the energy of the incident particle.

An important property of the approach suggested in this work is taking into account effects related to radiation lag and radiation of high-multipolarity photons. The influence of the multipole character of radiation manifests itself even at low incident particle energies $\varepsilon_1 \sim m_p c^2$ by asymmetry of the angular radiation distribution. The angular radiation distributions calculated by (32) are compared with the results obtained using the nonrelativistic dipole approximation and the experimental data obtained in [15] in Fig. 4. These curves describe scattering of protons with a 1.5 MeV kinetic energy on aluminum atoms. The emitted photon energies are in the range 5.18–5.67 keV, in which polarization bremsstrahlung predominates over the other radiation mechanisms [15]. The calculations were performed for hydrogen-like wave functions with an effective target charge of $Z_T = 12.7$ (see [48]). Note that, in contrast to the symmetrical curve of the nonrelativistic dipole approximation, the curve corresponding to cross section (32), which takes into account the multipole character of radiation, reproduces well the observed radiation shift in the incident particle direction. This result is already obtained by taking into account quadrupole corrections, which allows us to claim the validity of the approach suggested in this work and recommend it for calculating polarization bremsstrahlung cross sections.

An increase in the incident particle energy not only causes the appearance of high-multipolarity photon radiation effects but also increases the contribution of relativistic effects related to radiation lag, which results in still more substantial differences between the angular radiation distributions calculated in the relativistic and nonrelativistic approximations. The importance of taking these effects into account in considering polarization bremsstrahlung caused by collisions between relativistic incident particles and a heavy target follows from Fig. 5, where we plotted the spectral-angular radiation distribution profiles $\omega d^2\sigma/d\omega d\Omega_{\mathbf{k}}$ for $\varepsilon_1 = 3$ GeV proton collisions with the Al^{+12}, Ag^{+46}, and Au^{+78} ions and the specified emitted photon energies. In these plots (and also in Figs. 6 and 7), the length of the segment connecting the origin and a curve point equals the differential polarization bremsstrahlung cross section (in millibarn units) in the corresponding direction. The direction along the horizontal axis ($\theta_{\mathbf{k}} = 0$) is the direction of incident particle motion. The curves describing the contributions of cross section components proportional to the squares of the moduli of the $\alpha_l(\omega, q, k)$, $\beta_l^{(0)}(\omega, q, k)$, and $\beta_l^{(1)}(\omega, q, k)$ polarizabilities are also shown in Fig. 5. In contrast to the spectral distribution, the sum of these curves does not correspond to total cross section (32), which also takes into account cross terms.

The results obtained in this work show that taking into account relativistic effects and the effects related to the radiation of high-multipolarity photons noticeably

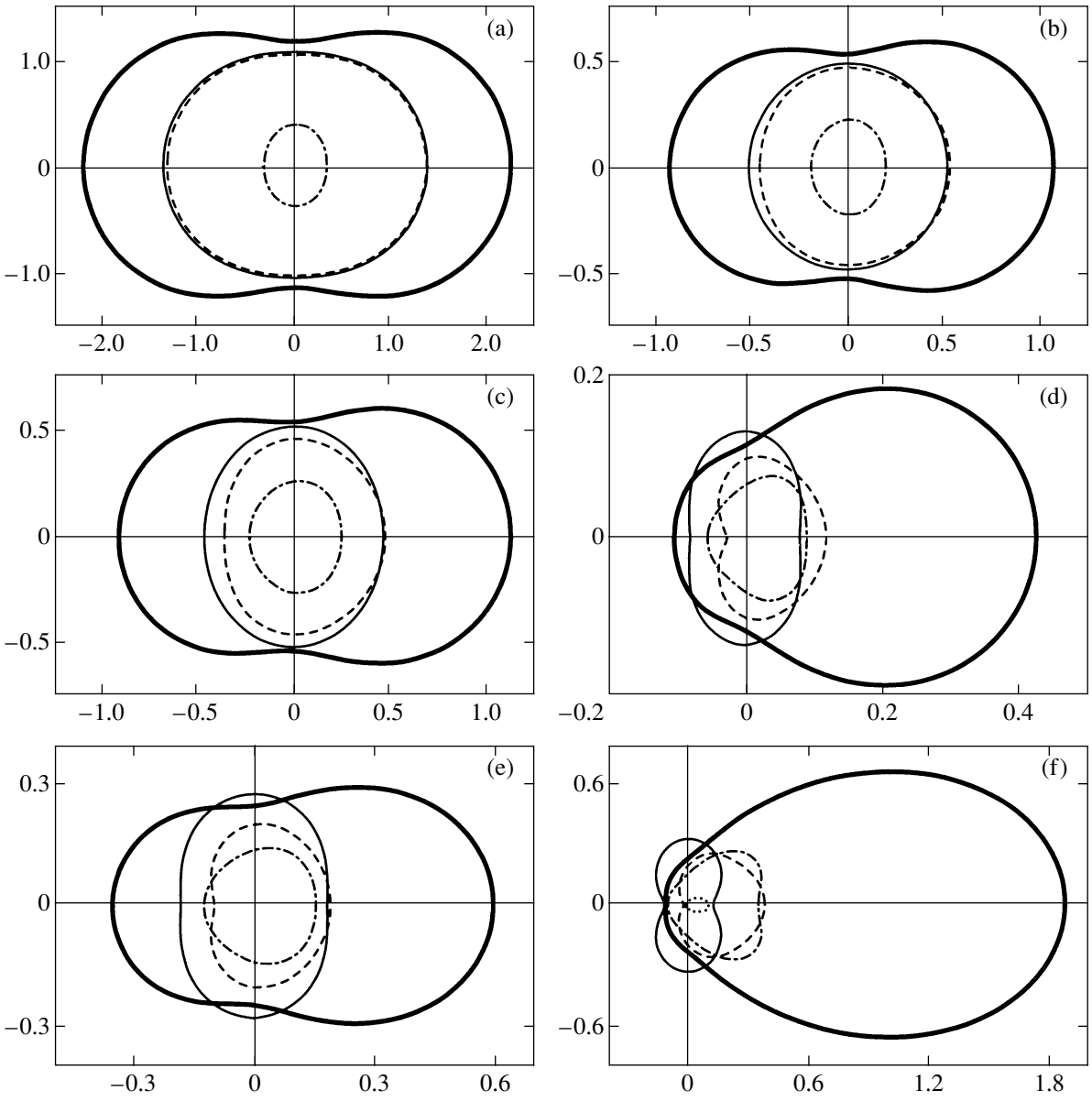

Fig. 5. Angular distribution $\omega d^2\sigma/d\omega d\Omega_{\mathbf{k}}$ for polarization bremsstrahlung arising in collisions between 3 GeV protons and (a and b) Al^{+12}, (c and d) Ag^{+46}, and (e and f) Au^{+78} ions for two emitted photon energies of (a, c, and e) 1.5*I* and (b, d, and f) 4*I* (see Fig. 3). The thick solid line is relativistic cross section (32), and the thin solid line was obtained in the nonrelativistic dipole approximation. The dashed, dotted, and dot-and-dash lines correspond to the contributions of the terms proportional to the squares of the moduli of the $\alpha_l(\omega, q, k)$, $\beta_l^{(0)}(\omega, q, k)$, and $\beta_l^{(1)}(\omega, q, k)$ polarizabilities, respectively [see (32)–(37)].

changes the angular distributions of emitted photons and makes these distributions appreciably asymmetric compared with the distributions of the nonrelativistic dipole approximation, which are typically symmetrical with respect to the $\theta_{\mathbf{k}} \longrightarrow \pi - \theta_{\mathbf{k}}$ operation. The relativistic angular distributions are shifted in the direction in which the incident particle moves, and the distributions become more asymmetric as the emitted photon energy increases.

The contributions of separate multipoles to the angular radiation distribution for proton collisions with the Ag^{+46} and Au^{+78} ions (at the same parameter values as in Fig. 5) are analyzed in Fig. 6. The curves shown in Fig. 6 were obtained by Eqs. (34)–(37) in which only the terms with l, $l' = 1$ for the dipole approximation, l, $l' = 1, 2$ for taking into account quadrupole corrections, and l, $l' = 1, 2, 3$ for octupole corrections were retained.

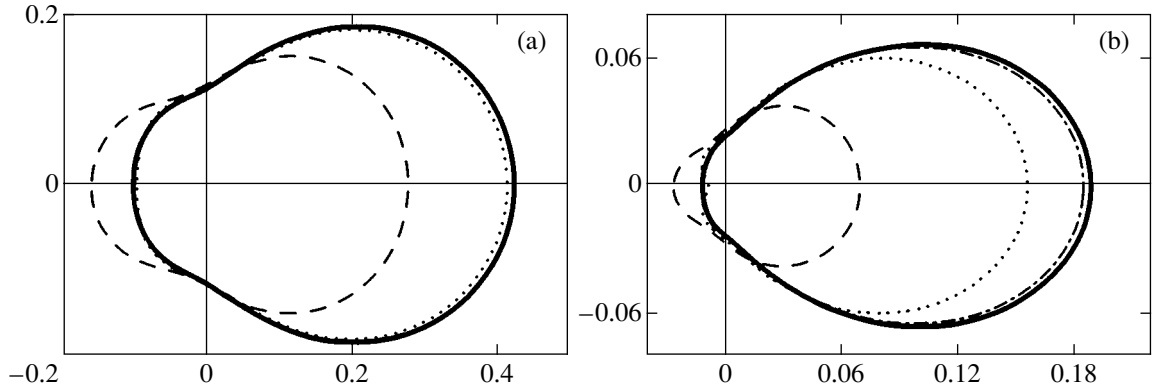

Fig. 6. Multipole contributions to the angular polarization bremsstrahlung distribution $\omega d^2\sigma/d\omega d\Omega_{\mathbf{k}}$ formed in collisions between protons of a 3 GeV energy with (a) Ag^{+46} and (b) Au^{+78} for emitted photon energies equal to four times the ionization potentials of the target 1*s* subshells. The dashed, dotted, and dot-and-dash lines are the angular distribution profiles taking into account the dipole, quadrupole, and octupole radiation contributions, respectively. The thick solid line corresponds to the exact result taking into account the contributions of the first five multipoles.

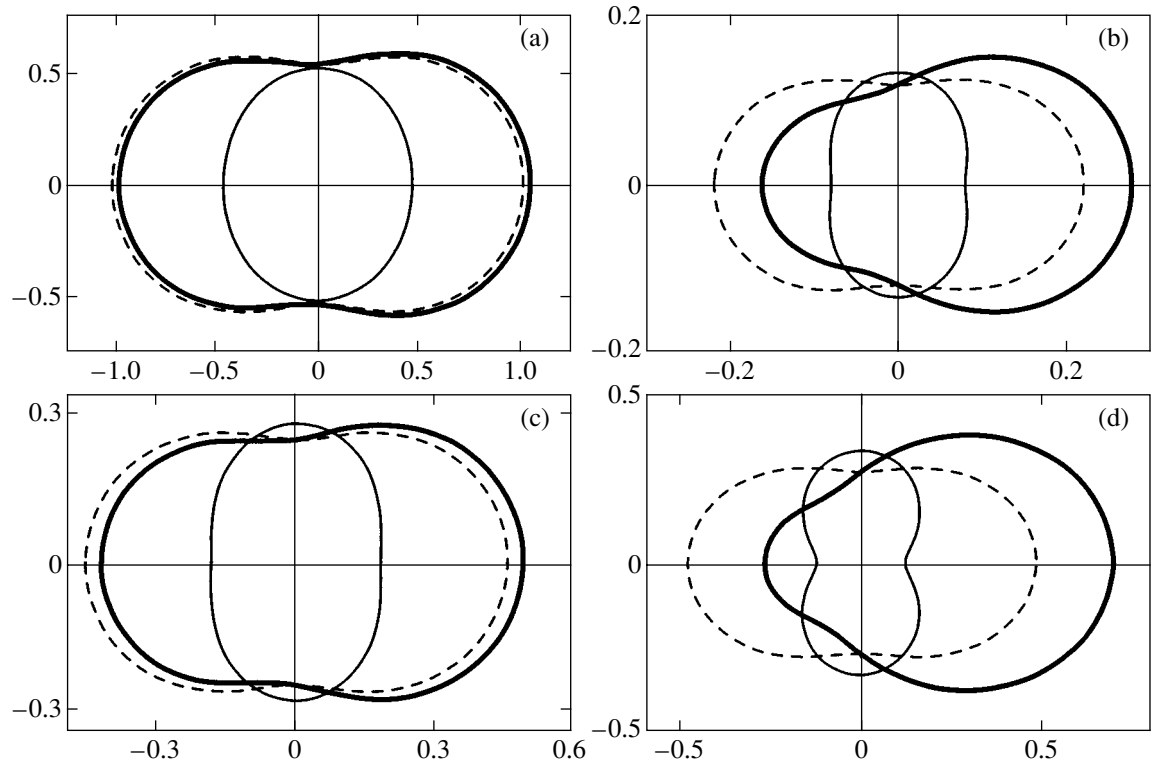

Fig. 7. Angular polarization bremsstrahlung distribution $\omega d^2\sigma/d\omega d\Omega_{\mathbf{k}}$ for collisions of 3 GeV protons with (a and b) Ag^{+46} and (c and d) Au^{+78} and two emitted photon energies of (a and c) 1.5*I* and (b and d) 4*I* (see Fig. 5). The thick solid line was obtained in the relativistic dipole approximation by (46), and the thin solid line was obtained in the nonrelativistic dipole approximation. The dashed line is the summed contribution to the cross section of terms with $l_k = 0$ and 2, respectively, in the dipole approximation [see (46)].

Note that, apart from the contributions of quadrupole and octupole corrections, which to a substantial extent determine asymmetric angular radiation distribution shapes, the curves obtained in the dipole approximation also have asymmetric shapes substantially different from those obtained in the nonrelativistic dipole approximation, because, in contrast to the nonrelativistic approximation, cross section (32) takes into account cross term contributions. These contributions are proportional to the products of polarizabilities

$\beta_1^{(0)}(\omega, q, k)$ and $\beta_1^{(1)}(\omega, q, k)$ and $\beta_1^{(0)}(\omega, q, k)$ and $\alpha_1(\omega, q, k)$ corresponding to interference between radiated photons of different types. Indeed, double differential cross section (32) for the dipole case can be written as

$$\left[\frac{d^2\sigma}{d\omega d\Omega_{\mathbf{k}}}\right]_{\text{dip}} = \left(\frac{d\sigma}{d\omega}\right)_{\text{dip}} \times (1 + a_1(\omega)P_1(\cos\theta_{\mathbf{k}}) + a_2(\omega)P_2(\cos\theta_{\mathbf{k}})), \quad (46)$$

where a_1 and a_2 are the angular anisotropy coefficients in the dipole approximation. These coefficients can be obtained in an explicit form from (32)–(37); the corresponding equations are, however, too cumbersome to be given here. Their dependence on the energy of emitted photons for scattering of 3 GeV protons on the Ag^{+46} and Au^{+78} ions is shown in Fig. 8 to shed light on the reason why the dipole curves shown in Fig. 6 are asymmetric. Indeed, while the contribution of the terms related to the a_1 coefficient is small for low emitted photon energies ($kr \ll 1$), it rapidly grows and becomes predominant as ω increases. This alongside higher multipole contributions results in the distributions shown in Fig. 6.

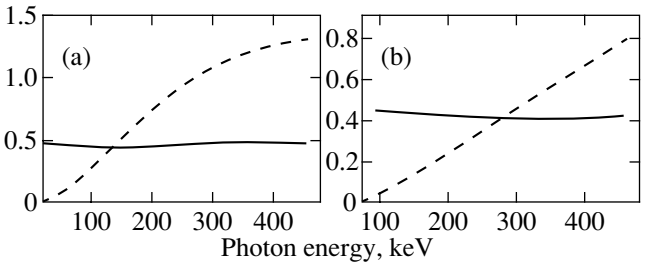

Fig. 8. Dependence of angular anisotropy coefficients for dipole cases a_1 (dashed line) and a_2 (solid line) [see (46)] on the energy of photons emitted in collisions between 3 GeV protons and (a) Ag^{+46} and (b) Au^{+78}.

4. CONCLUSION

We suggested a completely relativistic formalism for describing polarization bremsstrahlung caused by elastic collisions between charged particles and many-electron targets. This approach can be used to determine spectral and angular radiation distributions in wide ranges of collision energies and emitted photon frequencies. With slight modifications, it can also be applied to describe a wide class of collisional systems with internal structures in which relativistic effects play an important role. For instance, this approach can be used to describe polarization bremsstrahlung arising in relativistic collisions of nuclei. The dynamic polarization of colliding nuclei then causes photon emission by the polarization bremsstrahlung mechanism, and the major contribution is made by nondipole radiation (quadrupole radiation and radiation of higher multipolarities).

If the relativistic incident particle possesses an internal structure, we need not develop a new approach taking into account the polarization bremsstrahlung of this particle. The required formulas can be obtained from those given in this work by taking into account the special features caused by the Doppler effect and light aberration, as is done in [49, 50] for relativistic atom–atom collisions. In these works, the internal structure of colliding particles was described in the nonrelativistic dipole approximation.

The formalism described above was used to develop an effective method for calculating polarization bremsstrahlung cross sections. This method is based on the use of the hydrogen-like approximation for describing the dynamic atomic response. This approach is advantageous because the hydrogen-like model allows the polarization bremsstrahlung cross sections to be determined very accurately at photon frequencies exceeding the ionization potentials of inner target shells, which obviates the necessity of complex numerical calculations. Note that various approximations to the relativistic wave functions of the target can be used within the framework of the suggested approach. Note that the Hartree–Fock–Dirac approximation allows many-electron correlations to be taken into account in calculating the dynamic polarizabilities of targets. In addition, an effective approach to calculating polarizabilities can be based on the Sternheimer method. In essence, this method reduces summing over an infinite number of intermediate atomic states to solving inhomogeneous differential equations (e.g., see [51]).

The suggested method was used to calculate the spectral and angular characteristics of polarization bremsstrahlung arising when a heavy particle (proton) is scattered on the Al^{+12}, Ag^{+46}, and Au^{+78} hydrogen-like ions, which allowed us to study some general properties of polarization bremsstrahlung cross sections in the relativistic case. Taking into account relativistic and multipole effects was shown to result in substantial asymmetry of the angular distributions of emitted photons compared with the nonrelativistic dipole case and in noticeable changes in the spectral characteristics of polarization radiation. We also showed that the polarization bremsstrahlung cross section logarithmically increased as the energy of the incident particle grew. The results described in this work may be useful in setting up new polarization bremsstrahlung experiments and repeating those performed earlier but at a new technical level.

An interesting problem, which deserves special consideration, is an analysis of the asymptotic behavior of the completely relativistic bremsstrahlung cross section including the contributions of both the usual and the polarization mechanism in the region of high photon energies comparable with the rest mass of the electron. A nonrelativistic description of atomic response then

becomes inapplicable, and using more complex approaches to describe atom electrons excited to the continuum of negative frequencies becomes necessary.

An analysis of inelastic polarization bremsstrahlung is outside the scope of this work. The development of a completely relativistic formalism for describing these processes is of special interest.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 99-02-18294a) and the Russian Academy of Sciences (the Sixth Competition of Scientific Projects of Young Scientists in Basic and Applied Studies of the Russian Academy of Sciences, grant 44).

APPENDIX

The Representation of Generalized Polarizabilities with the Use of Relativistic Green's Function

The multipole bispinor structure of the one-particle relativistic Green's function is determined by the equation (e.g., see [52])

$$G_E(\mathbf{r}_2, \mathbf{r}_1) = \sum_{\varepsilon jlm} \frac{\Psi_{\varepsilon jlm}(\mathbf{r}_2)\Psi^{\dagger}_{\varepsilon jlm}(\mathbf{r}_1)\gamma^0}{E - \varepsilon(1 - i0)} = \frac{1}{r_2 r_1} \tag{47}$$

$$\times \sum_{jl} \begin{pmatrix} G^{(1)}_{Ejl}(r_2, r_1) Y_{ll}(\mathbf{n}_2, \mathbf{n}_1) & iG^{(2)}_{Ejl}(r_2, r_1) Y_{ll'}(\mathbf{n}_2, \mathbf{n}_1) \\ iG^{(3)}_{Ejl}(r_2, r_1) Y_{l'l}(\mathbf{n}_2, \mathbf{n}_1) & -G^{(4)}_{Ejl}(r_2, r_1) Y_{l'l'}(\mathbf{n}_2, \mathbf{n}_1) \end{pmatrix},$$

where

$$Y_{l_1 l_2}(\mathbf{n}_2, \mathbf{n}_1) = \sum_{m=-j}^{j} \Omega_{jl_1 m}(\mathbf{n}_2)\Omega^{\dagger}_{jl_2 m}(\mathbf{n}_1)$$

and the radial parts of the Green's function are expressed in terms of the radial functions of the large $g_{\varepsilon jl}(r)$ and small $f_{\varepsilon jl}(r)$ excited electron wave function components,

$$\begin{pmatrix} G^{(1)}_{Ejl}(r_2, r_1) & G^{(2)}_{Ejl}(r_2, r_1) \\ G^{(3)}_{Ejl}(r_2, r_1) & G^{(4)}_{Ejl}(r_2, r_1) \end{pmatrix} = \sum_{\varepsilon} \frac{1}{E - \varepsilon(1 - i0)} \tag{48}$$

$$\times \begin{pmatrix} g_{\varepsilon jl}(r_2) g^*_{\varepsilon jl}(r_1) & g_{\varepsilon jl}(r_2) f^*_{\varepsilon jl}(r_1) \\ f_{\varepsilon jl}(r_2) g^*_{\varepsilon jl}(r_1) & f_{\varepsilon jl}(r_2) f^*_{\varepsilon jl}(r_1) \end{pmatrix}.$$

Using (48) in (25) and (26), we can write the $\alpha_l(\omega, q, k)$ and $\beta^{(0,1)}_l(\omega, q, k)$ polarizabilities in terms of $G^{(1-4)}_{Ejl}(r_2, r_1)$ as follows:

$$\alpha_l(\omega, q, k) = \frac{e^2}{q\omega}(2l + 1) \sum_{\varepsilon_0 j_0 l_0} \sum_{j_n l_n} C^{(-1)}_{0n}(l)$$

$$\times \sum_{i=1,2} (-1)^i \int_0^\infty \int_0^\infty dr_1 dr_2 j_l(qr_1)(f^*_0(r_2), g^*_0(r_2)) \tag{49}$$

$$\times \begin{pmatrix} I^-_2(r_2) G^{(1)}_{\varepsilon_i j_n l_n}(r_2, r_1) & I^-_2(r_2) G^{(2)}_{\varepsilon_i j_n l_n}(r_2, r_1) \\ I^+_2(r_2) G^{(3)}_{\varepsilon_i j_n l_n}(r_2, r_1) & I^+_2(r_2) G^{(4)}_{\varepsilon_i j_n l_n}(r_2, r_1) \end{pmatrix} \begin{pmatrix} g_0(r_1) \\ f_0(r_1) \end{pmatrix},$$

$$\beta^{(0)}_l(\omega, q, k) = e^2(2l + 1) \sum_{\varepsilon_0 j_0 l_0} \sum_{j_n l_n} C^{(0)}_{0n}(l)$$

$$\times \sum_{i=1,2} \int_0^\infty \int_0^\infty dr_1 dr_2 j_l(qr_1) j_l(kr_2)(f^*_0(r_2), g^*_0(r_2)) \tag{50}$$

$$\times \begin{pmatrix} G^{(1)}_{\varepsilon_i j_n l_n}(r_2, r_1) & G^{(2)}_{\varepsilon_i j_n l_n}(r_2, r_1) \\ G^{(3)}_{\varepsilon_i j_n l_n}(r_2, r_1) & G^{(4)}_{\varepsilon_i j_n l_n}(r_2, r_1) \end{pmatrix} \begin{pmatrix} f_0(r_1) \\ g_0(r_1) \end{pmatrix},$$

$$\beta^{(1)}_l(\omega, q, k) = e^2(2l + 1) \sum_{\varepsilon_0 j_0 l_0} \sum_{j_n l_n} C^{(0)}_{0n}(l)$$

$$\times \sum_{i=1,2} \int_0^\infty \int_0^\infty dr_1 dr_2 (f^*_0(r_2), g^*_0(r_2)) \tag{51}$$

$$\times \begin{pmatrix} I^+_1(r_1) I^-_2(r_2) G^{(1)}_{\varepsilon_i j_n l_n}(r_2, r_1) & I^-_1(r_1) I^-_2(r_2) G^{(2)}_{\varepsilon_i j_n l_n}(r_2, r_1) \\ I^+_1(r_1) I^+_2(r_2) G^{(3)}_{\varepsilon_i j_n l_n}(r_2, r_1) & I^-_1(r_1) I^+_2(r_2) G^{(4)}_{\varepsilon_i j_n l_n}(r_2, r_1) \end{pmatrix}$$

$$\times \begin{pmatrix} f_0(r_1) \\ g_0(r_1) \end{pmatrix}.$$

Here, $\varepsilon_1 = \varepsilon_0 - \omega$, $\varepsilon_2 = \varepsilon_0 + \omega$, and the $I^{\pm}_{1,2}(r)$ function is defined as

$$I^{\pm}_1(r) = \frac{\kappa_0 - \kappa_n}{l(l + 1)} \left(\frac{1}{q} \frac{dj_l(qr)}{dr} + \frac{j_l(qr)}{qr} \right) \pm \frac{j_l(qr)}{qr}, \tag{52}$$

$$I^{\pm}_2(r) = \frac{\kappa_n - \kappa_0}{l(l + 1)} \left(\frac{1}{k} \frac{dj_l(kr)}{dr} + \frac{j_l(kr)}{kr} \right) \pm \frac{j_l(kr)}{kr}. \tag{53}$$

REFERENCES

1. A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics* (Nauka, Moscow, 1969; Wiley, New York, 1965).
2. V. B. Berestetskii, E. M. Lifshitz, and L. P. Pitaevskii, *Course of Theoretical Physics*, Vol. 4: *Quantum Electrodynamics* (Nauka, Moscow, 1989; Pergamon, New York, 1982).

3. M. Ya. Amus'ya, V. M. Buĭmistrov, B. A. Zon, V. N. Tsytovich, *et al.*, *Polarization Bremsstrahlung of Particles and Atoms* (Nauka, Moscow, 1987).
4. M. Ya. Amus'ya, A. S. Baltenkov, and A. A. Paĭziev, Pis'ma Zh. Éksp. Teor. Fiz. **24**, 366 (1976) [JETP Lett. **24**, 332 (1976)].
5. M. S. Pindzola and H. P. Kelly, Phys. Rev. A **14**, 204 (1976).
6. G. Wendin and K. Nuroh, Phys. Rev. Lett. **39**, 48 (1977).
7. V. M. Buĭmistrov and L. I. Trakhtenberg, Zh. Éksp. Teor. Fiz. **69**, 108 (1975) [Sov. Phys. JETP **42**, 54 (1975)].
8. V. M. Buĭmistrov and L. I. Trakhtenberg, Zh. Éksp. Teor. Fiz. **73**, 850 (1977) [Sov. Phys. JETP **46**, 447 (1977)].
9. A. Duboius and A. Maquet, Phys. Rev. A **40**, 4288 (1989).
10. A. Duboius, A. Maquet, and S. Jetzke, Phys. Rev. A **34**, 1888 (1986).
11. A. V. Korol', O. I. Obolenskiĭ, and A. V. Solov'ev, Zh. Tekh. Fiz. **69** (10), 7 (1999) [Tech. Phys. **44**, 1135 (1999)].
12. M. Ya. Amus'ya, M. Yu. Kuchiev, and A. V. Solov'ev, Pis'ma Zh. Tekh. Fiz. **10**, 1025 (1984) [Sov. Tech. Phys. Lett. **10**, 431 (1984)].
13. M. Ya. Amus'ya, M. Yu. Kuchiev, and A. V. Solov'ev, Pis'ma Zh. Tekh. Fiz. **11**, 1401 (1985) [Sov. Tech. Phys. Lett. **11**, 577 (1985)].
14. M. Ya. Amus'ya, M. Yu. Kuchiev, and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **89**, 1512 (1985) [Sov. Phys. JETP **62**, 876 (1985)].
15. K. Ishii and S. Morita, Phys. Rev. A **30**, 2278 (1984).
16. K. Ishii and S. Morita, Phys. Rev. A **31**, 1168 (1985).
17. K. Ishii, Nucl. Instrum. Methods Phys. Res. B **99**, 165 (1995).
18. M. C. Pacher and J. E. Miraglia, Phys. Rev. A **41**, 2574 (1990).
19. M. Ya. Amus'ya, A. V. Korol', M. Yu. Kuchiev, and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **88**, 383 (1985) [Sov. Phys. JETP **61**, 224 (1985)].
20. M. Ya. Amus'ya, A. V. Korol', and A. V. Solov'ev, Pis'ma Zh. Tekh. Fiz. **12**, 705 (1986) [Sov. Tech. Phys. Lett. **12**, 290 (1986)].
21. M. Ya. Amusia, Phys. Rep. **142**, 269 (1988).
22. M. Ya. Amusia and R. H. Pratt, Comments At. Mol. Phys. **28**, 247 (1992).
23. A. V. Korol and A. V. Solov'yov, J. Phys. B **30**, 1105 (1997).
24. M. C. Pacher and J. E. Miraglia, Phys. Rev. A **39**, 2905 (1989).
25. M. Ya. Amusia, L. V. Chernysheva, and A. V. Korol, J. Phys. B **23**, 2899 (1990).
26. M. Ya. Amusia and A. V. Korol, J. Phys. B **25**, 2383 (1992).
27. A. V. Korol, A. G. Lyalin, and A. V. Solov'yov, J. Phys. B **28**, 4947 (1995).
28. A. V. Korol', A. G. Lyalin, O. I. Obolenskiĭ, and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **114**, 458 (1998) [JETP **87**, 251 (1998)].
29. A. V. Korol', A. G. Lyalin, and A. V. Solov'ev, Opt. Spektrosk. **86**, 552 (1999) [Opt. Spectrosc. **86**, 486 (1999)].
30. A. V. Korol, A. G. Lyalin, R. H. Pratt, and A. V. Solov'yov, in *Proceedings of the 12th International Conference on Vacuum Ultraviolet Radiation Physics, San Francisco, 1998*, p. We116.
31. A. D. Gonzalez, M. C. Pacher, and J. E. Miraglia, Phys. Rev. A **37**, 4974 (1988).
32. A. V. Solov'yov, in *Proceedings of the 5th EPS Conference on Atomic and Molecular Physics, Edinburgh, 1995*, p. 711.
33. L. G. Gerchikov, A. V. Korol, S. A. Sheinerman, and A. V. Solov'yov, in *Proceedings of the 17th International Conference on X-ray and Inner-Shell Processes, Hamburg, Germany, 1996*, p. 167.
34. V. A. Astapenko, V. M. Buĭmistrov, Yu. A. Krotov, *et al.*, Zh. Éksp. Teor. Fiz. **88**, 1560 (1985) [Sov. Phys. JETP **61**, 930 (1985)].
35. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Zh. Éksp. Teor. Fiz. **117**, 496 (2000) [JETP **90**, 434 (2000)].
36. V. A. Astapenko, L. A. Bureeva, and V. S. Lisitsa, Zh. Éksp. Teor. Fiz. **117**, 906 (2000) [JETP **90**, 788 (2000)].
37. A. V. Korol, O. I. Obolensky, A. V. Solov'yov, and I. A. Solov'yov, J. Phys. B **34**, 1589 (2001).
38. V. Ambrose, C. A. Quarles, and R. Ambrose, Nucl. Instrum. Methods Phys. Res. B **124**, 457 (1997).
39. S. Portillo, C. A. Quarles, F. D. McDaniel, *et al.*, in *Proceedings of the 21st International Conference on the Physics of Electronic and Atomic Collisions, Sendai, Japan, 1999*, p. 574.
40. C. A. Quarles and S. Portillo, in *Applications of Accelerators in Research and Industry*, Ed. by J. L. Duggan and I. L. Morgan (American Inst. of Physics Press, New York, 1999), p. 174.
41. T. Ludziejewski, Th. Stohlker, S. Keller, *et al.*, J. Phys. B **31**, 2601 (1998).
42. I. Lindgren and J. Morrison, *Atomic Many-Body Theory* (Springer-Verlag, Berlin, 1986).
43. L. V. Chernysheva and V. L. Yakhontov, Comput. Phys. Commun. **119**, 232 (1999).
44. D. A. Varshalovich, A. N. Moskalev, and V. K. Khersonskii, *Quantum Theory of Angular Momentum* (Nauka, Leningrad, 1975; World Scientific, Singapore, 1988).
45. A. R. Sohval, J. P. Delvaille, K. Kalata, and H. W. Schnopper, J. Phys. B **8**, L426 (1975).
46. K. Ozawa, J. H. Chang, Y. Yamamoto, *et al.*, Phys. Rev. A **33**, 3018 (1986).
47. A. V. Solov'yov, Z. Phys. D **24**, 5 (1992).
48. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer-Verlag, Berlin, 1985).
49. M. Ya. Amus'ya, M. Yu. Kuchiev, and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **94** (1), 74 (1988) [Sov. Phys. JETP **67**, 41 (1988)].
50. M. Ya. Amus'ya and A. V. Solov'ev, Zh. Éksp. Teor. Fiz. **97**, 745 (1990) [Sov. Phys. JETP **70**, 416 (1990)].
51. P. M. Bergstrom, Jr., T. Suric, K. Pisk, and R. H. Pratt, Phys. Rev. A **48**, 1134 (1993).
52. S. A. Zapryagaev, N. L. Manakov, and V. G. Pal'chikov, *Theory of Multiply Charged Ions with One and Two Electrons* (Énergoatomizdat, Moscow, 1985).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 720–727.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 838–846.
Original Russian Text Copyright © 2002 by Parkhomenko.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Noninversive Partial Velocity Amplification of Radiation by Ions due to Their Rotation in a Magnetic Field

A. I. Parkhomenko

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

Received October 29, 2001

Abstract—It is shown that upon the application of an external magnetic field, a gas of ionized particles may experience noninversive partial velocity amplification of radiation by ions due to their Larmor rotation. In this case, virtually all ions may be in the ground state. It may happen that approximately half the number of ions in the medium amplify the incident radiation. The integrated absorption coefficient remains positive due to the enhancement of absorption of radiation by the other half of ions. Noninversive amplification of radiation takes place when the condition $\omega_c \gtrsim \Gamma^2/kv_T$ is satisfied (ω_c is the cyclotron frequency of ions in the magnetic field; Γ is the homogeneous half-width of the absorption line for ions, and kv_T is the Doppler width). In the case of interaction of atomic ions with radiation in the optical range, this corresponds to magnetic fields $B \gtrsim 600$ G (for the ion mass $M \sim 10$ amu). Noninversive partial velocity amplification of radiation is a "latent" effect in the sense that it disappears upon averaging over all velocity directions of ions. This effect is associated with the emergence of phase incursion of the induced dipole moment oscillations for ions moving in circular cyclotron orbits, which depends on the ion velocity. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

It is well known that the Larmor rotation of ions in a magnetic field may lead to a sharp change in their absorption line profile [1–3]. Namely, for observation across the magnetic field and for $\omega_c \gtrsim \Gamma$ (ω_c is the cyclotron frequency of ions and Γ is the homogeneous half-width of the absorption line for ions), the Doppler profile of the ion line splits into a number of equidistant peaks. The width of each peak is equal to the homogeneous absorption line width 2Γ, and the distance between adjacent peaks is equal to the cyclotron frequency ω_c of ions. If the average Larmor radius of an ion orbit is much smaller than the wavelength, the absorption line has a Lorentz profile with a homogeneous width 2Γ which may be much smaller than the Doppler width [1]. Such a narrowing of the line is due to the limitation of the spatial displacement of ions across the magnetic field (which eliminates the nonhomogeneous Doppler broadening) and is similar to the well-known collision-induced narrowing of spectral lines due to the Dicke effect [4–6].

It is natural to assume that the Larmor rotation of ions in a magnetic field also leads to a sharp change in the radiation absorption by ions with a fixed value of their velocity **v**. The corresponding theoretical calculations lead to a completely unexpected result. It turns out that groups of ions with definite directions of velocities in a magnetic field amplify the radiation incident on the medium even when all ions are in the ground state.

The present work is devoted to a theoretical analysis of this phenomenon.

2. BASIC RELATIONS

Let us consider a gas of ionized particles in a constant uniform magnetic field **B**. Let radiation in the form of a running monochromatic wave be resonantly absorbed during the *m*–*n* transition between the ground (*n*) and the first excited (*m*) levels of ions. We will confine our analysis to the simple case when the Zeeman splitting of the absorption line can be disregarded. For example, line splitting is absent in the normal Zeeman effect (the Landé *g* factors of the combining states *m* and *n* are identical) when radiation linearly polarized along the magnetic field **B** propagates at right angles to this field.

The radiation absorption probability $P(\mathbf{v})$ at the *m*–*n* transition ($P(\mathbf{v})$ is equal to the number of radiation absorption acts per unit time for a particle with a given velocity **v** in a unit interval of velocities) is defined by the nondiagonal element $\rho_{mn}(\mathbf{v})$ of the density matrix [6]:

$$P(\mathbf{v}) = -\frac{2}{N}\mathrm{Re}[iG^*\rho_{mn}(\mathbf{v})],$$
$$|G|^2 = \frac{B_{nm}I}{2\pi}, \quad B_{nm} = \frac{\lambda^2\Gamma_m}{4\hbar\omega}, \tag{1}$$

1063-7761/02/9404-0720$22.00 © 2002 MAIK "Nauka/Interperiodica"

where N is the concentration of ions, B_{nm} is the second Einstein coefficient for the m–n transition, I is the radiation intensity, ω and λ are the frequency and the wavelength of radiation, and Γ_m is the spontaneous decay rate for the excited state m. For a low radiation intensity under steady-state and spatially homogeneous conditions, $\rho_{mn}(\mathbf{v})$ is determined from the equation [6]

$$\left[\frac{\Gamma_m}{2} - i(\Omega - \mathbf{k}\cdot\mathbf{v}) + \omega_c[\mathbf{v}\times\mathbf{h}]\frac{\partial}{\partial\mathbf{v}}\right]\rho_{mn}(\mathbf{v}) = S_{mn}(\mathbf{v}) + iGNW(\mathbf{v}), \tag{2}$$

where

$$\omega_c = \frac{eB}{Mc}, \quad \Omega = \omega - \omega_{mn}, \tag{3}$$

$\mathbf{h} = \mathbf{B}/B$ is the unit vector in the direction of the magnetic field, $\mathbf{k}$ is the wave vector of radiation, $W(\mathbf{v})$ is the Maxwell velocity distribution, $S_{mn}(\mathbf{v})$ is the "nondiagonal" collision integral, ω_c is the cyclotron frequency of ions, e is the elementary electric charge, M is the ion mass, and ω_{mn} is the frequency of the m–n transition.

For the nondiagonal collision integral $S_{mn}(\mathbf{v})$ appearing in formula (2), we will use the approximation conventionally employed in nonlinear spectroscopy [6],

$$S_{mn}(\mathbf{v}) = -\left(\Gamma - \frac{\Gamma_m}{2}\right)\rho_{mn}(\mathbf{v}), \tag{4}$$

and indicating that collision completely breaks the phase of the oscillating dipole moment.

We will solve Eq. (2) in a system of coordinates in which the z axis is directed along the magnetic field $\mathbf{B}$ and the x axis is directed along the wave vector $\mathbf{k}$ (we assume that $\mathbf{k} \perp \mathbf{B}$). In the velocity space, it is convenient to go over to the cylindrical system of coordinates $v_\perp$, φ, v_z ($v_x = v_\perp\cos\varphi$, $v_y = v_\perp\sin\varphi$). In these coordinates, Eq. (2) taking into account Eq. (4) assumes the form

$$\left[\Gamma - i(\Omega - k_\perp v\cos\varphi) - \omega_c\frac{\partial}{\partial\varphi}\right]\rho_{mn}(\mathbf{v}) = iGNW(\mathbf{v}). \tag{5}$$

Solving this linear nonhomogeneous differential equation, we obtain the following expression for the radiation absorption probability $P(\mathbf{v})$ (1):

$$P(\mathbf{v}) = 2|G|^2 W(v_\perp)W(v_z) \times \mathrm{Re}\left[\exp(i\beta\sin\varphi)\sum_{n=-\infty}^{n=\infty}\frac{\exp(-in\varphi)J_n(\beta)}{\Gamma - i(\Omega - n\omega_c)}\right], \tag{6}$$

where

$$W(v_\perp) = \frac{1}{(\sqrt{\pi}v_T)^2}\exp\left(-\frac{v_\perp^2}{v_T^2}\right), \quad W(v_z) = \frac{1}{\sqrt{\pi}v_T}\exp\left(-\frac{v_z^2}{v_T^2}\right), \quad \beta = \frac{kv_\perp}{\omega_c}, \quad v_T = \sqrt{\frac{2k_BT}{M}}, \tag{7}$$

$J_n(\beta)$ is a Bessel function of the first kind, $W(v_\perp)$ and $W(v_z)$ are the Maxwell distributions over the transverse and longitudinal (relative to the magnetic field $\mathbf{B}$) components of velocity $\mathbf{v}$, v_T is the most probable velocity of ions, k_B is the Boltzmann constant, and T is the temperature. In zero magnetic field (for $\omega_c = 0$), the radiation absorption probability $P(\mathbf{v})$ is defined by the well-known formula [6]

$$P(\mathbf{v}) = \frac{2|G|^2\Gamma W(\mathbf{v})}{\Gamma^2 + (\Omega - \mathbf{k}\cdot\mathbf{v})^2}. \tag{8}$$

3. ABSORPTION LINE PROFILE

The absorption line profile is defined by the radiation absorption probability integrated over velocities:

$$P \equiv \int P(\mathbf{v})d\mathbf{v}.$$

In accordance with the well-known results [1–3], we obtain the following expression for P from Eq. (6):

$$P = 2|G|^2\Gamma\exp(-\mu)\sum_{n=-\infty}^{n=\infty}\frac{I_n(\mu)}{\Gamma^2 + (\Omega - n\omega_c)^2}, \quad \mu = \frac{(kv_T)^2}{2\omega_c^2}, \tag{9}$$

where $I_n(\mu)$ is a modified Bessel function [7]. It follows from Eq. (9) that the Larmor rotation of ions in a magnetic field may lead to the emergence of equidistant peaks (cyclotron resonances) in the absorption line profile. The distance between adjacent peaks is equal to the cyclotron frequency ω_c of ions, and the width of an individual peak is determined by the homogeneous width 2Γ of the absorption line.

In the case of Doppler absorption line broadening and a moderate magnetic field ($kv_T \gg \Gamma, \omega_c$), we can derive from Eq. (9) the following formula, which is valid in the radiation frequency detuning range $|\Omega| \lesssim kv_T$ [1]:

$$P = P_0\exp\left[-\left(\frac{\Omega}{kv_T}\right)^2\right] \times \frac{\sinh(2\pi\Gamma/\omega_c)}{\cosh(2\pi\Gamma/\omega_c) - \cos(2\pi\Omega/\omega_c)}, \tag{10}$$

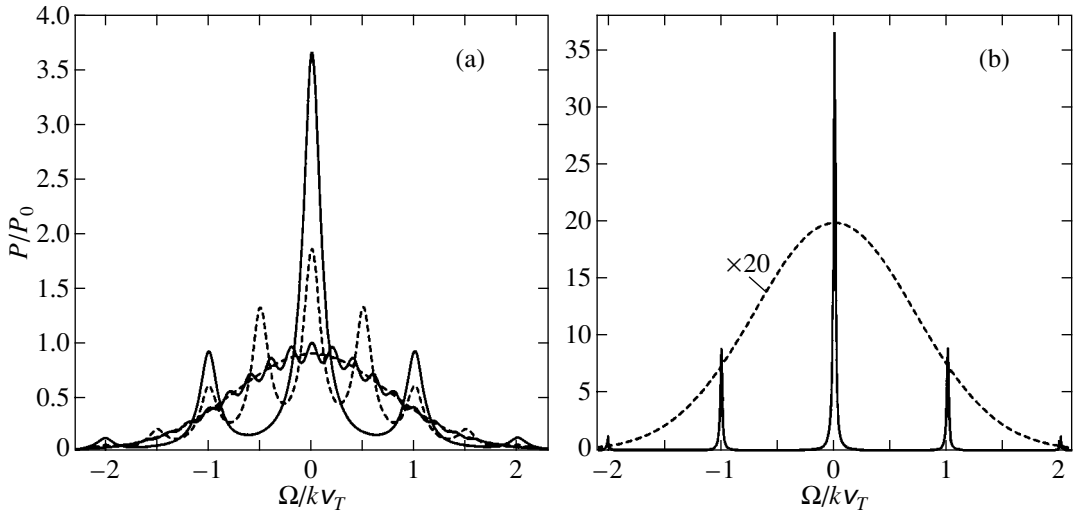

Fig. 1. Radiation absorption probability integrated over velocities as a function of the radiation frequency detuning for various values of magnetic field: (a): $\Gamma/kv_T = 0.1$, $\omega_c/\Gamma = 10, 5, 2, 0$; the values of parameter ω_c/Γ correspond to the arrangement of the curves in descending order near $\Omega = 0$; (b) $\Gamma/kv_T = 10^{-2}$; the solid curve corresponds to $\omega_c/\Gamma = 100$ and the dashed curve to $\omega_c = 0$.

where

$$P_0 = \frac{2\sqrt{\pi}|G|^2}{kv_T} \tag{11}$$

is the radiation absorption probability at the line center under Doppler broadening in zero magnetic field. The absorption line described by formula (10) has the form of a Doppler profile modulated by the periodic function of the radiation frequency detuning Ω with a period equal to ω_c. The oscillating function P (10) has peaks for $\Omega = n\omega_c$ and minima for $\Omega = (n + 1/2)\omega_c$. For $\omega_c \gg \Gamma$, the Doppler profile distinctly splits into a series of peaks, while, for $\omega_c \lesssim \Gamma$, the shape of the line differs from the Doppler contour by an exponentially small oscillating correction.

In the case of homogeneous absorption line broadening ($\Gamma \gg kv_T$) or strong magnetic fields (for $\omega_c \gg kv_T$ and for an arbitrary relation between Γ and kv_T), the line has a Lorentz profile of width 2Γ [1]:

$$P = \frac{2|G|^2\Gamma}{\Gamma^2 + \Omega^2}. \tag{12}$$

Figure 1 shows the dependence of the velocity absorption probability P integrated over velocities and calculated by formula (9) as a function of the radiation frequency detuning Ω for various values of the magnetic field.

4. DEPENDENCE OF RADIATION ABSORPTION PROBABILITY ON ION VELOCITY

The dependence of the radiation absorption probability $P(\mathbf{v})$ on the velocity component v_z in expression (6) is trivial and is manifested only through the Maxwell factor $W(v_z)$ (the magnetic field does not affect the motion of particles along the z axis). For this reason, we will be interested in the integrated characteristics $P(\mathbf{v}_\perp)$ and $P(\varphi)$:

$$P(\mathbf{v}_\perp) \equiv P(v_\perp, \varphi) = \int_{-\infty}^{\infty} P(\mathbf{v})dv_z,$$
$$P(\varphi) = \int_0^{\infty} P(v_\perp, \varphi)v_\perp dv_\perp. \tag{13}$$

Here, $P(\varphi)$ is the number of radiation absorption acts per unit time in a unit interval of angles per ion with a given value of φ between the direction of radiation $\mathbf{k}$ and the projection of the velocity of ions onto the plane perpendicular to the magnetic field. The function $P(\varphi)$ possesses the property

$$P(\Omega, \varphi) = P(-\Omega, \varphi \pm \pi). \tag{14}$$

In the case of strong magnetic fields ($\omega_c \gg kv_T$) or homogeneous broadening of the absorption line ($\Gamma \gg kv_T$), formula (6) for $P(\mathbf{v})$ is simplified considerably. In these cases, we have

$$P(\mathbf{v}) = PW(\mathbf{v})\left[1 + \frac{\Omega k v_\perp}{\Gamma}\tau(\varphi)\right], \tag{15}$$

$$P(v_\perp, \varphi) = PW(v_\perp)\left[1 + \frac{\Omega k v_\perp}{\Gamma}\tau(\varphi)\right], \tag{16}$$

$$P(\varphi) = \frac{P}{2\pi}\left[1 + \frac{\sqrt{\pi}}{2}\frac{\Omega k v_T}{\Gamma}\tau(\varphi)\right], \tag{17}$$

where

$$\tau(\varphi) = \frac{2\Gamma(\Gamma^2+\Omega^2)\cos\varphi - \omega_c(3\Gamma^2+\omega_c^2-\Omega^2)\sin\varphi}{[\Gamma^2+(\Omega-\omega_c)^2][\Gamma^2+(\Omega+\omega_c)^2]}. \quad (18)$$

In formulas (15)–(17), the radiation absorption probability P integrated over velocities is defined by formula (12). It follows from formulas (15)–(17) that the absorption probabilities $P(\varphi)$, $P(v_\perp,\varphi)$, and $P(\mathbf{v})$ may assume negative values. This is manifested most strongly for high magnetic fields for which $\omega_c \gg \Gamma$. In this case, the dependence of the factor $\tau(\varphi)$ on the radiation frequency detuning Ω has a resonant nature and the factor $\tau(\varphi)$) attains its maximal value equal to $\cos\varphi/2\Gamma$ for $|\Omega| = \omega_c$. In this case, the alternating (second) term in square brackets in formulas (15)–(17) for the radiation absorption probability is equal to $\Omega k v_T/2\Gamma^2$ in order of magnitude and may be considerably greater than unity in magnitude, having the negative sign. Thus, noninversive partial velocity amplification of radiation by ions due to their rotation in the magnetic field may take place. The intensity of radiation incident on the medium may be indefinitely low in this case; for this reason, almost all the particles may be in the ground state.

It should be noted that the sign-alternating term in square brackets in formulas (15)–(17), which is responsible for the emergence of noninversive partial velocity amplification of radiation, makes zero contribution to the absorption probability P integrated over velocities that determines the absorption line profile.

Figure 2 shows the $P(\varphi)$ dependence calculated by formula (17) for different values of frequency detuning near the cyclotron resonance ($\Omega = \omega_c$) in the case of Doppler broadening of the absorption line. The absorption probability $P(\varphi)$ assumes negative values in the interval of angles $\Delta\varphi \approx \pi$ in the vicinity of $\varphi \approx \pi$ (radiation is amplified by particles moving predominantly opposite to the direction of radiation). In the case shown in Fig. 2, the amplitude of positive and negative values of $P(\varphi)$ are identical and equal approximately to the radiation absorption probability P_0 at the line center in zero magnetic field in spite of large values of radiation frequency detuning ($\Omega \approx 5kv_T$). At the same time, the radiation absorption probability P integrated over velocities is low:

$$P/P_0 \approx \Gamma/25\sqrt{\pi}kv_T \approx 2.3\times10^{-4}.$$

In the case presented in Fig. 2, the factor $\Omega k v_T/2\Gamma^2 \approx 2\times10^4$; consequently, the radiation absorption probability $P(\varphi)$ increases upon the application of the field by more than four orders of magnitude. Thus, the following dramatic situation takes place in the medium. Approximately half the particles moving in a certain direction in the medium strongly absorb radiation, while the other half of the particles moving in the opposite direction strongly amplify radiation. However, the contributions from these two groups of particles to the integrated absorption probability are compensated almost completely, and the medium as a whole weakly absorbs radiation.

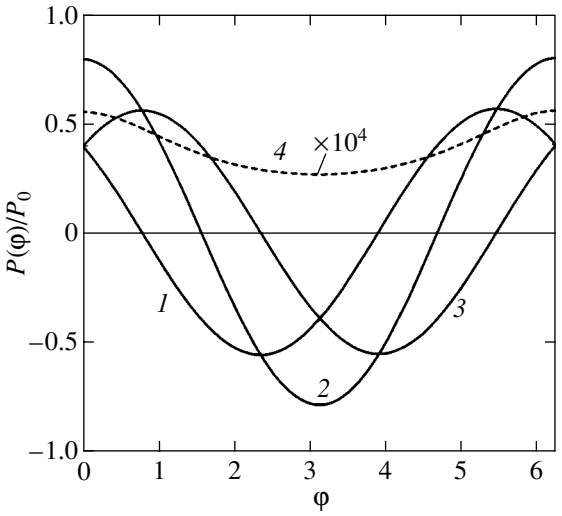

Fig. 2. Radiation absorption probability $P(\varphi)$ integrated over velocities $v_\perp$ and v_z as a function of the azimuthal angle φ for various values of radiation frequency detuning; $\Gamma/kv_T = 10^{-2}$; solid curves correspond to $\omega_c/kv_T = 5$; the dashed curve corresponds to $\omega_c = 0$ (calculations are based on formula (8)); $\omega_c/kv_T = 4.99$ (*1*), 5 (*2*, *4*), and 5.01 (*3*).

Formulas (15)–(17) do not describe the most interesting case of moderate magnetic fields with $\omega_c \ll kv_T$ for the Doppler line broadening ($kv_T \gg \Gamma$), and the radiation absorption probability in this case must be calculated by the exact formula (6). The corresponding dependence $P(\varphi)$ calculated by formula (6) is shown in Fig. 3. The values of $P(\varphi)$ were calculated only for positive values of radiation frequency detuning Ω in view of the fact that the function $P(\varphi)$ possesses property (14).

Figures 3a–3c illustrate the emergence of narrow resonances (in angle φ) of noninversive amplification of radiation for $\omega_c \ll kv_T$. In the cases presented in Figs. 3a and 3b, the absorption probability $P(\varphi)$ assumes negative values in small angular intervals $\Delta\varphi \approx 0.02$ and $\Delta\varphi \approx 1.6\times10^{-3}$, respectively. These cases are also interesting in that the absorption line profile for $\omega_c \lesssim \Gamma \ll kv_T$ differs from the Doppler profile by a small oscillating correction; nevertheless, the oscillations of $P(\varphi)$ are considerable. An analysis of expression (9) for P shows that the relative deviation ξ of the line profile from the Doppler profile for $\omega_c \lesssim \Gamma \ll kv_T$ can be estimated by the formula

$$\xi \approx \max\left[2\exp\left(-\frac{2\pi\Gamma}{\omega_c}\right), \frac{\omega_c^2}{5(kv_T)^2}\right]. \quad (19)$$

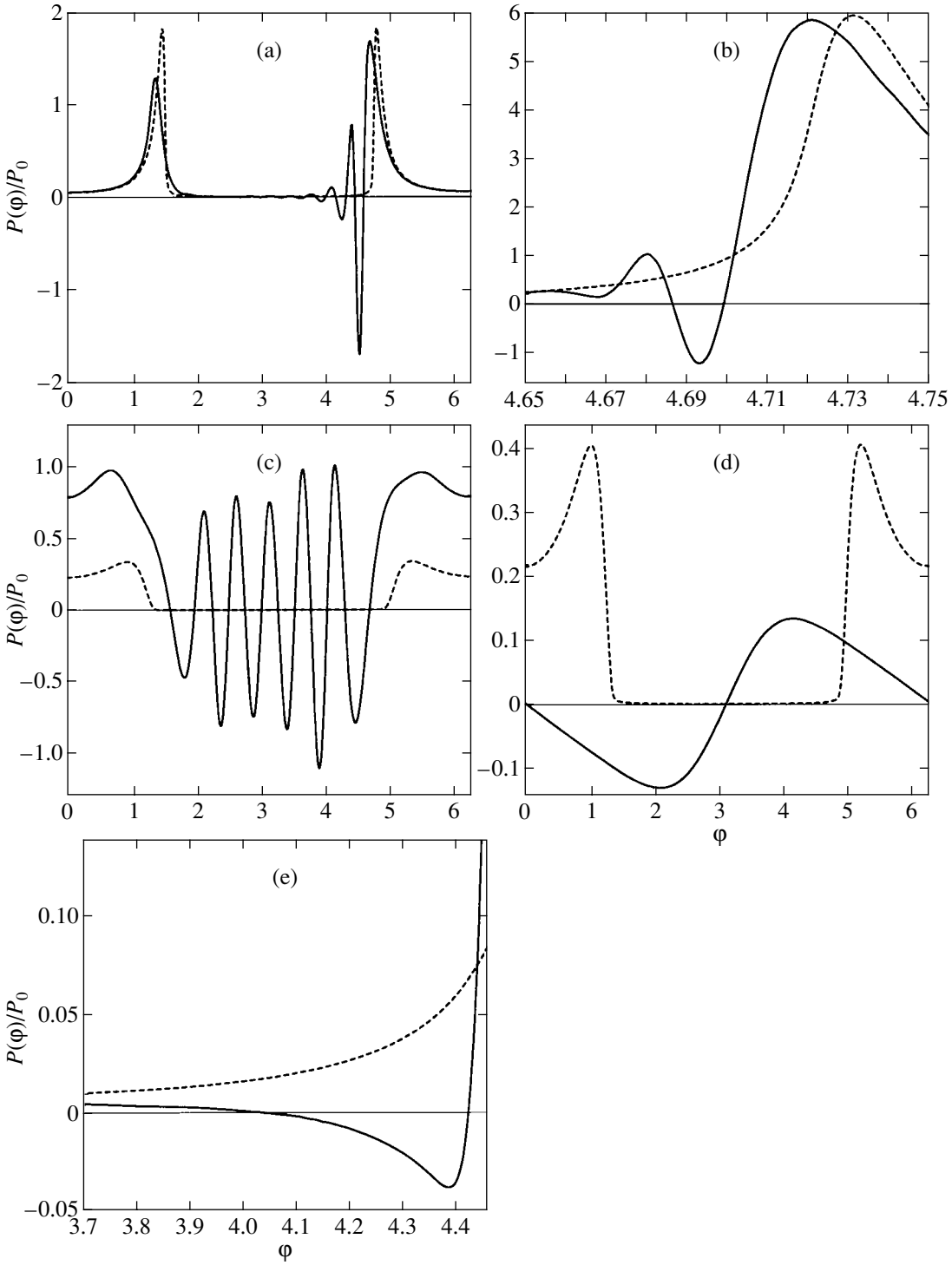

Fig. 3. Dependence $P(\varphi)$ for various values of the radiation frequency detuning and magnetic field; $\Gamma/kv_T = 10^{-2}$; the dashed curve corresponds to $\omega_c = 0$; (a) $\omega_c/\Gamma = 1$, $\Omega/kv_T = 0.1$; (b) $\omega_c/\Gamma = 0.02$, $\Omega/kv_T = 0.02$; (c) $\omega_c/\Gamma = 10$, $\Omega/kv_T = 0.6$; (d) $\omega_c/\Gamma = 100$, $\Omega/kv_T = 0.5$; (e) $\omega_c/\Gamma = 1$, $\Omega = 0$.

For the cases depicted in Figs. 3a and 3b, the values of ξ are equal to 4×10^{-3} and 10^{-8}, respectively; i.e., the line profile virtually coincides with the Doppler profile. Nevertheless, the amplitude of the negative values of $P(\varphi)$ is large and exceeds the radiation absorption probability P_0 at the center of the Doppler-broadened line in zero magnetic field.

A numerical analysis shows that, in the case of Doppler broadening ($\Gamma/kv_T \ll 1$), the effect of noninversive

partial velocity amplification of radiation emerges when the condition

$$\omega_c \gtrsim \frac{\Gamma^2}{kv_T} \tag{20}$$

is satisfied. In other words, the effect may appear in weak magnetic fields for which $\omega_c \ll \Gamma$. Figure 3b illustrates the emergence of the effect for

$$\omega_c = 2\Gamma(\Gamma/kv_T) = 0.02\Gamma.$$

Let us estimate the magnitude of the magnetic field required for the emergence of the effect. For the ion mass $M \sim 10$ amu, temperature $T \sim 10^3$ K, and radiation wavelength $\lambda \sim 0.5$ μm, the Doppler line width is $kv_T \approx 1.6 \times 10^{10}$ s^{-1}. Assuming that the homogeneous line width is determined by radiative decay, we set $\Gamma \sim 10^8$ s^{-1}. Then, we obtain the following estimate from relation (20): $B \gtrsim 600$ G.

The number of oscillations $N_{\rm osc}$ of function $P(\varphi)$ depends on the magnitude of the magnetic field and on the radiation frequency detuning. For $|\Omega| \lesssim 2kv_T$ and $\omega_c > \Gamma$, the number of oscillations can be estimated by the formula

$$N_{\rm osc} \sim \frac{|\Omega|}{\omega_c} \tag{21}$$

(see Fig. 3c).

For $\omega_c \gg \Gamma$, the Doppler absorption line profile distinctly splits into a number of peaks (see Fig. 1b). For radiation frequencies tuned to the interval between the central ($\Omega = 0$) and the first side ($|\Omega| = \omega_c$) peaks and for $\omega_c \gtrsim kv_T$, the dependence $P(\varphi)$ becomes sinusoidal (see Fig. 3d).

In the cases when $\Gamma \gg kv_T$ or $\omega_c \gg kv_T$, the radiation absorption probabilities $P(\varphi)$, $P(v_\perp, \varphi)$, and $P(\mathbf{v})$ may assume negative values, in accordance with formulas (15)–(17), only for a nonzero radiation detuning frequency $\Omega \neq 0$. However, in the most interesting case when $\omega_c, \Gamma \ll kv_T$, these absorption probabilities may assume negative values for $\Omega = 0$ also (see Figs. 3e and 4).

5. QUALITATIVE PATTERN

The effect of noninversive partial velocity amplification of radiation considered by us here is associated with the emergence of the ion-velocity-dependent phase incursion in the oscillations of the induced dipole moment of ions due to their motion in a circular cyclotron orbit.

Indeed, the concepts considered above can be used to obtain a qualitative estimate for the absorption probability $P(\mathbf{v}_\perp)$ which matches formula (16) for $P(\mathbf{v}_\perp)$ provided that $\omega_c \gg \Gamma, |\Omega|$. In order to obtain this estimate, we decompose the total velocity $\mathbf{v}$ of ions into the velocity $\mathbf{v}_z$ directed along the magnetic field $\mathbf{B}$ and velocity $\mathbf{v}_\perp$ perpendicular to $\mathbf{B}$:

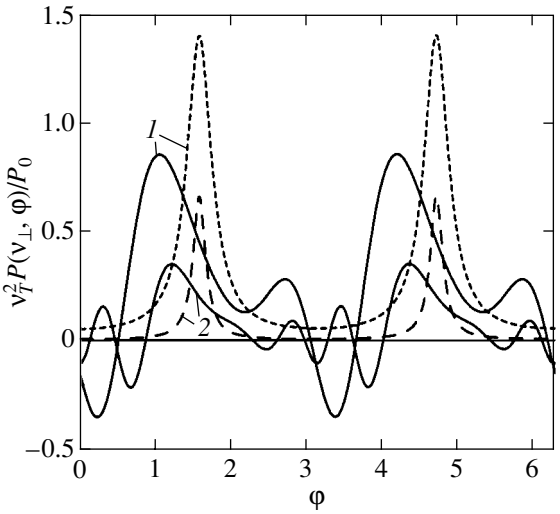

Fig. 4. Radiation absorption probability $P(v_\perp, \varphi)$ as a function of the azimuthal angle φ for various values of ion velocity v; $\Gamma/kv_T = 0.1$, $\Omega = 0$; solid curves correspond to $\omega_c/\Gamma = 1$; dashed curves correspond to $\omega_c = 0$; $v_\perp/v_T = 0.5$ (*1*) and 1 (*2*).

$$\mathbf{v} = \mathbf{v}_z + \mathbf{v}_\perp.$$

Since the magnetic field does not affect the motion of ions along $\mathbf{B}$, the effect of noninversive partial velocity amplification of radiation can be due only to singularities in the dependence of the absorption probability on velocity $\mathbf{v}_\perp$. Vector $\mathbf{v}_\perp$ rotates uniformly about the magnetic field direction with the angular velocity ω_c. In a strong magnetic field (for $\omega_c \gg \Gamma$), the particle has time to describe many turns between collisions (which cause phase incursion of the dipole moment induced by radiation). The Larmor rotation of ions does not lead to any phase incursion of the induced magnetic moment, but leads to a phase incursion for an ion moving in a circular cyclotron orbit. We assume that the magnetic field $\mathbf{B}$ is directed along the z axis, while the wave vector $\mathbf{k}$ of radiation is directed along the x axis. The position of an ion on a circular cyclotron orbit will be defined by the angle φ between the direction $\mathbf{k}$ of radiation and velocity $\mathbf{v}_\perp$. The phase incursion for the ion moving in a circular cyclotron orbit from point $A(x_0, \varphi_0)$ to point $B(x, \varphi)$ (see Fig. 5) is

$$\Delta\psi = k(x - x_0) = -k\rho_c(\sin\varphi - \sin\varphi_0), \tag{22}$$

where $\rho_c = v_\perp/\omega_c$ is the Larmor radius of the ion orbit. The angle φ specifying the direction of velocity $\mathbf{v}_\perp$ is fixed, while angle φ_0 varies from 0 to 2π in this case. The phase incursion averaged over angle φ_0 is

$$\overline{\Delta\psi} = -k\rho_c\sin\varphi. \tag{23}$$

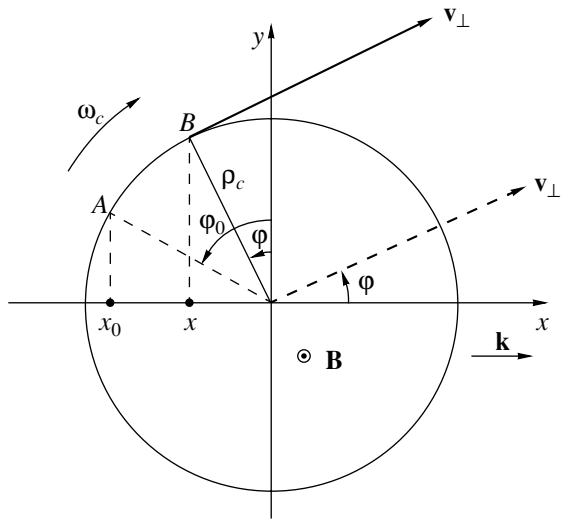

Fig. 5. Schematic diagram explaining formula (22) for the phase incursion for an ion moving in a circular cyclotron orbit.

In view of this phase incursion, the phase shift $\psi(\mathbf{v}_\perp)$ of oscillations (relative to the radiation field oscillations) of the macroscopic polarization created by ions moving at velocity $\mathbf{v}_\perp$ can be presented as the sum

$$\psi(\mathbf{v}_\perp) = \psi_0 + \overline{\Delta\psi} = \psi_0 - \frac{kv_\perp \sin\varphi}{\omega_c}, \qquad (24)$$

where the term ψ_0 is independent of velocity.

Let us now take into account the fact that the energy exchange between the field and the ensemble of particles moving with velocity $\mathbf{v}_\perp$ we are interested in is determined by the polarization component which is not synphase with the field and is proportional to $\sin\psi(\mathbf{v}_\perp)$ (see, for example, [8]). Consequently, the radiation absorption probability $P(\mathbf{v}_\perp)$ for ions with the fixed velocity $\mathbf{v}_\perp$ in a strong magnetic field (such that $kv_\perp/\omega_c \ll 1$) is given by

$$P(\mathbf{v}_\perp) \propto \sin\psi(\mathbf{v}_\perp) \approx \left(1 - \frac{kv_\perp}{\omega_c}\frac{\sin\varphi}{\tan\psi_0}\right) \times \sin\psi_0 \propto \left(1 - \frac{kv_\perp}{\omega_c}\frac{\sin\varphi}{\tan\psi_0}\right)P. \qquad (25)$$

While deriving the last relation in this formula, we took into account the fact that the radiation absorption probability P integrated over velocities is obviously proportional to $\sin\psi_0$.

In order to determine $\tan\psi_0$, we will use a visual classical model describing electrons in an atom as damped harmonic oscillators performing forced vibrations in the electric field of the wave. It is known from classical mechanics that the phase shift ψ_0 between the oscillator vibrations and the external driving force is given by the formula $\tan\psi_0 = \gamma/\Omega$, where γ is the damping coefficient (resonance curve half-width) and $\Omega = \omega - \omega_0$ is the detuning of the driving force frequency ω relative to frequency ω_0 of natural vibrations of the oscillator (see, for example, [9]). In the case of ions, we have $\gamma = \Gamma$, and the quantity Ω has the meaning of radiation frequency detuning relative to the transition frequency for a stationary particle and may have any sign. As a result, we obtain the following estimate for ions in a magnetic field from relation (25):

$$P(\mathbf{v}_\perp) \propto \left(1 - \frac{\Omega k v_\perp}{\Gamma\omega_c}\sin\varphi\right)P. \qquad (26)$$

The second sign-alternating term on the right-hand side of this relation appears due to the rotation of ions in the magnetic field. Its magnitude may be greater than unity, and, hence $P(\mathbf{v}_\perp)$ may be smaller than zero; i.e., ions in an external magnetic field may induce noninversive partial velocity amplification of radiation.

It should be noted that, for $\omega_c \gg \Gamma, |\Omega|$, relation (18) leads to

$$\tau(\varphi) = -\sin\varphi/\omega_c$$

and estimate (26) agrees with formula (16).

6. CONCLUSIONS

The main conclusion of the present work is that the Larmor rotation of ions in a magnetic field may lead to the emergence of noninversive partial velocity amplification of radiation by ions. The effect is associated with the emergence of an ion-velocity-dependent phase incursion in the oscillations of the induced dipole moment of ions due to their motion in a circular cyclotron orbit.

We have considered the case of absorption of low-intensity radiation during the transition between the ground and the first excited energy levels of ions. The formulas derived for the radiation absorption probability can be easily generalized to the case of transitions between excited levels of ions. For this purpose, the following substitution must be made in the initial equation (2) for $\rho_{mn}(\mathbf{v})$:

$$N \longrightarrow \rho_n^0 - \rho_m^0,$$

where ρ_n^0 and ρ_m^0 are the populations of the lower and upper levels in the absence of radiation, respectively (the effect of low-intensity radiation on the population of the levels can be neglected). Then, the right-hand side of formula (6) for $P(\mathbf{v})$ (as well as the formulas for $P(\mathbf{v}_\perp)$, $P(\varphi)$, and P following from Eq. (6)) will acquire an additional factor $(\rho_n^0 - \rho_m^0)/N$ which may be positive

or negative (in the case of population inversion). Since $P(\mathbf{v})$ may be smaller than zero in the case

$$\rho_n^0 - \rho_m^0 = N > 0$$

investigated by us here, $P(\mathbf{v})$ may also be greater than zero for $\rho_n^0 - \rho_m^0 < 0$. Thus, in the case of population inversion of operating energy levels (affected by radiation), partial velocity absorption of radiation by ions may take place due to their Larmor rotation in the magnetic field.

In the case of Doppler line broadening ($\Gamma/kv_T \ll 1$) the effect of noninversive partial velocity amplification of radiation by ions emerges, in accordance with relation (20), in magnetic fields $B \gtrsim B_0$, where

$$B_0 \approx \frac{10^{-4} M\Gamma^2}{kv_T} \tag{27}$$

(M is the ion mass in atomic units). For atomic ions with a mass $M \sim 10$ amu, homogeneous line half-width $\Gamma \sim 10^8$ s^{-1}, and Doppler width $kv_T \approx 1.6 \times 10^{10}$ s^{-1} (wavelength $\lambda \sim 0.5$ μm and temperature $T \sim 10^3$ K), we obtain $B_0 \approx 600$ G.

The threshold value of the magnetic field B_0 starting from which the noninversive amplification effect can be observed is the lower, the smaller the homogeneous line width. If the homogeneous width is determined by radiative decay, the minimum values of B_0 are attained for radiation absorbed at transitions between the Rydberg states of atomic ions or for radiation absorption at vibration-rotation transitions of molecular ions. In these cases, B_0 may be smaller than 1 G.

Noninversive partial velocity amplification of radiation is a "latent" effect in the sense that it vanishes as a result of averaging over all directions of velocities of ions:

$$\int_0^{2\pi} P(\mathbf{v})\, d\varphi > 0.$$

However, this does not mean that the latent effect cannot lead to new properties of spectroscopy and kinetics of ions in laser fields. In particular, we can expect that such a strong latent effect as the amplification of radiation by half the particles in a medium described in the present work will strongly affect the absorption of the test (probing and comparatively weak) field, and we cannot rule out the possibility of noninversive amplification of the test field (due to a comparatively stronger, but still low-intensity, field which does not change the population of energy levels) in a frequency range of the order of the Doppler line width. The solution of the corresponding theoretical problem is interesting as regards the experimental "manifestation" of the effect and requires separate analysis.

ACKNOWLEDGMENTS

The author is indebted to A.M. Shalagin for fruitful discussions and valuable critical remarks.

This work was supported financially by the Russian Foundation for Basic Research (project no. 01-02-17433).

REFERENCES

1. M. I. D'yakonov, Zh. Éksp. Teor. Fiz. **51**, 612 (1966) [Sov. Phys. JETP **24**, 408 (1967)].
2. G. I. Smirnov and D. A. Shapiro, Zh. Éksp. Teor. Fiz. **87**, 1639 (1984) [Sov. Phys. JETP **60**, 940 (1984)].
3. V. V. Antsiferov and G. I. Smirnov, *Physics of Coherent Radiation Processes in Plasma* (NGTU, Novosibirsk, 1995).
4. R. H. Dicke, Phys. Rev. **89**, 472 (1953).
5. S. G. Rautian and I. I. Sobel'man, Usp. Fiz. Nauk **90**, 209 (1966) [Sov. Phys. Usp. **9**, 701 (1967)].
6. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979).
7. *Handbook of Mathematical Functions*, Ed. by M. Abramowitz and I. A. Stegun (National Bureau of Standards, Washington, 1964; Nauka, Moscow, 1979).
8. S. Stenholm, *Foundations of Laser Spectroscopy* (Wiley, New York, 1984; Mir, Moscow, 1987).
9. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 1: *Mechanics* (Nauka, Moscow, 1988; Pergamon, New York, 1988).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 728–738.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 847–859.
Original Russian Text Copyright © 2002 by Parkhomenko.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Light-Induced Drift of Ions in a Magnetic Field

A. I. Parkhomenko

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

Received November 27, 2001

Abstract—The effect of a magnetic field on the ion drift in a weakly ionized gas under the combined action of the light-induced drift and light pressure is studied theoretically. It is shown that, under the action of light, a component of ion drift velocity transverse to the direction of propagation of radiation may appear in a weakly ionized gas upon the application of an external magnetic field. It is shown that the Lorentz force acting on ions in the magnetic field radically changes the dependence of the ion drift velocity on the radiation frequency detuning. It is predicted that the ion drift velocity component along the direction of radiation must reverse its sign upon an increase in the magnetic field and an anomalous light-induced drift may be observed. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

After the prediction of light-induced drift [1] and its first experimental observation [2], a large number of experimental and theoretical publications appeared on this subject (see, for example, [3–13] and the references cited therein). The effect consists in the emergence of a directed macroscopic flow of particles absorbing radiation in a mixture with buffer particles. The origin of this effect is as follows. As a result of the Doppler effect, the radiation produces a selective effect on the absorbing particles depending on their velocities by creating effective counterpropagating "beams" of particles in an excited and the ground state. In the buffer gas atmosphere, these beams experience different drags due to different transport collision frequencies for excited and nonexcited particles. As a result, the gas of absorbing particles acquires a directional motion as a single entity. The drift velocity is proportional to the relative difference $(\nu_n - \nu_m)/\nu_n$ in the transport collision frequencies of resonant particles with buffer particles in the ground (ν_n) and excited (ν_m) states. This forms the basis of one of the main scientific applications of the light-induced drift effect, viz., the measurement of the relative variation of transport collision frequencies during the excitation of particles.

The light-induced drift effect is one of the strongest effects of radiation on the translational motion of particles. Theoretically, the velocity of the light-induced drift may attain the value of thermal velocity in the case of laser pumping [4]. It has been shown experimentally that, as a result of the light-induced drift effect, atoms may move with a velocity of the order of several tens of meters per second [6]. The light-induced drift effect has been registered experimentally for nearly two dozen different objects (atoms and molecules). Light-induced drift is possible not only for atoms and molecules in a gaseous medium, but also for ions in a weakly ionized gas [14], conduction electrons in solids [15, 16], and Wannier–Mott excitons in semiconductors [17].

It is clear from simple physical considerations that an external magnetic field may strongly influence the light-induced drift of charged particles due to the Lorentz force acting on particles drifting in a magnetic field. The force aspect of the effect of a magnetic field on the light-induced drift of charged particles was not investigated until recently. This problem was considered for the first time in [18, 19] for the light-induced drift of ions. Unfortunately, the results obtained in [18, 19] cannot be considered as reliable since the effect of a magnetic field on light-induced drift is taken into account incorrectly in these publications.[1] In a recent publication [20], the force aspect of the effect of a magnetic field on the light-induced drift of ions is investigated theoretically in the limiting cases of homogeneous broadening of absorption lines ($\Gamma \gg kv_T$, where Γ is the homogeneous half-width of the absorption line for ions and kv_T is the Doppler width) or strong magnetic fields (for $\omega_c \gg kv_T$ and for an arbitrary relation between Γ and kv_T; ω_c is the cyclotron frequency of ions in the magnetic field). In [20], some interesting features of the light-induced drift of ions in a magnetic field were revealed, such as the emergence of a drift velocity component transverse to the direction of propagation of

[1] In [18, 19], the effect of a magnetic field on the drift velocity $\mathbf{u}(\mathbf{r}, t)$ of ions (in the notation adopted in these works) appearing in the equation of the ion flux (Eq. (8) in [18] and Eq. (1) in [19]) is disregarded. The equation for the ion flux analyzed in [18, 19] is in fact a modified first equation of the system of equations (8) of the present work (for $\mathbf{a} = 0$ and $\boldsymbol{\xi} = 0$) and has the following form in our notation: $\nu_n\mathbf{J} + (\nu_m - \nu_n)\mathbf{j}_0 = -(v_T^2/2)\nabla N + \omega_c\mathbf{J} \times \mathbf{h}$, where the flux $\mathbf{j}_0$ is independent of the magnetic field, in contrast to the flux $\mathbf{j}_m$ in (8).

radiation, sign reversal of the drift velocity component along the direction of radiation upon an increase in the magnetic field, and a radical change in the dependence of the drift velocity of ions on the radiation frequency detuning. However, the most interesting case of moderate magnetic fields ($\omega_c \ll kV_T$) cannot be described by the formulas derived in [20] for the Doppler line broadening ($kV_T \gg \Gamma$).

The present work aims at a theoretical analysis and study of the photoinduced ion drift in a magnetic field of an arbitrary magnitude for an arbitrary relation between the Doppler and homogeneous absorption line widths. As compared to [20], we use another approach to calculating the light-induced drift velocity, which makes it possible to remove the limitations imposed in [20] on the magnitude of the magnetic field and on the relation between Γ and kV_T. The force effect of the external magnetic field on the light-induced drift of ions is maximal in the case when the magnetic field is perpendicular to the drift velocity direction. It is this case that is analyzed in the present work.

The light-induced drift effect may exceed the well-known effect of light pressure in its manifestation by several orders of magnitude even under optimal conditions for this phenomenon. However, for specific objects and under certain conditions, the combined analysis of these effects might be required (in the case when the light-induced drift effect is "suppressed" for some reason or is comparable to or is even weaker than the effect of light pressure). For this reason, we will consider here the combined action of these effects.

2. INITIAL EQUATIONS

Let us consider a three-component weakly ionized gas consisting of electrons, singly charged positive ions of one species, and neutral atoms in a constant uniform magnetic field **B**. The collisions between charged particles in a weakly ionized gas are insignificant since the frequencies of collisions of electrons and ions with neutral atoms are much higher than the frequencies of collisions between these particles (this condition assumes the degree of gas ionization $\lesssim 10^{-4}$ at a temperature of the order of 0.1 eV [21]). Let us suppose that radiation in the form of a propagating monochromatic wave is resonantly absorbed during the m–n transition between the ground (n) and the first excited (m) states of ions. We will concentrate our attention only on the analysis of the force effect of the magnetic field on the ion drift; consequently, we confine our analysis to the simplest case when the Zeeman splitting of the absorption line can be disregarded. For example, there is no line splitting in the case of the simple Zeeman effect (the equality of Landé g factors of the mixing states m and n) for the transverse (relative to the magnetic field **B**) direction of propagation of radiation polarized linearly along **B**.

Under these conditions, the interaction between radiation and two-level particles (ions) taking into account the recoil effect can be described by the following equations for the density matrix [9, 22]:

$$\begin{aligned}
\left[\frac{d}{dt} + \Gamma_m\right]\rho_m(\mathbf{v}) &= S_m(\mathbf{v}) + NP(\mathbf{v} - \boldsymbol{\xi}),\\
\frac{d}{dt}\rho_n(\mathbf{v}) &= S_n(\mathbf{v}) + \hat{\Gamma}_m\rho_m(\mathbf{v}) - NP(\mathbf{v} + \boldsymbol{\xi}),\\
\left[\frac{d}{dt} + \frac{\Gamma_m}{2} - i(\Omega - \mathbf{k}\cdot\mathbf{v})\right]\rho_{mn}(\mathbf{v})&\\
= S_{mn}(\mathbf{v}) + iG[\rho_n(\mathbf{v} - \boldsymbol{\xi}) - \rho_m(\mathbf{v} + \boldsymbol{\xi})],&
\end{aligned} \tag{1}$$

where

$$\begin{aligned}
\frac{d}{dt} &\equiv \frac{\partial}{\partial t} + \mathbf{v}\frac{\partial}{\partial \mathbf{r}} + \mathbf{a}_i\frac{\partial}{\partial \mathbf{v}},\\
\mathbf{a}_i &= \frac{e\mathbf{E}}{M} + \omega_c\mathbf{v}\times\mathbf{h}, \quad \omega_c = \frac{eB}{Mc},\\
\hat{\Gamma}_m\rho_m(\mathbf{v}) &= \frac{\Gamma_m}{4\pi}\int\rho_m(\mathbf{v} + 2\xi\mathbf{n}_r)d\mathbf{n}_r,\\
NP(\mathbf{v}) &= -2\mathrm{Re}[iG^*\rho_{mn}(\mathbf{v})],\\
\boldsymbol{\xi} &= \frac{\hbar\mathbf{k}}{2M}, \quad |G|^2 = \frac{B_{nm}I}{2\pi},\\
B_{nm} &= \frac{\lambda^2\Gamma_m}{4\hbar\omega}, \quad \Omega = \omega - \omega_{mn}.
\end{aligned} \tag{2}$$

Here, $\mathbf{h} = \mathbf{B}/B$ is the unit vector in the direction of the magnetic field; $\rho_i(\mathbf{v})$is the velocity distribution of ions at level $i = m, n$; N is the total concentration of ions; $S_m(\mathbf{v})$, $S_n(\mathbf{v})$, and $S_{mn}(\mathbf{v})$ are the collision integrals for ions; ω, λ, and $\mathbf{k}$ are the frequency, wavelength, and the wave vector of radiation; Γ_m is the spontaneous decay rate for the excited state m; $\hat{\Gamma}_m\rho_m(\mathbf{v})$ is the integral operator describing the radiative transition of particles from the excited level m to the ground level n taking into account the change in the velocity of particles due to the recoil effect during spontaneous emission; $\mathbf{n}_r$ is the unit vector defining the direction of spontaneous radiation; ω_{mn} is the frequency of the m–n transition; B_{nm} is the second Einstein coefficient for the m–n transition; I is the radiation intensity; $P(\mathbf{v})$ is the number of radiation absorption acts per unit time for an ion with a preset velocity $\mathbf{v}$ from the unit velocity interval; $2\boldsymbol{\xi}$ is the recoil velocity of an ion after the absorption of a photon; ω_c is the cyclotron frequency of ions; e is the elementary electric charge; M is the ion mass; **B** is the magnetic induction; and **E** is the internal electric field strength in the medium.

The electric field **E** in the medium may emerge due to the directional motion of ions as a single entity as a result of light-induced drift and light pressure effects. Two different cases are possible here. If the concentration of charged particles is not high enough for the ionized gas to display the properties of a plasma (the

Debye radius r_d characterizing the spatial separation of charged particles is much larger than the characteristic size L of the system), electrons do not affect the ion drift, and field $\mathbf{E}$ in Eqs. (1) can be neglected.

If, however, the concentration of the charged particles is high enough for the ionized gas to manifest the properties of a plasma ($r_d \ll L$), the directional motion of ions must induce the directional motion of electrons in view of the quasineutrality condition for the plasma. This leads to the emergence of the electric field $\mathbf{E}$ compensating the force of friction between electrons and buffer particles (neutral atoms).

Thus, under the plasma conditions, the motion of electrons is matched with the motion of ions through the electric field $\mathbf{E}$, and Eqs. (1) should be supplemented with the equation for the electron distribution function $\rho_e(\mathbf{v})$:

$$\left[\frac{\partial}{\partial t} + \mathbf{v}\frac{\partial}{\partial \mathbf{r}} + \mathbf{a}_e\frac{\partial}{\partial \mathbf{v}}\right]\rho_e(\mathbf{v}) = S_e(\mathbf{v}), \tag{3}$$

where

$$\mathbf{a}_e = -\frac{e\mathbf{E}}{m} - \omega_e \mathbf{v}\times\mathbf{h}, \quad \omega_e = \frac{eB}{mc}, \tag{4}$$

m is the electron mass, ω_e is the electron cyclotron frequency, and $S_e(\mathbf{v})$ is the collision integral for electrons.

For the nondiagonal collision integral $S_{mn}(\mathbf{v})$ in Eqs. (1), we will use the following approximation, which is conventional in nonlinear spectroscopy [9, 22]:

$$S_{mn}(\mathbf{v}) = -\left(\Gamma - \frac{\Gamma_m}{2}\right)\rho_{mn}(\mathbf{v}), \tag{5}$$

indicating that collisions completely shift the phase of the oscillating dipole moment (Γ is the homogeneous absorption line half-width for ions).

Inelastic collision-related processes (ionization, recombination, etc.) are insignificant in the problem under investigation (the effective frequencies of ionization and recombination are smaller than the frequencies of elastic collisions); for this reason, we will confine the subsequent analysis to the inclusion of only elastic collisions of ions and electrons with buffer particles (neutral atoms).

It is well known [5, 7, 9] that many experimental results of investigation of light-induced drift are successfully described by the relevant "standard" theory with velocity-independent transport frequencies of collisions between resonant and buffer particles. The strong deviation from the "standard" theory (the so-called anomalous light-induced drift[2] is observed only when the difference between the transport collision frequencies for resonant particles as a function of velocity $\mathbf{v}$,

$$\Delta\nu(v) \equiv \nu_m(v) - \nu_n(v),$$

at combining levels (not affected by radiation) reverses its sign [12, 13]. In the present work, we assume that $\Delta\nu(v)$ as a function of velocity v does not reverse its sign; i.e., the light-induced drift of ions is successfully described by the theory with the velocity-independent transport collision frequencies. In this case, the following relation holds for the first moment of the diagonal collision integrals [9]:

$$\begin{aligned} \int \mathbf{v} S_i(\mathbf{v}) d\mathbf{v} &= -\nu_i \mathbf{j}_i, \\ \mathbf{j}_i &= \int \mathbf{v}\rho_i(\mathbf{v}) d\mathbf{v}, \quad i = m, n, e, \end{aligned} \tag{6}$$

where ν_i is the mean transport frequency of collisions, $\mathbf{j}_m$ and $\mathbf{j}_n$ are the fluxes of ions in states m and n, and $\mathbf{j}_e$ is the electron flux. For ions ($i = m, n$), the mean transport frequency is connected with the diffusion coefficient D_i for ions in state i through the following simple formula:

$$\nu_i = \frac{v_T^2}{2D_i}, \quad v_T = \sqrt{\frac{2k_B T}{M}}, \tag{7}$$

where v_T is the most probable velocity of ions, T is the temperature, and k_B is the Boltzmann constant. The diffusion coefficient for electrons ($i = e$) is given by

$$D_e = v_e^2/2\nu_e,$$

where v_e is the most probable velocity of electrons.

3. EQUATIONS FOR PARTICLE FLUXES

In order to calculate the drift velocity of ions, it is convenient to go over in the subsequent analysis from the kinetic equations (1), (3) to the equations for particle fluxes (hydrodynamic equations). We multiply the first two equations in (1) and Eq. (3) by $\mathbf{v}$ and then integrate with respect to $\mathbf{v}$. Taking into account Eqs. (6), we obtain

$$\begin{aligned} &\left(\frac{\partial}{\partial t} + \nu_n\right)\mathbf{J} + \frac{1}{M}\sum_{\alpha\beta}\mathbf{e}_\alpha \frac{\partial}{\partial x_\beta}P_{\alpha\beta} + (\nu_m - \nu_n)\mathbf{j}_m \\ &\quad = \mathbf{a}N + \omega_c \mathbf{J}\times\mathbf{h} + 2NP\boldsymbol{\xi}, \\ &\left(\frac{\partial}{\partial t} + \Gamma_m + \nu_m\right)\mathbf{j}_m + \frac{1}{M}\sum_{\alpha\beta}\mathbf{e}_\alpha \frac{\partial}{\partial x_\beta}P^{(m)}_{\alpha\beta} \\ &\quad = \mathbf{a}N_m + \omega_c \mathbf{j}_m\times\mathbf{h} + NP\boldsymbol{\xi} + N\int \mathbf{v}P(\mathbf{v})d\mathbf{v}, \end{aligned} \tag{8}$$

[2] In 1992, an unexpectedly strong deviation of the frequency dependence of the drift velocity from a dispersion-like curve was discovered [10] during the investigation of the light-induced drift of C_2H_4 molecules in the buffer gas Kr. An anomalous spectral profile of the drift velocity was observed with three zeros instead of one as per the theory of light-induced drift with velocity-independent transport collision frequencies prevailing at that time. The departure from the prediction of the theory was so strong that the effect was called the "anomalous" light-induced drift.

$$\left(\frac{\partial}{\partial t} + \nu_e\right)\beta\mathbf{j}_e + \frac{1}{M}\sum_{\alpha\beta}\mathbf{e}_\alpha\frac{\partial}{\partial x_\beta}P^{(e)}_{\alpha\beta} = \mathbf{a}N_e + \omega_c\mathbf{j}_e\times\mathbf{h} = 0,$$

where

$$P = \int P(\mathbf{v})d\mathbf{v}, \quad \mathbf{J} = \mathbf{j}_m + \mathbf{j}_n, \quad N = N_m + N_n,$$
$$\mathbf{a} = \frac{e\mathbf{E}}{M}, \quad \beta = \frac{m}{M},$$
$$P_{\alpha\beta} = P^{(m)}_{\alpha\beta} + P^{(n)}_{\alpha\beta}, \quad P^{(i)}_{\alpha\beta} = M_i\int v_\alpha v_\beta\rho_i(\mathbf{v})d\mathbf{v}, \tag{9}$$
$$N_i = \int\rho_i(\mathbf{v})d\mathbf{v}, \quad i = m, n, e,$$

P is the number of radiation absorption acts per unit time for an ion, $\mathbf{J}$ is the total ion flux, N_m and N_n are the ion concentrations in states m and n, N_e is the electron concentration, $\mathbf{e}_\alpha$ is the unit vector in the direction of the coordinate axis x_α, v_α is the component of velocity $\mathbf{v}$ along the x_α axis, $P^{(i)}_{\alpha\beta}$ is the momentum flux density tensor for ions in states m and n ($M_i = M$) and for electrons ($M_i = m$), $\mathbf{a}$ is the acceleration of ions due to the internal electric field $\mathbf{E}$, and $\beta = m/M$ is the ratio of the electron and ion masses.

In order to simplify the problem, we confine our analysis to the condition of weak radiation intensity, assuming that the rate of induced transitions is smaller than the rate Γ_m of the radiative decay of the excited level m ($P \ll \Gamma_m$). Under these conditions, we can disregard the term $\mathbf{a}N_m$ in the second equation from (8), which is quadratic in the radiation intensity. Under standard and spatially homogeneous conditions taking into account the quasineutrality of the plasma ($N_e = N$), Eqs. (8) assume the form

$$\nu_n\mathbf{J} + (\nu_m - \nu_n)\mathbf{j}_m = \mathbf{a}N + \omega_c\mathbf{J}\times\mathbf{h} + 2NP\boldsymbol{\xi},$$
$$(\Gamma_m + \nu_m)\mathbf{j}_m = \omega_c\mathbf{j}_m\times\mathbf{h} + NP\boldsymbol{\xi} + N\int\mathbf{v}P(\mathbf{v})d\mathbf{v}, \tag{10}$$
$$\beta\nu_e\mathbf{j}_e + \mathbf{a}N + \omega_c\mathbf{j}_e\times\mathbf{h} = 0.$$

By definition, the drift velocity of ions is $\mathbf{u} \equiv \mathbf{J}/N$ and can be found from the system of equations (10) through the zeroth ($P = \int P(\mathbf{v})d\mathbf{v}$) and first ($\int\mathbf{v}P(\mathbf{v})d\mathbf{v}$) moments of the probability $P(\mathbf{v})$ of radiation absorption per unit time by an ion with a preset velocity $\mathbf{v}$.

In the case when the concentration of charged particles is not sufficiently high for the ionized gas to display the properties of a plasma (gas conditions, $r_d \gg L$), electrons do not affect the ion drift, and we can set $\mathbf{a} = 0$ in Eqs. (10). In this case, the drift velocity of ions can be determined from the first two equations of system (10).

If the concentration of charged particles is high enough for the ionized gas to display the plasma properties (plasma conditions, $r_d \ll L$), the ion acceleration $\mathbf{a}$ due to the internal electric field $\mathbf{E}$ cannot be disregarded any longer in Eqs. (10), and the drift velocity must be determined from the system of three equations (10).

The formulas for the drift velocity under the gas conditions can be easily derived from the formulas for the drift velocity under the plasma conditions by substituting the effective collision frequency $\tilde{\nu}_n$ by ν_n (see Eq. (18)). Let us therefore consider the drift of ions under the plasma conditions first.

The continuity equations for ions and electrons (which follow from Eqs. (1) and (3) integrated over $\mathbf{v}$ taking into account the relation

$$\int S_i(\mathbf{v})d\mathbf{v} = 0, \quad i = m, n, e,$$

reflecting the conservation of the number of particles in elastic collisions) combined with the quasineutrality condition for the plasma ($N_e = N$) lead to the condition

$$\mathrm{div}\mathbf{J} = \mathrm{div}\mathbf{j}_e, \tag{11}$$

which indicates the relation between the ion and electron fluxes flowing into each volume element. In a magnetic field, condition (11) may also be satisfied for $\mathbf{J} \neq \mathbf{j}_e$ due to the anisotropy of the mobility and diffusion coefficients for charged particles.

Let us now determine the relation between fluxes $\mathbf{J}$ and $\mathbf{j}_e$ for the case of ion drift under the action of a plane light wave propagating at right angles to the magnetic field. For this purpose, we consider the following configuration in the cylindrical system of coordinates ρ, φ, z with the z axis directed along the uniform magnetic field $\mathbf{B}$. Let a cylindrical monochromatic wave with the wave vector $\mathbf{k}$ perpendicular to the z axis diverge radially from a radiation source extended along the z axis in a homogeneous unbounded plasma. Then, the light-induced drift velocity of ions is a function of radius ρ only, and Eq. (11) leads to the equality of the radial components of fluxes, $\mathbf{J}_\rho = \mathbf{j}_{e\rho}$; i.e., the drift of ions and electrons along the vector $\mathbf{k}$ is ambipolar. Under steady-state conditions, the internal electric field $\mathbf{E}$ emerging in the plasma due to the ion drift is vortex-free (curl $\mathbf{E} = 0$) and, hence, the azimuthal component of the field $\mathbf{E}_\varphi = 0$. Thus, condition (11) and relation $\mathbf{E}_\varphi = 0$ in the case under investigation are equivalent to the conditions

$$\mathbf{j}_{e\parallel} = \mathbf{J}_\parallel, \quad \mathbf{a} = \frac{\mathbf{k}}{k}a, \tag{12}$$

where the symbol "$\parallel$" indicates the vector component directed along $\mathbf{k}$.

Let us now take into account the fact that individual segments of a cylindrical wave, which are smaller than the distance from the source of radiation, behave approximately as plane waves with constant radiation

intensity. The light-induced drift of particles in these regions can be regarded as a drift under the action of a plane light wave. Thus, conditions (12) also hold for the ion drift induced by a plane light wave, which is considered by us here.

Solving the system of equations (10) under conditions (12), we find that, for the direction of propagation of radiation transverse to the magnetic field (for $\mathbf{k} \perp \mathbf{B}$), the drift velocity of ions is equal to the sum of two mutually perpendicular components $\mathbf{u}_\parallel$ and $\mathbf{u}_\perp$:

$$\mathbf{u} = \mathbf{u}_\parallel + \mathbf{u}_\perp, \tag{13}$$

where component $u_\parallel$ is parallel to the wave vector $\mathbf{k}$, while component $u_\perp$ is perpendicular to $\mathbf{k}$ and $\mathbf{B}$:

$$\mathbf{u}_\parallel = \frac{\mathbf{k}}{k}u_\parallel, \quad \mathbf{u}_\perp = \mathbf{n}u_\perp, \quad \mathbf{n} = \frac{\mathbf{k}\times\mathbf{B}}{kB}. \tag{14}$$

The components $u_\parallel$ and $u_\perp$ of the drift velocity $\mathbf{u}$ along directions $\mathbf{k}$ and $\mathbf{n}$ are given by the formulas

$$u_\parallel \equiv \frac{\mathbf{k}\cdot\mathbf{u}}{k}\tau_\sigma$$

$$\times\left\{\left[1-\frac{\omega_c^2}{\nu_n(\Gamma_m+\nu_m)}\right]Q_\parallel - \frac{\omega_c(\Gamma_m+\nu_m+\nu_n)}{\nu_n(\Gamma_m+\nu_m)}Q_\perp\right\} \tag{15}$$

$$+ u_{0r}\left\{1+\frac{(\nu_n-\nu_m)\left[1-\dfrac{\omega_c^2}{\nu_n(\Gamma_m+\nu_m)}\right]}{2\left(\Gamma_m+\nu_m+\dfrac{\omega_c^2}{\Gamma_m+\nu_m}\right)}\right\},$$

$$u_\perp \equiv \mathbf{n}\cdot\mathbf{u} = \tau_\sigma\left\{\frac{\omega_c(\Gamma_m+\nu_m+\tilde{\nu}_n)}{\nu_n(\Gamma_m+\nu_m)}Q_\parallel\right.$$

$$\left.+\left[\frac{\tilde{\nu}_n}{\nu_n}-\frac{\omega_c^2}{\nu_n(\Gamma_m+\nu_m)}\right]Q_\perp\right\} + u_{0r}\frac{\omega_c}{\nu_n} \tag{16}$$

$$\times\left\{1+\frac{(\nu_n-\nu_m)(\Gamma_m+\nu_m+\tilde{\nu}_n)}{2\left[\Gamma_m+\nu_m+\dfrac{\omega_c^2}{\Gamma_m+\nu_m}\right](\Gamma_m+\nu_m)}\right\},$$

where

$$u_{0r} = \frac{2\xi P}{\tilde{\nu}_n+\omega_c^2/\nu_n},$$

$$\tau_\sigma = \frac{\nu_n-\nu_m}{\left[\tilde{\nu}_n+\dfrac{\omega_c^2}{\nu_n}\right]\left[\Gamma_m+\nu_m+\dfrac{\omega_c^2}{\Gamma_m+\nu_m}\right]}, \tag{17}$$

$$\tilde{\nu}_n = \nu_n + \beta\nu_e + \frac{\omega_c^2}{\beta\nu_e}, \quad Q_\parallel = \int\frac{\mathbf{k}}{k}\mathbf{v}P(\mathbf{v})d\mathbf{v},$$

$$Q_\perp = \int\mathbf{n}\cdot\mathbf{v}P(\mathbf{v})d\mathbf{v}.$$

Formulas (15)–(17) describe the drift of ions under the plasma conditions ($r_d \ll L$).

Under the gas conditions ($r_d \gg L$), the drift velocity of ions is defined by the same formulas (15)–(17) with the substitution

$$\tilde{\nu}_n = \nu_n. \tag{18}$$

Thus, the evaluation of drift velocity boils down to the calculation of the probability $P(\mathbf{v})$ of radiation absorption per unit time by an ion with a given velocity $\mathbf{v}$.

4. PROBABILITY OF RADIATION ABSORPTION

The probability $P(\mathbf{v})$ of radiation absorption at the m–n transition is determined by the nondiagonal density matrix element $\rho_{mn}(\mathbf{v})$. In the case of a low radiation intensity ($P \ll \Gamma_m$), we can disregard the population of the excited level ($\rho_m(\mathbf{v}) = 0$) in the equation for $\rho_{mn}(\mathbf{v})$ in system (1) and assume that the velocity distribution of populations in the ground state is close to the Maxwell distribution ($\rho_n(\mathbf{v}) = NW(\mathbf{v})$), where $W(\mathbf{v})$ is the Maxwell distribution). For a low radiation intensity, we can also disregard the internal electric field $\mathbf{E}$ in the equation for $\rho_{mn}(\mathbf{v})$ in system (1) and assume that $\mathbf{a}_i = \omega_c\mathbf{v}\times\mathbf{h}$. In this case, under steady-state and spatially homogeneous conditions, we obtain from Eqs. (1) taking into account relation (5) the following expression in the approximation linear in the small parameter ξ/v_T:

$$\left[\Gamma - i(\Omega-\mathbf{k}\cdot\mathbf{v}) + \omega_c\mathbf{v}\times\mathbf{h}\frac{\partial}{\partial\mathbf{v}}\right]\rho_{mn}(\mathbf{v}) = iGNW(\mathbf{v})\left[1+\frac{2\mathbf{v}\cdot\boldsymbol{\xi}}{v_T^2}\right]. \tag{19}$$

We will solve this equation in a system of coordinates in which the z axis is directed along the magnetic field $\mathbf{B}$ and the x axis is directed along the wave vector $\mathbf{k}$ (we assume that $\mathbf{k} \perp \mathbf{B}$). In the velocity space, it is convenient to go over to the cylindrical system of coor-

dinates $\mathbf{v}_\perp$, φ, v_z ($v_x = v_\perp\cos\varphi$, $v_y = v_\perp\sin\varphi$). In these coordinates, Eq. (19) assumes the form

$$\left\{\Gamma - i(\Omega - kv_\perp\cos\varphi) - \omega_c\frac{\partial}{\partial\varphi}\right\}\rho_{mn}(\mathbf{v}) = iGNW(v_\perp)W(v_z)\left[1 + \frac{2\xi v_\perp\cos\varphi}{v_T^2}\right], \quad (20)$$

where

$$W(v_\perp) = \frac{1}{(\sqrt{\pi}v_T)^2}\exp\left(-\frac{v_\perp^2}{v_T^2}\right), \qquad W(v_z) = \frac{1}{\sqrt{\pi}v_T}\exp\left(-\frac{v_z^2}{v_T^2}\right), \quad (21)$$

$W(v_\perp)$ and $W(v_z)$ being the Maxwell distributions over the transverse and longitudinal (relative to the magnetic field **B**) components of velocity **v**.

Solving this linear nonhomogeneous differential equation, we obtain the following expression for the radiation absorption probability:

$$P(\mathbf{v}) = 2|G|^2W(v_\perp)W(v_z)\mathrm{Re}\left\{\exp\left(i\frac{kv_\perp}{\omega_c}\sin\varphi\right)\times\sum_{n=-\infty}^{n=\infty}\frac{\exp(-in\varphi)J_n\left(\frac{kv_\perp}{\omega_c}\right)\left[1 + n\frac{2\xi}{v_Tkv_T}\frac{\omega_c}{}\right]}{\Gamma - i(\Omega - n\omega_c)}\right\}, \quad (22)$$

where $J_n(x)$ is a Bessel function of the first kind. In zero magnetic field (for $\omega_c = 0$), the radiation absorption probability $P(\mathbf{v})$ is defined by the well-known formula [9, 22]

$$P(\mathbf{v}) = \frac{2|G|^2\Gamma W(\mathbf{v})}{\Gamma^2 + (\Omega - \mathbf{k}\cdot\mathbf{v})^2}. \quad (23)$$

For the radiation absorption probability integral over velocities, $P = \int P(\mathbf{v})d\mathbf{v}$, which determines the profile of the absorption line, we obtain, using Eq. (22),

$$P = 2|G|^2\Gamma\exp(-\mu)\times\sum_{n=-\infty}^{n=\infty}\frac{\left[1 + n\frac{2\xi}{v_Tkv_T}\frac{\omega_c}{}\right]I_n(\mu)}{\Gamma^2 + (\Omega - n\omega_c)^2}, \quad \mu = \frac{(kv_T)^2}{2\omega_c^2}, \quad (24)$$

where $I_n(\mu)$ is a modified Bessel function. If we disregard the recoil effects, Eq. (24) leads to the known [23] expression for the spectral line profile for ions in a magnetic field. It follows from Eq. (24) that the Larmor rotation of ions in a magnetic field may lead to the emergence of equidistant peaks (cyclotron resonances) in the absorption line profile. The distance between adjacent peaks is equal to the cyclotron frequency ω_c of ions, and the width of an individual peak is determined by the homogeneous width 2Γ of the absorption line. For $kv_T \gg \Gamma$, the absorption line described by formula (24) has the form of a Doppler contour modulated by the periodic function of the radiation frequency detuning Ω with a period equal to ω_c. The oscillating function P (24) has peaks for $\Omega = n\omega_c$ and minima for $\Omega = (n + 1/2)\omega_c$. For $\omega_c \gg \Gamma$, the Doppler contour distinctly splits into a series of peaks, while, for $\omega_c \lesssim \Gamma$, the shape of the line differs from the Doppler contour by an exponentially small oscillating correction [23].

A nontrivial aspect is the effect of the magnetic field on the behavior of the radiation absorption probability $P(\mathbf{v})$ as a function of velocity **v**. An analysis of expression (22) shows that, for some values of velocity **v**, the function $P(\mathbf{v})$ may assume negative values. In addition, the integral characteristics

$$P(v_\perp, \varphi) = \int_{-\infty}^{\infty}P(\mathbf{v})dv_z, \qquad P(\varphi) = \int_0^\infty P(v_\perp, \varphi)v_\perp dv_\perp \quad (25)$$

may also assume negative values (here, $P(\varphi)$ is the number of radiation absorption acts per unit time in a unit interval of angles per ion with a given value of φ between the direction of radiation **k** and the projection of the velocity of ions onto the plane perpendicular to the magnetic field). In other words, in a gas of ionized particles in an external magnetic field, a partial (in the directions of velocity) noninversive amplification of radiation by ions due to their Larmor rotation may emerge upon the application of the external magnetic field. In this case, virtually all ions may be in the ground state. The partial (in velocities) noninversive amplification is a "latent" effect in the sense that it disappears as a result of averaging over all directions of ion velocities (the radiation absorption probability P integrated over velocities is positive). A detailed analysis of this effect would be of interest, but it is beyond the scope of this research.

5. DRIFT VELOCITY

Using formula (22), we derive the following expression for the quantities $Q_\parallel$ and $Q_\perp$ defined in Eqs. (17) and appearing in formulas (15) and (16) for drift velocity:

$$Q_\parallel = \frac{2|G|^2\omega_c}{k}\exp(-\mu)\times\sum_{n=-\infty}^{n=\infty}\frac{\left[1 + n\frac{2\xi}{v_Tkv_T}\frac{\omega_c}{}\right]\Gamma nI_n(\mu)}{\Gamma^2 + (\Omega - n\omega_c)^2}, \quad (26)$$

$$Q_\perp = \frac{2|G|^2\omega_c}{k}\exp(-\mu)\sum_{n=-\infty}^{n=\infty}\frac{\left[1+n\frac{2\xi}{v_T}\frac{\omega_c}{kv_T}\right](n\omega_c-\Omega)[(|n|-\mu)I_n(\mu)+\mu I_{|n|+1}(\mu)]}{\Gamma^2+(\Omega-n\omega_c)^2}. \quad (27)$$

This completes the calculation of the drift velocity of ions. The drift velocity can be determined from formulas (15) and (16) by substituting the radiation absorption probability P from Eq. (24) and the quantities $Q_\parallel$ and $Q_\perp$ defined by Eqs. (26) and (27).

According to these formulas, the right-hand sides of expressions (15) and (16) for $u_\parallel$ and $u_\perp$ can be written in the form of the sum of two terms:

$$u_\parallel = u_{L\parallel} + u_{r\parallel}, \quad u_\perp = u_{L\perp} + u_{r\perp}, \quad (28)$$

where velocities $u_{L\parallel}$ and $u_{L\perp}$ differ from zero only for $\nu_m \neq \nu_n$ and are independent of the recoil velocity 2ξ of an ion upon the absorption of a photon (light-induced drift), while velocities $u_{r\parallel}$ and $u_{r\perp}$ differ from zero only for $\xi \neq 0$ (light pressure). Thus, the drift velocity $\mathbf{u}$ (13) can also be presented as the sum of the drift velocities $\mathbf{u}_L$ and $\mathbf{u}_r$ associated with the effects of light-induced drift ($\mathbf{u}_L$) and light pressure ($\mathbf{u}_r$):

$$\mathbf{u} = \mathbf{u}_L + \mathbf{u}_r. \quad (29)$$

The formulas for velocities $\mathbf{u}_L$ and $\mathbf{u}_r$ derived as a result of the relevant grouping of terms depending on and independent of ξ in Eqs. (15) and (16) are obvious and will not be given here.

In the case of strong magnetic fields ($\omega_c \gg kv_T$) or in the case of homogeneous broadening of the absorption line ($\Gamma \gg kv_T$), the formulas for drift velocity derived in the present work are simplified considerably and coincide with the formulas derived earlier in [20] using the Grad method of solution of kinetic equations if we disregard the recoil effect.

Figures 1–3 show the results of calculations of drift velocity obtained from formulas (15) and (16) by substituting Eqs. (24), (26), and (27) into them. In all the figures, we choose the unit of velocity equal to the quantity

$$u_R = \frac{2\xi P_0}{\nu_n}, \quad P_0 = \frac{2\sqrt{\pi}|G|^2}{kv_T}, \quad (30)$$

which is equal to the maximal (for $\Omega = 0$) value of the ion drift velocity $\mathbf{u}_r$ under the action of light pressure in zero magnetic field for $\nu_m = \nu_n$. The quantity P_0 is the radiation absorption probability at the center of the line for Doppler broadening in zero magnetic field. The ratio of the maximal values of velocities $\mathbf{u}_L$ and $\mathbf{u}_r$ in zero magnetic fields in the case of Doppler broadening is characterized by the parameter A:

$$\frac{|(\mathbf{u}_L)_{\omega_c\to 0}|_{\max}}{u_R} \approx \frac{v_T}{5.4\xi}\frac{|\nu_m-\nu_n|}{\Gamma_m+\nu_n} \equiv A. \quad (31)$$

For values of the parameters used for calculating the curves in Figs. 1 and 3, the value of $A \approx 600$; i.e., in zero magnetic field, the drift velocity $\mathbf{u}_L$ associated with light-induced drift effect is 600 times the velocity $\mathbf{u}_r$ of the drift induced by light pressure.

Figures 1 and 2 illustrate the dependence of the drift velocity of ions on the detuning of the radiation frequency Ω and on the magnetic field under the gas conditions ($r_d \gg L$). Figure 1a shows that, with increasing magnetic field, the drift velocity component $u_\parallel$ along the direction of radiation reverses its sign. Sign reversal occurs in the cyclotron frequency range

$$\omega_c \sim \sqrt{\nu_n(\Gamma_m+\nu_m)}. \quad (32)$$

Since the condition $|u_{L\parallel}| \gg |u_{r\parallel}|$ holds for the values of parameters used for calculating the curves in Fig. 1a (and, hence $u_\parallel \approx u_{L\parallel}$), the curves in Fig. 1a in fact illustrate the dependence of the drift velocity $u_{L\parallel}$ of the light-induced drift on the radiation frequency detuning Ω. Curves *1* and *3* in Fig. 1a correspond to the conventional light-induced drift with a typical dispersion-like frequency dependence $u_{L\parallel}(\Omega)$ of the drift velocity (which is equal, except for the sign, to the frequency derivative of the absorption line profile) with a single zero at the zero value of the radiation frequency detuning. Curve *2* with three zeros in Fig. 1a corresponds to anomalous light-induced drift [10–13] with a sharp deviation of the frequency dependence of the drift velocity $u_{L\parallel}(\Omega)$ from a dispersion-like curve. Anomalous light-induced drift and the change in the drift direction are observed for a cyclotron frequency of ions, whose order of magnitude is determined by relation (32). An analysis shows that the interval $\Delta\omega_c$ of the cyclotron frequency values in which anomalous light-induced drift takes place is approximately equal to $\Delta\omega_c \approx 0.1\nu_n$.

It was mentioned above that, in zero external fields, the anomalous light-induced drift is completely determined by the dependence of transport collision frequencies on the velocity v of resonant particles, the anomaly appearing only when the difference in the transport frequencies of collisions at mixing levels reverses its sign as a function of v. The results of the present work show that anomalous light-induced drift of ions in an external magnetic field may also appear for velocity-independent transport collision frequencies.

The physical origin of the change in the direction of the drift velocity of ions upon an increase in the magnetic field can be explained from the following qualitative considerations. In zero magnetic field, the ion drift

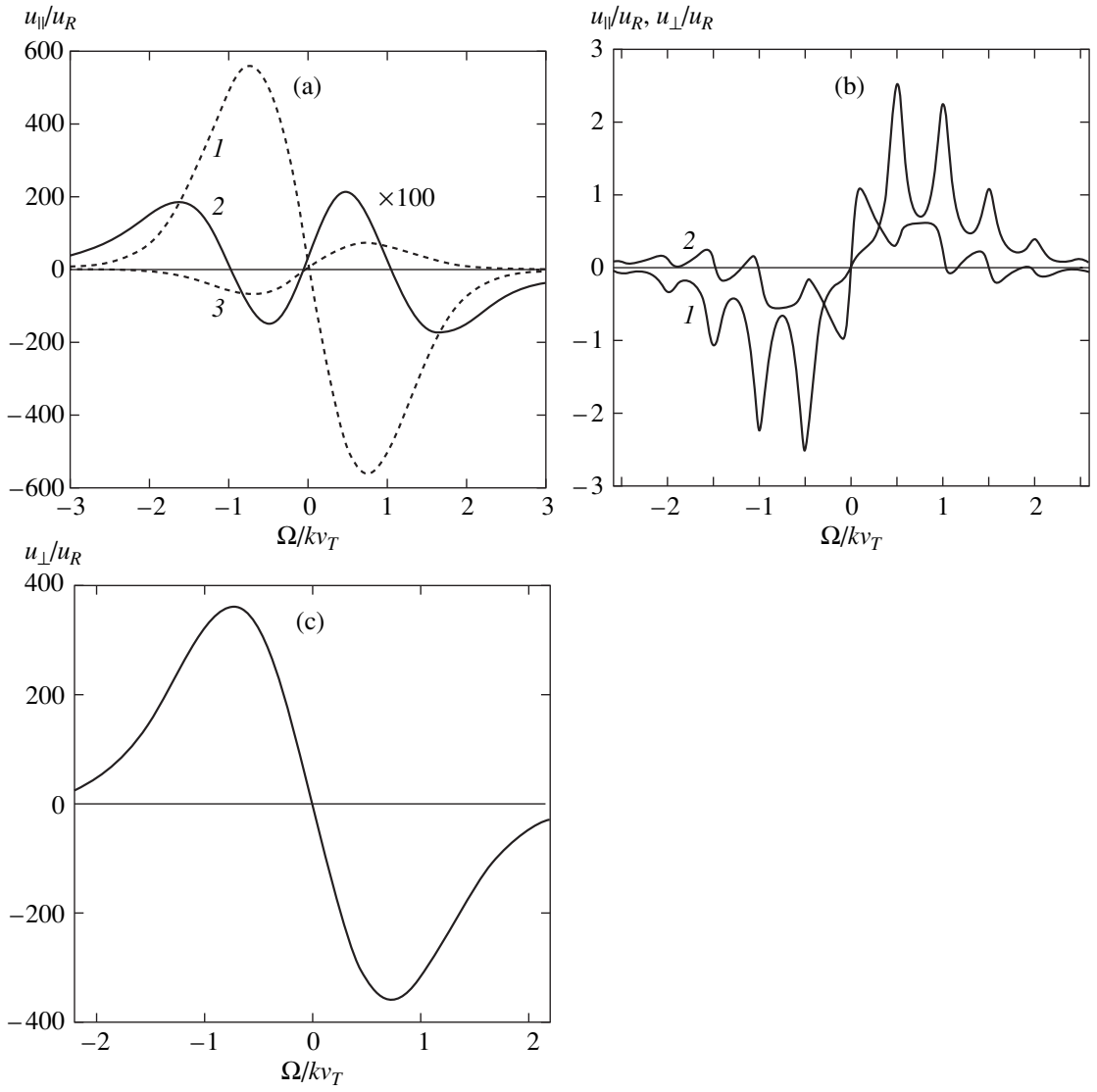

Fig. 1. Drift velocity of ions as a function of the radiation frequency detuning for various values of magnetic field (gas conditions, $r_d \gg L$): $\Gamma/kv_T = 0.1$, $(\nu_m - \nu_n)/\nu_n = 0.1$, $\nu_n/\Gamma = 0.2$, $\Gamma_m/\nu_n = 0.5$, $\nu_T/\xi = 5 \times 10^4$; (a) $\omega_c = 0$ (*1*), $\omega_c/\Gamma = 0.2485$ ($\omega_c/\nu_n = 1.2425$) (*2*); $\omega_c/\Gamma = 0.4$ ($\omega_c/\nu_n = 2$) (*3*); (b) $\omega_c/\Gamma = 5$ ($\omega_c/\nu_n = 25$), $u_\parallel/u_R$ (*1*), $u_\perp/u_R$ (*2*); (c) $\omega_c/\Gamma = 0.15$ ($\omega_c/\nu_n = 0.75$).

velocity $u_\parallel$ is proportional to the difference $\nu_n - \nu_m$ in the transport frequency of collisions of ions in the ground and in excited states with buffer particles. In the presence of a magnetic field, the diffusion coefficient D_{iB} for ions in state i across the magnetic field is

$$D_{iB} = \bar{v}^2/2\nu_{iB},$$

where the quantity

$$\nu_{iB} = \nu_i + \omega_c^2/\nu_i$$

has the meaning of the effective transport frequency of collisions of ions in state i with buffer particles in the presence of a magnetic field [21]. Consequently, in a magnetic field perpendicular to the direction of propagation of radiation, we can expect that (approximately)

$$u_\parallel \propto \nu_{nB} - \nu_{mB} \propto (\nu_m - \nu_n)(\omega_c^2 - \nu_m\nu_n).$$

It follows hence that, in an increasing magnetic field, the drift velocity component along the direction of radiation reverses its sign. The change in the direction of the drift is associated with the sign reversal of the difference $\nu_{nB} - \nu_{mB}$ in the effective transport frequencies of collisions of ions upon an increase in the magnetic field.

For the case of $\omega_c \lesssim \Gamma$ depicted in Fig. 1a, the shape of the absorption line for ions differs from the Doppler profile by an exponentially small oscillation correction [23]. Consequently, no oscillations are observed in the

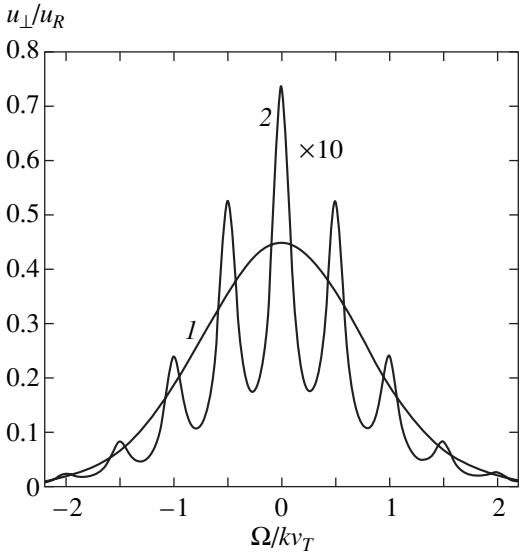

Fig. 2. Drift velocity associated with light pressure as a function of the radiation frequency detuning (gas conditions, $r_d \gg L$): $\Gamma/kv_T = 0.1$, $\nu_m = \nu_n$, $\nu_n/\Gamma = 0.2$, $\Gamma_m/\nu_n = 0.5$; $\omega_c/\Gamma = 0.2$ ($\omega_c/\nu_n = 1$) (*1*); $\omega_c/\Gamma = 5$ ($\omega_c/\nu_n = 25$) (*2*).

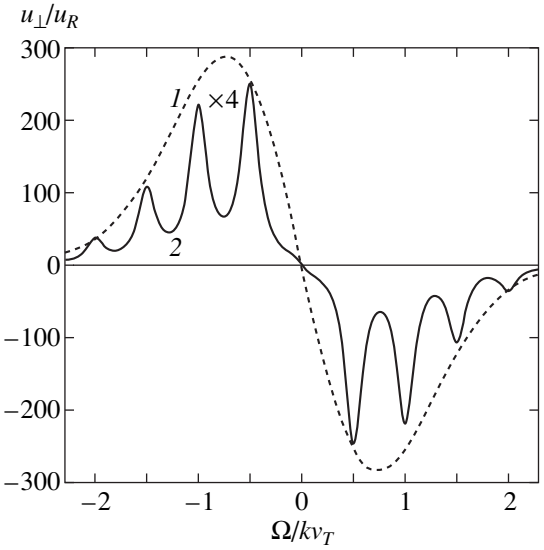

Fig. 3. Drift velocity of ions as a function of the radiation frequency detuning for various values of the magnetic field (plasma conditions, $r_d \ll L$). $\Gamma/kv_T = 0.1$, $(\nu_m - \nu_n)/\nu_n = 0.1$, $\nu_n/\Gamma = 0.2$, $\Gamma_m/\nu_n = 0.5$, $\nu_T/\xi = 5 \times 10^4$, $\beta\nu_e/\nu_n = 10^{-2}$; $\omega_c/\Gamma = 0.3$ ($\omega_c/\nu_n = 1.5$) (*1*); $\omega_c/\Gamma = 5$ ($\omega_c/\nu_n = 25$) (*2*).

dependence of the drift velocity on Ω. For $\omega_c \gg \Gamma$, the Doppler profile distinctly splits into a number of peaks [23], and the drift velocity displays oscillations as a function of Ω (see Fig. 1b).

Figure 1c shows the drift velocity component transverse to the wave vector as a function of Ω. The absolute (in Ω and ω_c) maximum of velocity $u_\perp$ is attained for $\omega_c \sim \nu_n$ and is close to the absolute maximum of the light-induced drift velocity in zero magnetic field (this can be seen from a comparison of curve *1* in Fig. 1a with Fig. 1c).

In the case of equality of the transport collision frequencies in the ground and excited states ($\nu_m = \nu_n$), no light-induced drift is observed, and ions drift only under the effect of light pressure. In this case, the dependence of the drift velocity on Ω repeats the shape of the absorption line (see Fig. 2).

Upon a transition from the gas conditions to the plasma conditions, the dependence of the drift velocity component $u_\parallel$ along the direction of radiation on Ω does not change, but its magnitude decreases (by a factor of $(\tilde{\nu}_n\nu_n + \omega_c^2)/(\nu_n^2 + \omega_c^2)$ in accordance with formula (15)). This decrease is associated with the ambipolar nature of the drift along the direction of radiation, leading to a drag exerted by electrons on the ion drift (in a strong magnetic field **B**, the transverse (relative to the field) diffusion coefficient for electrons is smaller than the transverse diffusion coefficient for ions by a factor of $\nu_n/\beta\nu_e$ [21]).

The dependence of the drift velocity component $u_\perp$ transverse to the direction of radiation on Ω changes upon a transition from the gas conditions to the plasma conditions. In weak magnetic fields (for $\omega_c \lesssim \Gamma$), the dependence $u_\perp(\Omega)$ remains virtually the same as under the gas conditions and is dispersion-like (curve *1* in Fig. 3). As the magnetic field increases (for $\omega_c \gg \Gamma$), the form of the $u_\perp(\Omega)$ dependence changes upon a transition from the gas conditions to the plasma conditions. In strong magnetic fields (for $\omega_c \gg \Gamma$), the magnitude of the drift velocity $u_\perp$ under the plasma conditions is considerably larger than under the gas conditions (this can be seen from a comparison of curves *2* in Fig. 1b and Fig. 3).

Let us now determine the ambipolar electric field **E** formed automatically in the plasma for leveling out the fluxes of oppositely charged particles along the direction of radiation. Using Eqs. (10) together with conditions (12), we find that

$$\mathbf{E} = -\frac{\mathbf{u}_\parallel}{\mu_{e\perp}}, \quad \mu_{e\perp} = \frac{\mu_e}{1 + \omega_e^2/\nu_e^2}, \quad \mu_e = \frac{e}{m\nu_e}, \quad (33)$$

where $\mu_{e\perp}$ is the electron mobility in a direction transverse to the magnetic field [21] and μ_e is the electron mobility in zero magnetic field.

Let us estimate the value of **E**. In weak magnetic fields ($\omega_c = \beta\omega_e \ll \beta\nu_e$), the electron mobility $\mu_{e\perp} \approx \mu_e$, and the maximum value of the drift velocity is $|\mathbf{u}_\parallel|_{\max} \approx Au_R$ (see formula (31)). The ambipolar electric field strength in this case is

$$|\mathbf{E}| \approx \frac{Au_R}{\mu_e} = \frac{\hbar k P_0}{e}\frac{\beta\nu_e}{\nu_n}A. \quad (34)$$

This leads to the following estimate for the radiation wavelength $\lambda \sim 0.5$ μm, the radiation absorption probability at the line center $P_0 \sim 10^7$ s^{-1}, and the values of $\beta\nu_e/\nu_n \sim 10^{-2}$ and $A \approx 600$ (see Eq. (31)): $|\mathbf{E}| \sim 5 \times 10^{-3}$ V/cm. As the magnetic field increases (so that $\omega_c^2 \gtrsim \beta\nu_e\nu_n$, which corresponds to $\tilde{\nu}_n \gtrsim \nu_n$), the drift velocity decreases,

$$|\mathbf{u}_\parallel|_{\max} \sim Au_R\beta\nu_e\nu_n/\omega_c^2,$$

but the drag effect of electrons becomes stronger,

$$\mu_{e\perp} \sim \mu_e\beta^2\nu_e^2/\omega_c^2.$$

Consequently, the ambipolar electric field strength increases by a factor of $\nu_n/\beta\nu_e$ and may attain values $|\mathbf{E}| \sim 0.5$ V/cm for $\nu_n/\beta\nu_e \sim 100$.

CONCLUSIONS

In the present work, we analyzed theoretically the force action of an external magnetic field on the ion drift under the combined action of the light-induced drift and light pressure effects under the conditions when this action is maximal and is manifested in "pure" form (the Zeeman splitting of the absorption line is absent). The force action attains its maximum value in the case when the magnetic field is perpendicular to the direction of propagation of radiation; it is precisely in this case that it can be singled out in "pure" form (there is no line splitting in the case of the normal Zeeman effect with radiation propagating across the magnetic field and polarized linearly along the magnetic field). The formulas for the drift velocity of ions derived in the present work are valid for an arbitrary relation between the Doppler and homogeneous widths of the absorption line and for an arbitrary magnitude of the magnetic field.

As the magnetic field increases, the ion drift velocity component along the direction of radiation reverses its sign. It follows from Eq. (32) that this effect can only be observed in magnetic fields

$$B \sim 10^{-4}M\sqrt{\nu_n(\Gamma_m + \nu_m)}, \qquad (35)$$

where M is the ion mass in atomic units. It follows hence that the value of the magnetic field required for observing this effect experimentally is the smaller, the lower the gas pressure and the rate of spontaneous decay of the excited state of the ion. For the transport frequency of ion collisions $\nu_n \sim 10^5$ s^{-1} (which corresponds to a gas pressure of ~0.01 Torr), the radiation constant $\Gamma_m \sim 10^7$ s^{-1}, and the ion mass $M \sim 10$ amu, we obtain the following estimate from relation (35): $B \sim 10^3$ G.

The drift velocity component transverse to the direction of radiation emerges for indefinitely weak magnetic fields. For $\omega_c \lesssim \nu_n$, its magnitude can be estimated using the formula

$$|\mathbf{u}_\perp| \sim (\omega_c/\nu_n)|\mathbf{u}_d|,$$

where $\mathbf{u}_d$ is the ion drift velocity in zero magnetic field. The magnitude of the transverse drift velocity component may attain the value $|\mathbf{u}_d|$ even in quite weak magnetic fields ($B \sim 100$ G for $\nu_n \sim 10^5$ s^{-1} and the ion mass $M \sim 10$ amu).

Under laboratory conditions, photoinduced ion drift may be manifested in the form of an electric current (photoinduced current [14]). A potential difference $V \sim |\mathbf{E}|L$ will emerge between the endfaces of a cell with a weakly ionized gas, where L is the cell length and $\mathbf{E}$ is the ambipolar electric field in the cell, which emerges due to ion drift induced by light. For $|\mathbf{E}| \sim 5 \times 10^{-3}$ V/cm (see the estimate following formula (34)) and $L \sim 10$ cm, a potential difference $V \sim 0.05$ V emerges between the endfaces of the cell. A conductor connecting the opposite ends of the cell will carry a current $I \sim V/R$, where R is the internal resistance of the plasma. Since $R \sim L/eNSm_{e\perp}$, where S is the cross-sectional area of the cell, we obtain the following estimate, taking into account Eq. (33): $I \sim |\mathbf{u}_\parallel|eNS$. For the drift velocity $|\mathbf{u}_\parallel| \sim 10$ cm/s, the ion concentration $N \sim 10^{11}$ cm^{-3}, and $S \sim 1$ cm^2, we obtain $I \sim 10^{-7}$ A.

The results obtained in the present work may be interesting for astrophysical applications in connection with the phenomenon of chemically peculiar stars, which is widely discussed in the literature [24–27]. One of the main hypotheses explains the anomalies in the chemical composition of all peculiar stars by the separation of chemical elements in their atmospheres through the mechanism of selective drift of atoms and ions under the action of radiation emitted by the star [24–27]. Both light pressure [25–28] and the light-induced drift [27, 29, 30] were considered as possible reasons for the drift in the atmospheres of such stars. Chemically peculiar stars include so-called magnetic stars [24–27] with strong (up to 3×10^4 G) large-scale magnetic fields predominantly of a dipole nature. It was shown in the present work that the magnetic field radically changes the pattern of the light-induced drift and, hence, may strongly affect the separation of chemical elements in the atmospheres of magnetic stars.

ACKNOWLEDGMENTS

The author is indebted to A.M. Shalagin for numerous discussions and valuable remarks and to F.Kh. Gel'mukhanov and L.V. Il'ichev for fruitful discussions and interest in this research.

This work was supported financially by the Russian Foundation for Basic Research (project no. 01-02-17433).

REFERENCES

1. F. Kh. Gel'mukhanov and A. M. Shalagin, Pis'ma Zh. Éksp. Teor. Fiz. **29**, 773 (1979) [JETP Lett. **29**, 711 (1979)].
2. V. D. Antsygin, S. N. Atutov, F. Kh. Gel'mukhanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **30**, 262 (1979) [JETP Lett. **30**, 261 (1979)].
3. F. Kh. Gel'mukhanov and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **78**, 1674 (1980) [Sov. Phys. JETP **51**, 839 (1980)].
4. A. K. Popov, A. M. Shalagin, V. M. Shalaev, and V. Z. Yakhnin, Zh. Éksp. Teor. Fiz. **80**, 2175 (1981) [Sov. Phys. JETP **53**, 1134 (1981)].
5. G. Nienhuis, Phys. Rep. **138**, 151 (1986).
6. S. N. Atutov, I. M. Ermolaev, and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **92**, 1215 (1987) [Sov. Phys. JETP **65**, 679 (1987)].
7. H. G. C. Werij and J. P. Woerdman, Phys. Rep. **169**, 145 (1988).
8. P. L. Chapovskiĭ, Izv. Akad. Nauk SSSR, Ser. Fiz. **53**, 1069 (1989).
9. S. G. Rautian and A. M. Shalagin, *Kinetic Problems of Nonlinear Spectroscopy* (North–Holland, Amsterdam, 1991).
10. G. J. van der Meer, J. Smeets, S. P. Pod'yachev, and L. J. F. Hermans, Phys. Rev. A **45**, R1303 (1992).
11. F. Yahyaei-Moayyed and A. D. Streater, Phys. Rev. A **53**, 4331 (1996).
12. F. Kh. Gel'mukhanov, A. I. Parkhomenko, T. I. Privalov, and A. M. Shalagin, J. Phys. B **30**, 1819 (1997).
13. A. I. Parkhomenko, Zh. Éksp. Teor. Fiz. **116**, 1587 (1999) [JETP **89**, 856 (1999)].
14. F. Kh. Gel'mukhanov and A. M. Shalagin, Kvantovaya Élektron. (Moscow) **8**, 590 (1981).
15. É. M. Skok and A. M. Shalagin, Pis'ma Zh. Éksp. Teor. Fiz. **32**, 201 (1980) [JETP Lett. **32**, 184 (1980)].
16. A. M. Dykhne, V. A. Roslyakov, and A. N. Starostin, Dokl. Akad. Nauk SSSR **254**, 599 (1980) [Sov. Phys. Dokl. **25**, 741 (1980)].
17. A. I. Parkhomenko, Fiz. Tverd. Tela (Leningrad) **25**, 2374 (1983) [Sov. Phys. Solid State **25**, 1363 (1983)].
18. S. Dattagupta, R. Ghosh, and J. Singh, Phys. Rev. Lett. **83**, 710 (1999).
19. J. Singh, R. Ghosh, and S. Dattagupta, Phys. Rev. A **61**, 025402 (2000).
20. A. I. Parkhomenko, Pis'ma Zh. Éksp. Teor. Fiz. **74**, 172 (2001) [JETP Lett. **74**, 154 (2001)].
21. V. E. Golant, A. P. Zhilinskii, and I. E. Sakharov, *Fundamentals of Plasma Physics* (Atomizdat, Moscow, 1977; Wiley, New York, 1980).
22. S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in Atomic and Molecular Spectra* (Nauka, Novosibirsk, 1979).
23. M. I. D'yakonov, Zh. Éksp. Teor. Fiz. **51**, 612 (1966) [Sov. Phys. JETP **24**, 408 (1966)].
24. S. B. Pikel'ner and V. L. Khokhlova, Usp. Fiz. Nauk **107**, 389 (1972) [Sov. Phys. Usp. **15**, 395 (1972)].
25. V. L. Khokhlova, Itogi Nauki Tekh., Ser. Astron. **24**, 233 (1983).
26. *Physics of Space: Small Encyclopedia*, Ed. by R. A. Syunyaev (Sovetskaya Éntsiklopediya, Moscow, 1986), p. 360.
27. *Physical Encyclopedia*, Ed. by A. M. Prokhorov (Bol'shaya Ross. Éntsiklopediya, Moscow, 1998), Vol. 5, p. 409.
28. G. Michaud, Astrophys. J. **160**, 641 (1970).
29. S. N. Atutov and A. M. Shalagin, Pis'ma Astron. Zh. **14**, 664 (1988) [Sov. Astron. Lett. **14**, 284 (1988)].
30. K. A. Nasyrov and A. M. Shalagin, Astron. Astrophys. **268**, 201 (1993).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 739–744.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 860–866.
Original Russian Text Copyright © 2002 by Kompan, Novak, Kulik.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Spectra of Second-Order Raman Scattering in Porous Silicon

M. E. Kompan*, I. I. Novak, and V. B. Kulik

Ioffe Physicotechnical Institute, Russian Academy of Sciences, St. Petersburg, 194021 Russia
**e-mail: kompan@solid.ioffe.rssi.ru*

Received December 10, 2001

Abstract—Spectra of second-order Raman scattering in porous silicon are investigated. A band shift towards lower energies in second-order spectra is observed, as well as the correlation between the values of band shift in first- and second-order spectra. It is demonstrated that the observed effect cannot be interpreted using the conventional concepts of the mechanisms of scattering in microcrystalline samples. An interpretation of the revealed effect is suggested. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Porous silicon (por-Si) has been one of the "explosive" subjects in solid-state physics in the 1990s. In less than ten years after the publication by Canham [1], the number of publications devoted to this material exceeded by almost an order of magnitude the number of publications on the subject of high-temperature superconductivity (see the analysis of the dynamics of publications by Parkhutik [2]). Quite a few of these papers were devoted to the investigation of Raman scattering. The readily observed shift of the scattering band towards lower energies (compared with the position of this band in the spectrum of a crystalline material), the accepted interpretation of this effect [3], and the possibility of directly relating the observed shift to the size of small regions defined the high activity in precisely this direction.

In the majority of studies, the scattering of light in porous silicon was treated as a particular case of scattering by an aggregate of semiconductor objects of nanometer size. The adopted qualitative approach consisted in this case in the following: It was assumed that the smallness of the region in which the scattering occurs partly removes the prohibition of the light scattering by phonons with $k \neq 0$. Moreover, in contrast to a bulk material, the scattering became allowed in some interval of wave vectors (Δk) in the vicinity of the dispersion curve maximum ($k = 0$). The width of the interval Δk was estimated from the uncertainty relation $\Delta x \Delta k < 1$, where Δx is the linear dimension of a small spatial region. For regions several nanometers in size, this interval is of the order of $k_{max}/10$; even with the relative slope of the dispersion curve of silicon in the region of maximum ($k = 0$), this corresponded to the energy range from several to several tens of inverse centimeters. Therefore, the spectral shift and broadening were readily detected experimentally.

The quantitative interpretation of this phenomenon in almost all cases was based on the approach developed by Campbell and Faucheet [3]. In accordance with this approach, the vibration excited in a small spatially bounded element is treated as a wave packet of vibrations of different types allowed for propagation in an unbounded medium. Because waves of all types with $k \neq 0$ have a lower energy than that at the center of the Brillouin zone, this treatment leads to the same conclusions as the qualitative treatment; namely, the scattering band for nanometer-size objects must be broadened and shifted towards lower energies. The calculations in [3] were performed for two-, one-, and zero-dimensional objects, and the calculation results were compared with the results of experiments in films, thin filaments, and small spherical samples.

Note that both the above-described qualitative approach and the quantitative calculation in [3] treat the effect of the smallness of size on the process of scattering and actually proceed from the assumption of the invariability of the phonon spectrum of the material in nanometer-size objects. However, there exist reasons for which the intrinsic energies of phonon vibrations and their damping in nanometer-size objects may differ from the respective parameters for bulk materials. Some of the mechanisms leading to such differences were treated by Gorelik *et al.* [4] for nanoparticles of diamond and germanium.

In the case of second-order scattering, the small photon momentum is transferred to a pair of phonons; the momentum of each one of the latter may be other than zero. This removes the restriction $\Delta k = 0$ (common for first-order scattering) and the respective selection rules. As a result, phonons from the entire Brillouin zone may be involved in scattering, and the second-order spectrum largely reflects the entire phonon frequency spectrum of crystal (transferred to the doubled-frequency region). The clearly defined singularities in

second-order spectra correspond to the points of the dispersion curve characterized by a singularity in the density of states, $(dE/dn)(E) \longrightarrow 0$. Such singularities, in particular, show up in spectra at energy values equal to combinations of energies of phonon branches on the edges of the Brillouin zone. Results of theoretical analysis and experimental data for second-order scattering in crystalline silicon may be found in [5, 6].

Only several publications are available in which references are made to the observation of two-phonon scattering in porous silicon. Spectra of two-phonon scattering were first recorded for samples of porous silicon on silicon substrates (see, for example, [7, 8]). Because the effect of the substrate on the result remained unclear, the authors of those papers only established the fact of the presence of a band characteristic of two-phonon scattering; they could also point to a high intensity of this band compared with the corresponding spectra for crystalline silicon. Note that the nature of one specific mechanism of amplification of Raman scattering in porous silicon, defined by confinement in an electron subsystem, was treated by us in [9]. Shu-Lin Zhang *et al.* [10] recorded scattering spectra for free samples of porous silicon in a wide spectral region including the spectral region of second-order scattering. Without giving concrete definitions, Shu-Lin Zhang *et al.* [10] restricted themselves to making a remark that second-order spectra do not correspond to the existing theory of Raman effect in microcrystalline objects.

Indeed, as was mentioned above, first- and second-order scattering spectra are governed by different regularities. Therefore, a comparison of first- and second-order scattering spectra may produce independent information about the phonon spectrum of nanometer-size silicon objects. At present, however, it is not understood which special features must show up in second-order Raman scattering spectra for porous silicon and how informative such experiments may be.

2. EXPERIMENTAL PROCEDURE AND INVESTIGATED SAMPLES

We recorded scattering spectra using a RAMALOG-5 spectrometer with a triple monochromator. The spectral width of the slit in the experiments was 7 cm^{-1}. The spectra were excited by a continuous argon laser (λ = 488 nm). In order to identify the contribution made by weak signals against the noise background, computer accumulation of spectra was used. The investigations were performed at room temperature.

Different samples of porous silicon prepared by different production techniques were used in the experiments. Samples on a substrate of crystalline silicon were prepared from compensated polycrystalline silicon by the procedure of [11], i.e., by chemical etching with preliminary ion bombardment of initial silicon.

Most of the experiments were performed with free samples prepared from degenerate *n*-type silicon with the initial resistivity of 0.4–0.9 Ω cm and anodizing current density of 50 mA/cm^2, and illuminated by the focused light of a halogen lamp. The etching was performed in an HF : ethanol (1 : 1) solution. In the final phase of etching, the current density was raised by approximately an order of magnitude, which resulted in the separation of samples. The thus obtained samples were fairly large plates (with an area of up to 0.5 cm^2) with a clean bright surface. Under conditions of ultraviolet light excitation, the samples exhibited luminescence in the red-orange spectral region, and the thinnest samples were further characterized by an appreciable transparency in the same region.

It is known that porous silicon may be realized in the most diverse morphologies. The structure of the employed free samples of porous silicon was investigated using the atomic force microscopy. These investigations revealed, even within a single sample, the presence of regions with strongly differing characteristics of microrelief on an outwardly uniform surface. Therefore, in analyzing the results, we do not attempt to correlate the observed band shifts in the spectra with the presence of silicon clusters of certain size and shape. We only believe that, in our experiments in light scattering, some inhomogeneous combination of nanometer-size clusters was investigated for each sample. However, the use of samples of different types with different scattering spectra enables one to state with greater assurance that the observed results are typical.

3. EXPERIMENTAL RESULTS

The observed second-order scattering spectra for porous silicon were similar to second-order scattering spectra for crystalline silicon, but differed from the latter spectra by the shape and position of the bands.

The most pronounced feature of the second-order spectrum was the characteristic trapezoidal band in the 900–1000 cm^{-1} range. In all of the obtained second-order scattering spectra for porous silicon, the position of this band was shifted relative to its position in a crystalline material towards lower energies. Examples of second-order scattering spectra in the 900–1000 cm^{-1} range are given in Fig. 1. Given for comparison in the top part of the figure is a spectrum for crystalline silicon; the remaining spectra are for different samples of porous silicon. The interpretations of singular points, given on the top spectrum, are borrowed from [12].

A marked tendency is observed for the obtained spectra: in the case of minor shifts, the band shape is close to its shape for crystalline silicon; for relatively large shifts, the band is deformed and loses its characteristic trapezoidal shape. Note that, in all cases, the spectra for porous silicon samples are broadened and shifted to the lower energy region.

Also recorded for the same samples of porous silicon were first-order scattering spectra. The position of the band of fundamental scattering in first-order spectra (Fig. 2) was typical of porous silicon: the fundamental band (in the vicinity of 520 cm^{-1}) was also shifted to the lower energy region relative to the existing bands in the spectrum of crystalline material and broadened.

Scattering spectra for samples of porous and crystalline silicon in the region of about 300 cm^{-1} are given in Fig. 3. One can see in the figure that, in this case as well, the characteristic band in the spectrum of two-phonon absorption is shifted towards lower energies relative to the position of the respective band for a crystalline material.

A new experimental fact was provided by the observed correlation between the values of the band shift for first- and second-order scattering. In the case of samples investigated by us, the band shift in second-order spectra was observed for samples characterized by a greater shift in the first-order spectrum as well. The positions of peaks in the first-order spectra and singularities in the second-order spectra (given in Fig. 1) are compared for four samples in Fig. 4. Because the band shape in the demonstrated spectra develops from almost trapezoidal to bell-shaped, it does not appear possible to trace the positions of singular points. The values of the band middle and the spectral positions of the band edges are given in Fig. 4 for characterization of the bands in each spectrum (i.e., in the spectrum of a concrete sample). The positions of the edges were provided by the coordinates of the point with the maximal derivative of intensity along the coordinate; for the top graph, in which the band edges decrease linearly, this was the middle of the respective region. The same graph gives the estimated values of line broadening. Because the demonstrated band was initially a superposition of several components, the measure of broadening of the band components was provided by the minimal radius of curvature (in appropriate units) of details of the bands of a concrete spectrum. The data in Fig. 4 demonstrate clearly that the band shift in the 900–1000 cm^{-1} range in the second-order spectrum has the same sign as the shift of the band of fundamental scattering in the first-order spectrum and correlates with the latter in magnitude. These data are indicative of another tendency, namely, that the band shift is accompanied by the broadening of the components. This fact is well known for first-order spectra; however, for second-order scattering spectra, it was recorded for the first time.

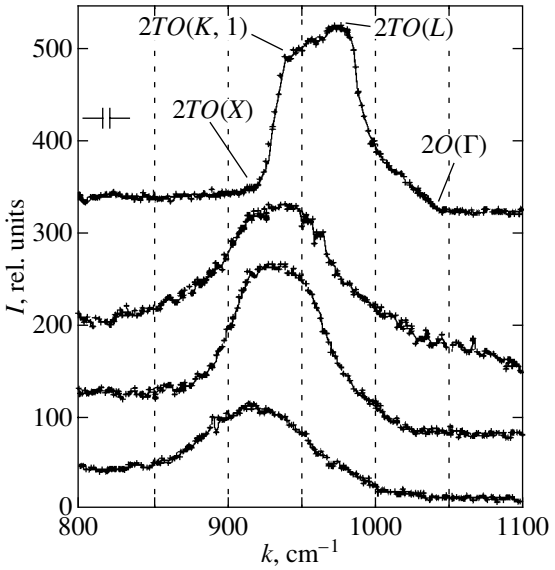

Fig. 1. The band position in the 900–1000 cm^{-1} range in second-order scattering spectra for crystalline silicon (top curve) and for different samples of porous silicon (I is the scattered light intensity, and k is the wave number). The band shift to the lower energy region is clearly discernible for the spectra of porous silicon. The interpretation of the positions of singular points on the top curve is borrowed from [12].

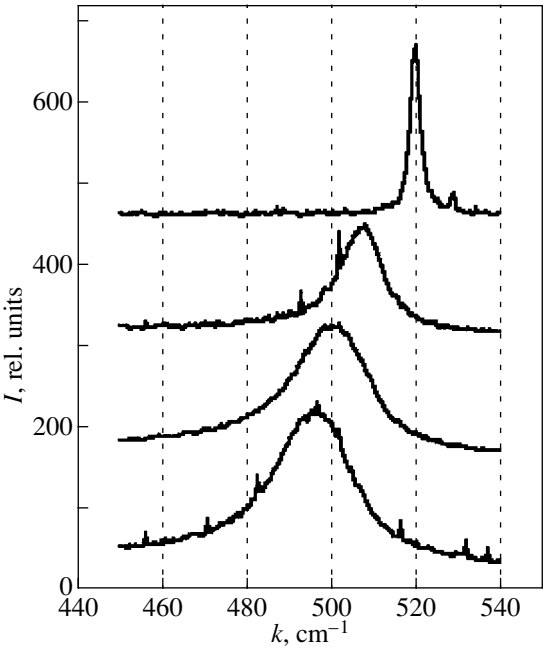

Fig. 2. The band position in the region of 440–540 cm^{-1} in the first-order scattering spectra for crystalline silicon (top curve) and for different samples of porous silicon. The band shift to lower energies is clearly discernible for the spectra of porous silicon.

4. DISCUSSION OF THE RESULTS

So, we have found that, in the case of porous silicon, the bands in second-order Raman spectra are shifted to the lower energy region compared with the respective bands for crystalline silicon, and this shift correlates with the analogous band shift and broadening for porous silicon in a first-order spectrum. From the standpoint of the adopted model, which attributes the known shift in the first-order spectrum to the effect of spatial confinement on scattering spectra, this result is unexpected and needs to be analyzed. One can further see

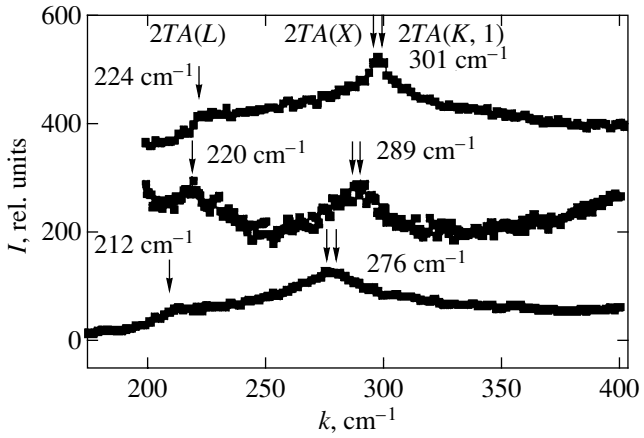

Fig. 3. Comparison of the position of singularities in second-order scattering spectra in the vicinity of 300 cm^{-1} for silicon (top curve) and porous silicon (two bottom curves). The interpretation of singular points is likewise borrowed from [12].

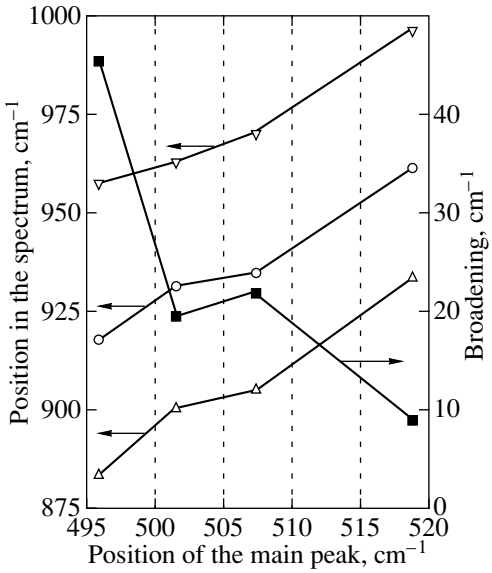

Fig. 4. The correlation between the positions of bands in first-order scattering spectra (plotted on the horizontal axis) and the positions of characteristic points in second-order scattering spectra for the same samples (plotted on the vertical axis). The hollow triangles and circles indicate the edges and the middle of the band for the spectrum of each sample, respectively; the solid squares indicate the estimated broadening of the components in cm^{-1} on the right-hand vertical scale.

fairly clearly that the band broadening increases, while the distance between the components does not vary or varies insignificantly compared with the observed shift.

We will, first of all, demonstrate that the concepts that are usually used for interpreting the data on Raman scattering in porous silicon [3] in no way help to explain the obtained results for second-order spectra. Of course, in the case of light scattering from nanometer objects, an uncertainty of the momentum of a scattering phonon must exist for any points of the phase diagram. However, if the phase shift in the 900–1000 cm^{-1} range was caused by the phonon momentum uncertainty (i.e., the mechanism suggested by Campbell and Faucheet [3]), the shift would have to have the sign opposite to that of the band shift in the first order, because the slopes of the dispersion branches of optical vibrations have opposite signs at the point $k = 0$ and on the boundaries of the Brillouin zone.

In fact, however, this is also incorrect. In the case of second-order scattering, the light scattering by phonons with $k \neq 0$ is not forbidden. Hence, it follows that the main mechanism providing for the band shift of first-order scattering, i.e., the violation of the selection rule $\Delta k = 0$, is not essential for the formation of spectra of second-order scattering. Nevertheless, the experimental results unambiguously demonstrate similar shifts and broadenings in both first- and second-order spectra for porous silicon.

Of other possible reasons for line shifts in scattering spectra, we must, no doubt, discuss the possible deformation of the material being investigated. The effect of pressure on the position of bands in second-order spectra may be assumed to be known. For silicon, the band in the vicinity of 1000 cm^{-1} must shift towards higher energies with increasing pressure with a coefficient of approximately 1 cm^{-1}/kbar. It is difficult to *a priori* assume a certain type of deformation of the material in the case of porous silicon. There is almost no question that the deformation must be nonuniform; however, the experimental spectrum is not simply broadened: the band is clearly shifted into the lower energy region. In view of the known correlations [13], this must correspond to the tensile stresses in the material. However, no such conclusion may be made independently of the data on the position of other bands.

The effect of deformation on the spectral position of the band in the first-order spectrum for silicon is also known. This band (520 cm^{-1}) must also shift into the lower frequency region under tensile stresses [13]. Thereby, the assumption of the existence of tensile deformation would automatically explain the correlation between the shifts of the above-identified bands in first- and second-order spectra.

However, the band in the vicinity of 300 cm^{-1} has the opposite sign of deformation constant ($dk/dP \approx -0.4$ cm^{-1} $kbar^{-1}$) and would have to shift towards higher energies in the case of tensile strain. The experimental results demonstrate quite the opposite. Both bands in the second-order spectrum, in the vicinity of 300 cm^{-1} and 900–1000 cm^{-1}, shift into the lower energy region. It is by virtue of the identical sign of shift for these two bands that we are forced to eliminate the deformation of the material from the possible reasons for the observed effect.

On the other hand, it is the coordinated band shift, especially, with the conservation of the band width in the vicinity of 900–1000 cm^{-1}, that gives a clue to understanding the reason for this phenomenon. We will take into account the fact that a second-order spectrum is formed in the case of superposition of scattering from all phonon branches, and the position of singularities in the spectrum is reflective of the energy of singular points of the phonon diagram. Then, the observed general shift towards lower energies must be understood as a consequence of the decrease (on the average) of the elasticity of the material for vibrations in the respective frequency region. However, the possible general reason for the reduction of elastic constants of porous material is almost evident: strictly speaking, this reason lies in the obvious decrease in the elasticity of the porous material. The elastic forces returning an element of the medium under vibrations to the initial position are made up of the forces counteracting the compression on one side of this element and the forces counteracting the tension on the opposite side of the element. In the case of a porous medium, a major part of the material lies on the surface of pores (or clusters), and, in the case of vibrations, the restoring force will act on the surface layers on only one side; this must unambiguously lead to the reduction of the natural vibration frequencies. A qualitative interpretation of this may also be that the surface modes start being admixed to the volume oscillation in a porous medium. More rigorous assertions are impossible in this situation because of the inhomogeneity of the porous structure; however, by and large, the suggested reason for the observed uniform shift of scattering bands to the lower energy region appears valid.

But now the results for second-order scattering, in turn, must be fitted to the known data on one-phonon Raman scattering. The problem is that the variations of the phonon frequencies, which, according to the assumption made, are responsible for the band shifts in the second order, must have brought about a variation of the entire dispersion diagram of lattice vibrations; this, in turn, inevitably must have an effect on the first-order scattering spectra. However, as was already mentioned, the observed band shift in the first-order spectrum was already unambiguously interpreted as a manifestation of phonon confinement [3]. The numerical values of the observed band shifts in the first and second orders are comparable in magnitude, and this prevents one from manipulating the concepts of the smallness of one of the contributions or of their compensation.

Nevertheless, the formulated problem finds its solution in view of the pattern of vibrations at the point $k = 0$ of the phonon diagram. At this point, the optical vibrations are the vibrations ($d\omega/dk = 0$) of one sublattice relative to another. In the case of vibrations of this type, the macroscopic environment for an individual element of the medium is of little significance; the restoring force for each atom is defined by the interaction with the nearest neighbors. Hence it follows that the effect of porosity of the medium must not affect considerably the natural frequencies in the vicinity of the dispersion curve maximum.

The vibrations at points on the edges of the Brillouin zone are not strictly optical. In this case, the vibrations partly include the combined vibrations of the sublattices as well, i.e., the vibrations of the material as a single whole. In turn, vibrations of this type must be sensitive to the presence of internal boundaries (surfaces) of the material, as was pointed out above. The natural frequencies of vibrations of this type must be lower in the vicinity of the surface.

In total, it is this that brings about the observed effects. In the case of first-order scattering, scattering from the region in the vicinity of extremum ($k = 0$) is observed. For vibrations of this type, the porosity of the material is of small significance, and the natural frequencies in the region in the vicinity of $k = 0$ remain unchanged. On the other hand, the very possibility of one-phonon scattering of light is associated with the condition $k = 0$, and a partial violation of the prohibition of scattering by phonons with $k \neq 0$ leads to the shift and broadening of the fundamental band of scattering in the spectrum under observation. The singularities in second-order scattering spectra correspond to the frequencies of the phonon branches on the boundaries of the Brillouin zone. In this case, there is no effect of the prohibition of transitions with $k \neq 0$; on the other hand, the phonon frequencies of the material prove to be sensitive to the presence of pores, and it is another mechanism that provides for the band shift towards lower energies. By and large, it turns out that the mechanisms of band shifts in first- and second-order spectra are different; however, in both cases, the mechanisms are “triggered” because the material being investigated is not a continuous medium but presents a conglomerate of clusters or a porous medium. This general reason, while acting via different mechanisms, leads to the presence of correlation in band shifts, which was revealed by us.

5. CONCLUSIONS

Thus, as a result of a series of experiments, it has been demonstrated that the bands in second-order spectra of light scattering are shifted towards lower energies compared with second-order spectra for a crystalline material. These results agree with the published results of separate experiments by other researchers. An explanation of the revealed effect has been suggested. In our experiments, data on the variation of the spectrum of phonon frequencies for porous silicon have been obtained for the first time.

We have also found a clearly defined correlation between the band shifts in first- and second-order scattering spectra and provided an explanation of the hier-

archy of mechanisms responsible for band shifts in spectra of different orders.

ACKNOWLEDGMENTS

We are grateful to I.A. Merkulov for discussions of the experimental results.

REFERENCES

1. L. T. Canham, Appl. Phys. Lett. **57**, 1046 (1990).
2. V. Parkhutik, J. Porous Mater. **7**, 363 (2000).
3. I. H. Campbell and P. M. Faucheet, Solid State Commun. **58**, 739 (1986).
4. V. S. Gorelik, A. V. Igo, and S. N. Mikov, Zh. Éksp. Teor. Fiz. **109**, 2141 (1996) [JETP **82**, 1154 (1996)].
5. K. Uchinikura, T. Sekine, and E. Matsuura, J. Phys. Chem. Solids **35**, 171 (1974).
6. W. Windl, P. Pavone, K. Karch, *et al.*, Phys. Rev. B **48**, 3164 (1993).
7. H. Munder, C. Andrzejak, M. G. Berger, *et al.*, Thin Solid Films **221**, 27 (1992).
8. I. Gregora, B. Champagnon, and A. Halimaoui, J. Appl. Phys. **75**, 3034 (1994).
9. M. E. Kompan, I. I. Novak, V. B. Kulik, and N. A. Kamakova, Fiz. Tverd. Tela (St. Petersburg) **41**, 1320 (1999) [Phys. Solid State **41**, 1207 (1999)].
10. Shu-Lin Zhang, Xin Wang, Kouk-san Ho, *et al.*, J. Appl. Phys. **76**, 3016 (1994).
11. V. I. Beklemishev, V. M. Gontar', V. V. Levenets, *et al.*, Élektron. Prom-st, No. 2, 36 (1994).
12. B. A. Weinstein, G. J. Piemarini, J. D. Barnett, and S. Block, Science **176**, 284 (1972).
13. *Light Scattering in Solids*, Ed. by M. Cardona and G. Guntherodt (Springer-Verlag, Berlin, 1984; Mir, Moscow, 1986), Vol. IV.

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 745–750.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 867–874.
Original Russian Text Copyright © 2002 by Kraĭnov, Smirnov.

PLASMA, GASES

Charge Composition of a Cluster Plasma upon Irradiation of Large Atomic Clusters by the Field of a Superatomic Femtosecond Laser Pulse

V. P. Kraĭnov[a, *] and M. B. Smirnov[b]

[a]*Moscow Physicotechnical Institute, Dolgoprudnyĭ, Moscow oblast, 141700 Russia*
[b]*Max-Born Institute, 12489, Berlin, Germany*
**e-mail: krainov@online.ru*

Received December 22, 2001

Abstract—A theory is developed for calculating the charge composition of a cluster plasma produced upon irradiation of large atomic clusters by the field of a superatomic femtosecond laser pulse. The theory is based on the overbarrier process of a successive multiple internal ionization of atomic ions inside a cluster accompanied by the external field ionization. Collision ionization is also taken into account in the calculations. The theory is illustrated by the example of a cluster consisting of 10^6 xenon atoms irradiated by a 50-fs laser pulse with a peak intensity of 2×10^{18} W/cm^2. In this case, the Xe^{26+} ions dominate. The amounts of atomic xenon ions with multiplicity up to 31 are calculated. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interaction of superpower laser pulses with large clusters consisting of noble gas atoms substantially differs from the interaction of such pulses with isolated atoms. Upon multiple ionization of atoms in the cluster, a strong internal field is produced, which itself can perform a further internal ionization of the atomic ions.

This problem is rather complicated because the ionized cluster is simultaneously expanded (during the laser pulse) and the electronic component of the cluster is heated up to several keV. In addition, external ionization occurs, i.e., the escape of the produced photoelectrons from the cluster. As a result, matter at the focus of the laser beam immediately after the laser pulse termination represents a rather homogeneous plasma consisting of free electrons and multiply charged atomic ions.

The calculation of the spectral distribution of produced atomic ions over their charges is an important theoretical problem. This problem is complicated by the fact that atomic ions are "stripped" at the leading edge of the laser pulse not only by the laser-pulse field but also due to collisions of atomic ions with fast electrons inside the cluster. At the same time, a standard statistical approach based on the Saha distribution [1] cannot be applied because the laser-pulse duration is several tens of femtoseconds, and slow recombination processes have no time to proceed during the laser pulse.

In the initial part of the leading edge of the laser pulse, a successive overbarrier ionization of atoms [2] and then of atomic ions by the laser field commonly takes place. In this case, the collision ionization is rather weak because the kinetic energy of produced photoelectrons is small. This ionization is observed only near the maximum of the laser pulse.

The external ionization of the cluster is mainly field (cold) ionization. The role of thermal ionization according to the Richardson–Dashman law is negligible because of a great positive charge of the ionized cluster, which prevents the thermal evaporation of electrons from the cluster surface.

The aim of our paper is to determine the charge state of atomic ions in the cluster plasma upon irradiating clusters by the field of a superatomic femtosecond laser pulse.

We considered xenon clusters as typical objects. They are formed during the adiabatic flowing out of gaseous xenon compressed preliminarily to a pressure of several tens of atmospheres through a nozzle into vacuum, which is accompanied by a temperature drop. The size of the clusters increases with pressure. We assume that matter in the cluster is in a liquid state. Thus, the density of liquid xenon is 3.52 g/cm^3 (the temperature of the transition from gaseous to liquid state is −107.1°C), which significantly differs from the density of solid xenone: 2.7 g/cm^3 (the solidification temperature is −111.9°C). The noble gas atoms in clusters are attracted to each other by van der Waals forces.

Therefore, the radius of a liquid xenon cluster (which is assumed spherical, as confirmed by experiments on Rayleigh scattering of light by clusters) containing a million atoms equals 245 Å. The skin depth exceeds this value, so that we can consider that the electromagnetic field of the laser pulse freely propagates through an individual cluster. Of course, when there are many such clusters in the focus of the laser beam, laser

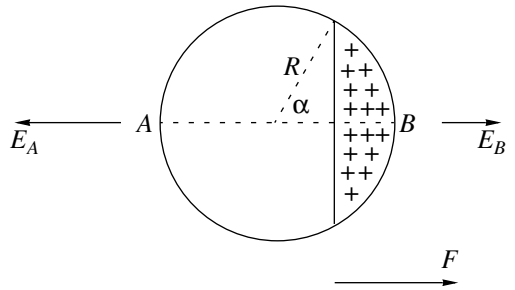

Fig. 1. Charged and neutral regions of the ionized cluster irradiated by a laser pulse.

radiation is strongly absorbed, as was observed experimentally [3].

During the multiple ionization, the concentration of free electrons inside the cluster plasma becomes rather high. The penetration of laser radiation inside the plasma (already at the trailing edge of the laser pulse, when the clusters virtually disappear and the cluster plasma becomes homogeneous) ceases because the laser radiation frequency becomes lower than the plasma frequency

$$\sqrt{\frac{4\pi N_e e^2}{m_e}}.$$

Here, N_e is the concentration of free electrons in the plasma.

The above discussion demonstrates a variety of processes occurring upon the interaction of superpower ultrashort laser pulses with large atomic clusters. Therefore, to analyze these processes, models are required which would allow the determination of the charge composition of atomic ions at the leading edge of the laser pulse and in the cluster plasma with good accuracy (before the development of recombination processes). This is important for a further study of the line electromagnetic emission of multiply charged atomic ions in the far X-ray range [4].

2. MODEL OF OVERBARRIER MULTIPLE INTERNAL AND EXTERNAL CLUSTER IONIZATION

Our approach to the internal multiple ionization of atoms in large van der Waals clusters by a superstrong field of a laser field is based on the Bethe model [5] of overbarrier ionization. Because we often used this model in our previous papers [6–8], we will describe it here only briefly. To produce an atomic ion with the charge Z and ionization potential E_Z in a cluster at some instant of time t, an electric field of the strength

$$F(t) = \frac{E_Z^2}{4Z} \tag{1}$$

is required. Hereafter, we use the atomic system of units, in which the electron charge and mass and the Planck constant are equal to unity.

However, the electric field inside the ionized cluster does not coincide with the external field of a laser pulse. Free electrons inside the cluster (their total number is ZN, where N is the number of atoms in the cluster) are rather rapidly displaced by the laser field oppositely to the field direction (Fig. 1). After a half-period, they are rapidly displaced in the opposite direction due to the absence of inertia. It is assumed in [9] that the spherical shape of the electronic subsystem is conserved during such displacements. This could be observed in the case of a strong surface tension existing in the system. However, there is no reason to assume this in our case. In our model, electrons are simply displaced oppositely to the laser field direction.

Therefore, the ionized cluster consists of two regions: in its neutral part the electrons and atomic ions are located, and in the charge part of the cluster only atomic ions are found (Fig. 1). Let us assume that the interface between these regions is flat. Of course, this is a certain approximation, because the interface surface is in fact bent towards the charged region, its convexity being determined by the condition that the tangential component of the electric field would be zero over the entire surface. This condition precludes the movement of free electrons. However, our aim is to determine an additional electric field produced by the charged region at the remotest point of the cluster (point A in Fig. 1), which will only slightly change upon small bending of the interface.

We could assume that the electrons are distributed not as shown in Fig. 1 but over the entire cluster, however, nonuniformly, their number in the left part of the cluster being greater than that in the right part, as in the case of volume plasma oscillations. In this case, the electrically neutral part of the cluster would be absent altogether. However, such a variant is less probable than that described above because plasma always tends to become electrically neutral. In a neutral cluster, surface plasma Mie oscillations would be excited [10].

The electric field strength produced at the point A in the charged region of the cluster can be readily calculated:

$$E_A = \frac{NZ}{R_2}\left(1 - 3\cos^2\frac{\alpha}{2} + 2\cos^3\frac{\alpha}{2}\right). \tag{2}$$

Here, R is the cluster radius, and the angle α is shown in Fig. 1. The condition

$$F(t) = E_A \tag{3}$$

means that the force with which an electron is ejected by the laser field outside is equal to the force with which this electron is attracted by the positively charged region of the ionized cluster.

The electric field strength at the extreme right point B (Fig. 1) produced by the charged region of the cluster can be readily calculated using electrostatic laws:

$$E_B = \frac{NZ}{R^2}\left(3 - 2\sin\frac{\alpha}{2}\right)\sin^2\frac{\alpha}{2}. \quad (4)$$

This field should be added to the external field $F(t)$ because it enhances the internal ionization of atomic ions in the cluster in accordance with the well-known "ignition" model [11].

However, we should bear in mind that, first, this field does not act in the neutral region of the cluster and, second, it gives only the maximum value of the total field in the charged region, whereas the total field will have other values at other points of the charged region. The first factor is most important. To take this factor into account, we propose to reduce the "ignition" field (4) by multiplying it by the ratio of the volume V of the charged region to the cluster volume $4\pi R^3/3$. The volume of the charged region is

$$V = \frac{1}{3}\pi R^3(2 - 3\cos\alpha + \cos^3\alpha). \quad (5)$$

The charge of this region (the charge of the ionized cluster) is

$$Q = \frac{NZ}{4}(2 - 3\cos\alpha + \cos^3\alpha). \quad (6)$$

By excluding the angle α from Eqs. (2), (3), and (6), we can obtain the universal relation between the external field strength $F(t)$ (in units of ZN/R^2) and the charge Q of the ionized cluster (in units of NZ). This relation shown in Fig. 2 allows one to calculate the degree of external ionization of any cluster for a given strength of the electric field of a laser pulse.

Therefore, the effective strength of the ignition field enhancing the internal ionization has the form

$$F_{\text{eff}} = \frac{NZ}{4R^2}(2 - 3\cos\alpha + \cos^3\alpha) \times\left(3 - 2\sin\frac{\alpha}{2}\right)\sin^2\frac{\alpha}{2}. \quad (7)$$

By adding this field to the external field $F(t)$, we obtain a real field producing the internal ionization of atomic ions in the ionized cluster, which we should equate to $E_Z^2/4Z$, according to the Bethe condition, to obtain finally

$$\frac{E_Z^2}{4Z} = \frac{NZ}{4R^2}(2 - 3\cos\alpha + \cos^3\alpha) \times\left(3 - 2\sin\frac{\alpha}{2}\right)\sin^2\frac{\alpha}{2} + \frac{NZ}{R^2}\left(1 - 3\cos^2\frac{\alpha}{2} + 2\cos^3\frac{\alpha}{2}\right). \quad (8)$$

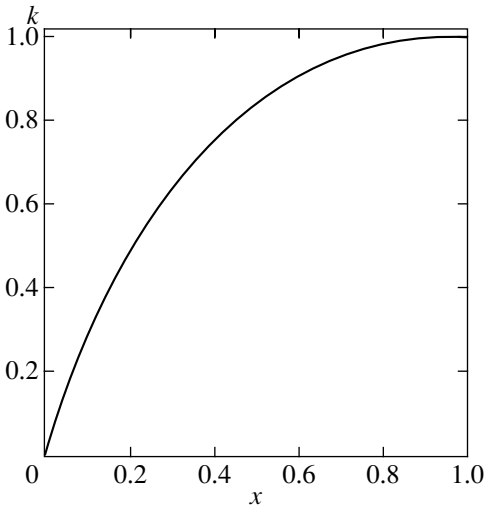

Fig. 2. Universal dependence of the fraction $k = Q/NZ$ of electrons escaping from a cluster on the laser-pulse field strength $x = FR^2/NZ$.

This equation allows us to calculate the angle α from the known ionization potential E_Z of the given atomic ion and its current radius $R(t)$, which increases with time due to the Coulomb explosion of the ionized cluster.

Knowing the angle α, we can find the charge Q of the ionized cluster from expression (6). Then, using relations (2) and (3), we can calculate the field strength $F(t)$,

$$F(t) = \frac{NZ}{R^2}\left(1 - 3\cos^2\frac{\alpha}{2} + 2\cos^3\frac{\alpha}{2}\right), \quad (9)$$

and the instant of time t at which the internal ionization occurred with the formation of atomic ions with the charge Z. For this purpose, we will use the relation

$$F(t) = F_0\exp\left(-\frac{t^2}{\tau^2}\right) \quad (10)$$

with the known values of the laser-pulse field amplitude F_0 and the laser pulse duration τ.

Now, we calculate the increase in the radius of the ionized cluster caused by its Coulomb expansion by using Newton's law for the movement of an atomic ion on the cluster surface

$$\frac{d^2R}{dt^2} = \frac{Q(t)Z(t)}{R^2(t)}. \quad (11)$$

Table 1 presents the results of calculations performed for a xenon cluster containing $N = 10^6$ atoms interacting with a 50-fs (FWHM) Gaussian laser pulse with a peak intensity of 2×10^{18} W/cm^2. These values

Table 1. Dynamics of the internal and external ionization of a xenon cluster containing $N = 10^6$ atoms irradiated by a 50-fs laser pulse with a peak intensity of 2×10^{18} W/cm^2

Z	E_Z, eV	$-t$, fs	F, au.	Q, 10^6	R, au.
1, $5p^6$	12.1	100	0.028	0.022	464
2, $5p^5$	21.1	96	0.044	0.036	464
3, $5p^4$	32.1	92	0.026	0.054	464
4, $5p^3$	46.7	87	0.108	0.086	464
5, $5p^2$	59.7	85	0.135	0.108	464
6, $5p^1$	71.8	83	0.162	0.129	464
7, $5s^2$	92.1	79	0.223	0.176	465
8, $5s^1$	106	78	0.258	0.206	465
9, $4d^{10}$	171	69	0.535	0.413	473
10, $4d^9$	202	66	0.653	0.518	478
11, $4d^8$	233	64	0.777	0.626	482
12, $4d^7$	263	62	0.894	0.731	487
13, $4d^6$	294	60	1.023	0.846	494
14, $4d^5$	325	58	1.143	0.969	502
15, $4d^4$	358	56	1.276	1.110	512
16, $4d^3$	390	55	1.406	1.270	518
17, $4d^2$	421	53	1.526	1.400	531
18, $4d^1$	452	52	1.641	1.570	538
19, $4p^6$	549	47	2.191	2.11	590
20, $4p^5$	583	46	2.285	2.59	603
21, $4p^4$	618	45	2.429	2.86	617
22, $4p^3$	651	44	2.548	3.13	632
23, $4p^2$	701	42	2.782	3.55	667
24, $4p^1$	737	41	2.915	4.08	687
25, $4s^2$	819	38	3.374	4.90	758
26, $4s^1$	897	35	3.770	6.40	846
26, $4s^1$	897	32	4.200	8.43	954
26, $4s^1$	897	29	4.650	10.9	1081
26, $4s^1$	897	26	5.100	14.0	1228
26, $4s^1$	897	23	5.550	17.3	1395
26, $4s^1$	897	20	5.940	20.7	1580
26, $4s^1$	897	17	6.330	23.7	1782
26, $4s^1$	897	14	6.660	25.6	2000
26, $4s^1$	897	11	6.950	26.0	2230
26, $4s^1$	897	7	7.230	26.0	2470
26, $4s^1$	897	4	7.360	26.0	2720
26, $4s^1$	897	0	7.430	26.0	3060

of the parameters are typical for experiments [12]. In this case, the distribution of the envelope of the laser field strength has the form

$$F(t) = 7.43\exp\left(-\frac{t^2}{\tau^2}\right) \text{[au]}. \quad (12)$$

Here, $\tau = 42.47$ fs.

One can see from Table 1 that the external ionization first lags strongly behind the internal ionization, and the charge of the ionized cluster slowly increases. However, already at the leading edge of the laser pulse, there comes the instant of time (–11 fs) when all the electrons that have escaped from xenon atoms are removed from the cluster, so that the latter contains only positively charged xenon ions (mainly with the charge $Z = 26$).

The cluster size already strongly increases at the leading edge of the laser pulse due to the Coulomb expansion. Thus, at the instant of time $t = 0$, which corresponds to the peak value of the laser-pulse intensity, the cluster diameter is more than six times larger than its initial value. The typical distance to an adjacent cluster is 10–20 cluster diameters. Therefore, clusters already disappear during the laser pulse, and the cluster plasma becomes virtually spatially homogeneous.

3. TUNNEL AND COLLISION IONIZATION OF ATOMIC IONS IN A CLUSTER

We described the internal ionization of atoms and atomic ions in the previous section within the framework of the overbarrier ionization. For the given value of the peak intensity of a laser pulse, a further overbarrier ionization is impossible. However, other types of ionization are possible. Consider first the tunnel ionization of xenon ions Xe^{26+}.

The absolute probability of tunnel ionization can be calculated analytically using expressions from [13]:

$$W = t\left(\frac{e}{\pi}\right)^{3/2}\frac{\sqrt{3}(2E_Z)^{9/4}}{Z^{5/2}}\left(\frac{16eE_Z^2}{ZF}\right)^{2Z/\sqrt{2E_Z}-3/2} \times \exp\left(-\frac{2(2E_Z)^{3/2}}{3F}\right). \quad (13)$$

Here, t is the time of action of the given electric field strength F.

By using Eq. (13), one should bear in mind that tunnel ionization appears in the total electric field, which includes an external field and the Coulomb field of an ionized cluster (the ignition model that was discussed in the previous section). The maximum field is achieved at point B (Fig. 1). Table 2 presents the values of the total field at different times. One can see that the field first increases, reaches the maximum equal to $F_{\text{eff}} = 14.43$ au at $t = -23$ fs, and then decreases. Because the probability of tunnel ionization exponentially depends on the electric field strength, it is this instant of time and the time in its vicinity that are the most important. By substituting the values $F_{\text{eff}} = 14.43$ au, $Z = 26$, $E_Z = 33.0$ au, and $t = 10$ fs into (13), we obtain $W = 10^{-8}$.

Such a small ionization probability means that tunnel ionization does not result in the production of atomic xenon ions with multiplicity exceeding 26.

Another mechanism is based on inelastic collisions of fast electrons with atomic ions inside a cluster,

resulting in the knocking out of an electron from the atomic ion. The cross section for collision ionization is described, with good accuracy, by the Lotz formula [14]

$$\sigma = 2.17 f_i \frac{\ln(E_e/E_Z)}{E_e E_Z} \text{ au.} \quad (14)$$

Here, E_e is the kinetic energy of a colliding electron (in au), E_Z is the bound energy of a multiply charged atomic ion (in au), and f_i is the number of electrons in the valence shell of the given multiply charged ion.

The amplitude of oscillations of an electron in the laser field at the leading edge of a laser pulse at the time instant $t = -35$ fs is

$$a = \frac{F}{\omega^2} = 1160 \text{ au},$$

$$E_e = \frac{F^2}{4\omega^2} = 1090 \text{ au}.$$

By substituting $f_i = 10$ and $E_{27} = 50.9$ au for the first electron knocked out from the 3*d* shell, we obtain from (14) $\sigma = 0.012$ au. The probability of an ionizing collision per unit time is determined by the expression

$$w = N_e \sigma \sqrt{2E_e}. \quad (15)$$

Here, N_e is the electron concentration inside the cluster, which is equal to

$$N_e = \frac{NZ - Q}{4\pi R^3/3}. \quad (16)$$

By substituting the values presented in Table 1 for the given instant of time, we find $w = 4.3 \times 10^{-4}$. By multiplying this value by the time $t = 10$ fs, during which, according to data presented in Table 1, free electrons still remain inside the cluster, and by the number of ions inside the cluster (i.e., by $N = 10^6$), we obtain the number of ions with the charge 27, $N_{27} = 160000$.

The numbers of atomic ions with charges 28, 29, 30, and 31 are calculated similarly. They rapidly decrease because each successive atomic ion is produced due to ionization of the previous ion (the probability of simultaneous multiple ionization of an atomic ion with the charge $Z = 26$ is negligible). The ionization potentials of multiply charged xenon ions were taken from [15]. The results are presented in Table 3.

Therefore, our approach allows us to calculate the charge composition of atomic ions upon irradiation of large clusters by the field of a superatomic femtosecond laser pulse. The charge composition changes at the leading edge of the laser pulse and becomes fixed near the pulse maximum. After the pulse termination, only free expansion of the homogeneous cluster takes place (for several nanoseconds).

Table 2. Maximum strength of the effective field at different instants of time at the leading edge of a laser pulse

$-t$, fs	F_{eff}, au.
35	12.70
32	13.46
29	13.98
26	14.38
23	14.43
20	14.25
17	13.79
14	13.03
11	12.18
7	11.49
4	10.87

Table 3. Charge composition of atomic ions in a xenon cluster containing $N = 10^6$ atoms after irradiation by a superstrong ultrashort laser pulse

Z	E_Z, eV	N_Z
26	897	816800
27	1385	160000
28	1491	20800
29	1587	2200
30	1684	190
31	1781	10

4. CONCLUSIONS

The interaction of large xenon clusters with superatomic femtosecond laser pulses was experimentally studied in [3, 12, 16, 17]. In these papers, X-ray radiation of multiply charged xenon ions caused by transitions between the discrete levels of these ions was analyzed. The energy of these transitions give information on the charge composition of the cluster plasma.

The authors of [16] observed the following X-ray transitions: Xe^{26+}: $3d^9 4f \longrightarrow 3d^{10}$; Xe^{27+}: $3d^8 4f \longrightarrow 3d^9$; Xe^{28+}: $3d^7 4f \longrightarrow 3d^8$, and Xe^{29+}: $3d^6 4f \longrightarrow 3d^7$. The maximum yield of photons was observed for Xe^{26+} ions. The yield of ions decreased with increasing degree of ionization. The average number of atoms in a cluster was about 3×10^6. This completely agrees with the results of our calculations presented in Table 3. Rhodes *et al.* [18] explain high charge states of xenon ions by the coherent motion of electron bunches inside the ionized cluster. In this case, the probability of electron knocking out from atomic ions increases. However, it is not clear in this model what confines the electron bunches together inside a small volume, preventing them from coming apart due to the Coulomb repulsion. Another explanation proposed by the group

of Ditmire [19] is that resonance excitation of surface plasma electron Mie oscillations occurs, resulting in a strong heating of electrons in the cluster. However, in the case of multiply charged ions, as was pointed out in [16], the free electron density is so high that the Mie frequency greatly exceeds the laser radiation frequency, so that the resonance cannot exist even upon a significant expansion of the cluster.

A recent paper [17] by Rhodes' group devoted to the analysis of X-ray transitions in the L shell of Xe^{27+} and Xe^{28+} ions is based on the model [18].

Finally, the authors of [3, 12] observed Xe ions with charges from 26 to 29, as well as with lower charges. The aim of our further calculations is to explain the intensity of X-ray lines in the spectra of such ions.

ACKNOWLEDGMENTS

The authors thank W. Becker and P. Nickles for useful advice. One of the authors (M. S.) thanks the Humboldt Foundation for financial support. This work was supported by the Russian Foundation for Basic Research (project no. 01-02-16056) and by the US Civilian Research and Development Foundation (CRDF grant MO-011-0).

REFERENCES

1. B. M. Smirnov, *Physics of Ionized Gases* (Wiley, New York, 2001).
2. N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms* (Springer-Verlag, Berlin, 2000).
3. M. Schnürer, S. Ter-Avetisyan, H. Stiel, *et al.*, Eur. Phys. J. D **14**, 331 (2001).
4. T. Auguste, P. D'Oliveira, S. Hulin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 54 (2000) [JETP Lett. **72**, 38 (2000)].
5. H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms* (Academic, New York, 1957; Fizmatgiz, Moscow, 1960).
6. V. P. Kraĭnov and M. B. Smirnov, Zh. Éksp. Teor. Fiz. **120**, 555 (2001) [JETP **93**, 485 (2001)].
7. V. P. Kraĭnov and M. B. Smirnov, Zh. Éksp. Teor. Fiz. **119**, 719 (2001) [JETP **92**, 626 (2001)].
8. V. P. Krainov and A. S. Roshchupkin, J. Phys. B **34**, L297 (2001).
9. P. B. Parks, T. E. Cowan, R. B. Stephens, and E. M. Campbell, Phys. Rev. A **63**, 063203 (2001).
10. V. P. Kraĭnov and M. B. Smirnov, Usp. Fiz. Nauk **170**, 969 (2000).
11. C. Rose-Petruck, K. J. Schafer, K. R. Wilson, and C. P. J. Barty, Phys. Rev. A **55**, 1182 (1997).
12. S. Ter-Avetisyan, M. Schnürer, H. Stiel, *et al.*, Phys. Rev. E **64**, 036404 (2001).
13. M. V. Ammosov, N. B. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1986)].
14. W. Z. Lotz, Z. Phys. **216**, 241 (1968).
15. R. D. Cowan, *The Theory of Atomic Structure and Spectra* (Univ. of California Press, Berkeley, 1981).
16. H. Honda, E. Miura, K. Katsura, *et al.*, Phys. Rev. A **61**, 023201 (2000).
17. W. A. Schroeder, T. R. Nelson, A. B. Borisov, *et al.*, J. Phys. B **34**, 297 (2001).
18. K. Boyer, B. D. Thomson, A. McPherson, and C. K. Rhodes, J. Phys. B **27**, 4373 (1994).
19. T. Ditmire, T. Donnelly, A. M. Rubenchik, *et al.*, Phys. Rev. A **53**, 3379 (1996).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 751–758.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 875–883.
Original Russian Text Copyright © 2002 by Meilikhov, Farzetdinova.

SOLIDS
Structure

Ground States for Lattices of Ferromagnetic Granules with Dipolar Magnetic Interaction

E. Z. Meilikhov* and R. M. Farzetdinova

Institute of Molecular Physics, Russian Research Centre Kurchatov Institute, Moscow, 123182 Russia
**e-mail: meilikhov@imp.kiae.ru*

Received October 8, 2001

Abstract—The magnetic phase diagrams of 2D and 3D regular lattices formed by nonspherical single-domain ferromagnetic granules featuring a dipolar magnetic interaction are studied. The energy of a magnetic state of such systems is calculated using an approximate expression for the pair interaction of nonspherical granules. The character of the magnetic ground state of the system is determined by three geometric parameters: (i) the eccentricity of granules; (ii) the ratio of periods of the rectangular (2D) or tetragonal (3D) lattice; and (iii) the ratio of a lattice period to a granule size. In contrast to the case of lattices formed by point (or spherical) magnetic moments, in which the ground state is always antiferromagnetic or frustrated (for triangular lattices), the ground state of a 2D lattice composed of nonspherical granules can be ferromagnetic. The magnetic phase diagrams of the systems studied are constructed in the space of the above geometric parameters. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Can ferromagnetism exist at zero temperature in a system of magnetic moments featuring dipole–dipole interaction? An answer to this question, requiring exact calculation of the energy of a long-range magnetic interaction involving all dipoles in the system, is not easy to obtain even for a system composed of point magnetic moments. For a pair of parallel point magnetic moments μ at a distance of r from each other, the magnetic interaction energy is given by the formula

$$w = \frac{\mu^2}{r^3}(1 - 3\cos^2\theta),$$

where θ is the angle between the direction of moments and the line connecting the two points. The sign of this energy varies depending on the angle θ: for $|\theta| < \theta_0$ or $|\pi - \theta| < \theta_0$, where $\theta_0 = \arccos(1/\sqrt{3}) \approx 55°$, the interaction retains parallelism of the magnetic moments (i.e., favors ferromagnetic ordering), otherwise the antiparallel (antiferromagnetic) dipole configuration becomes energetically favorable. The state of a system containing a large number of dipoles is determined by the competition of these trends. Calculations show that the ground magnetic state of one-dimensional chains [1], two-dimensional (2D) square [2] and rectangular [3] lattices, and also 3D cubic lattices [4, 5] composed of point magnetic dipoles is not ferromagnetic: the dipolar (non-exchange) ferromagnetism in such systems is impossible.

This conclusion is also valid for the systems composed of homogeneously magnetized (single-domain) spherical granules, because the field of each granule coincides with the field of an equivalent point dipole placed at the center and the dipolar magnetic interaction energy of a pair of such granules coincides with that of two equivalent point dipoles [6].

A different situation is observed for homogeneously magnetized granules possessing a nonspherical shape. Below, we restrict the consideration to granules having the shape of an elongated ellipsoid of revolution with a magnetic moment orientation fixed (due to a significantly large shape anisotropy) in the direction of the major semiaxis.[1]

If the shape of magnetic granules deviates considerably from the sphere, the magnetic field at small distances from its surface differs significantly from that of an equivalent dipole placed at the center. For a granule having the shape of an elongated ellipsoid, this situation is illustrated in Fig. 1, constructed using formulas (1) and (2) presented below. As can be seen, the points close to the equatorial plane (in which the other granules preferring antiferromagnetic ordering would occur), the field is significantly smaller than that of the equivalent dipole. This implies that the system exhibits a greater tendency to ferromagnetism as compared to the case of spherical granules. Therefore, we may expect that an increase in the "nonsphericity" of granules will inspire a transition from antiferromagnetic to ferromagnetic ground state of the system. This paper is devoted to an analysis of this question.

[1] The latter condition reflects a typical experimental situation, whereby systems of this type encountered in practice usually represent the ordered sets of quasi-ellipsoidal granules with parallel axes. This circumstance eliminates the necessity of considering the vortex magnetic states of various types typical of the systems of free dipoles [2].

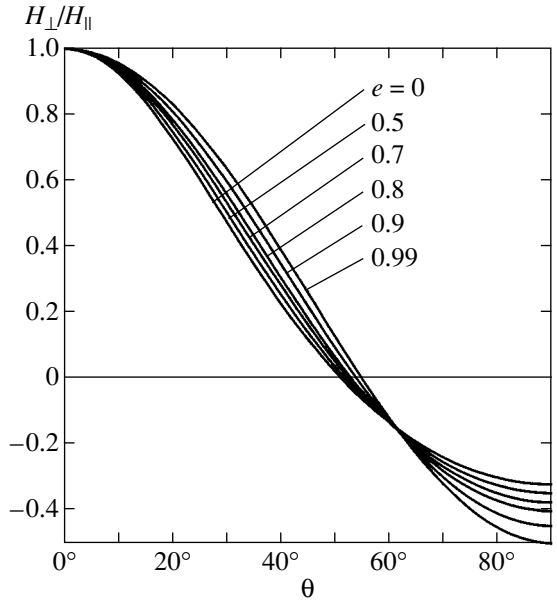

Fig. 1. The angular distribution of the magnetic field at a distance $R = 2.2a$ from the center of an ellipsoidal granule (θ is the angle between the long axis of the ellipsoid and the direction to the point of observation, a is the major semiaxis, and e is the eccentricity of the ellipsoid).

2. THE ENERGY OF PAIR DIPOLAR INTERACTION OF ELLIPSOIDAL MAGNETIC GRANULES

A constant magnetic field $\mathbf{H}(\mathbf{r})$ can be characterized by a magnetic potential ψ such that $\mathbf{H}(\mathbf{r}) = -\nabla\psi$. Therefore, the problem of a magnetic field created by a homogeneously magnetized ellipsoid is equivalent to the problem of a conducting ellipsoid exposed to an external homogeneous electric field, a solution to which is known (see, e.g., [7]). For an elongated ellipsoid of revolution with semiaxes a and b ($a > b$) and the eccentricity $e = \sqrt{1 - b^2/a^2}$, the magnetic field potential in the cylindrical coordinates x and ρ (the x axis coincides with the axis of revolution) is described by the formulas

$$\psi = \frac{3Mx}{e^3 a^3}(\mathrm{Arth}\,t - t), \quad t = \frac{e}{\sqrt{1+\xi}}, \tag{1}$$

where M is the magnetic moment of the ellipsoidal granule and ξ is the greater root of the equation

$$\frac{(\rho/a)^2}{1 - e^2 + \xi} + \frac{(x/a)^2}{1+\xi} = 1. \tag{2}$$

For a strongly elongated ellipsoidal granule ($e \longrightarrow 1$), Eqs. (1) and (2) give a simple relationship between the magnetic field strengths $H_\perp$ and $H_\parallel$ at the points situated in the equatorial plane and on the axis at a distance R from the center:

$$\frac{H_\perp}{H_\parallel} = \frac{\mathrm{Arth}(1/\sqrt{1+r^2}) - (1/\sqrt{1+r^2})}{\mathrm{Arth}(1/r) - [r/(r^2-1)]}, \tag{3}$$

where $r = R/a$. For $R \gg a$, this relationship yields a well-known result for the point dipole ($H_\perp/H_\parallel = -1/2$), while for $R = 2a$ (the case of contacting granules) we obtain $H_\perp/H_\parallel = -0.29$. Thus, the relative magnitude of the field in the equatorial plane is significantly smaller for an ellipsoidal magnetic granule than for the point dipole or a spherical particle.

Using Eqs. (1) and (2), we can determine the magnetic energy w of the pair dipole–dipole interaction of identical homogeneously magnetized ellipsoidal granules. Let the center of one granule coincide with the center of coordinates and that of the other occur at the point with the coordinates (x_0, y_0, z_0). Assuming that the magnetic moments of both granules are oriented along their major axes parallel to the x axis, we can write

$$\begin{aligned} w &\propto \int \frac{\partial\psi(x, y, z)}{\partial x} dx\,dy\,dz \\ &= \int \frac{\partial\psi(x', y', z')}{\partial x'} dx'\,dy'\,dz', \end{aligned} \tag{4}$$

where the integration is performed over the volume of the second granule (the coordinates $x' = x - x_0$, $y' = y - y_0$, $z' = z - z_0$ refer to the coordinate system with the origin at the center of the second granule, obtained by parallel transfer of the initial coordinate system).

Unfortunately, integral (4) cannot be exactly expressed in an analytical form. This circumstance practically hinders the possibility of numerically calculating the magnetic interaction energy for systems containing a large number of granules (as will be seen below, the numerical values of the interaction energy can be obtained at a required precision, provided that the number of pair contacts taken into consideration is on the order of 10^4). Therefore, it is important to find an approximate, but still sufficiently accurate, analytical expression for the energy.

As is clear from considerations in the Introduction section, the ground state can be ferromagnetic only in a system of rather strongly elongated granules (with an eccentricity $e \sim 1$). In this case, a change in the potential $\psi(x', y', z')$ in the direction perpendicular to the major axis is relatively small and we can use the corresponding expansion into a Taylor series

$$\begin{aligned} &\partial\psi(x', y', z')/\partial x' \\ &= \psi'_{0x} + [\psi''_{0x}x' + \psi''_{0y}y' + \psi''_{0z}z'] \\ &+ (1/2)[\psi'''_{0x}x'^2 + \psi'''_{0y}y'^2 + \psi'''_{0z}z'^2] \end{aligned} \tag{5}$$

$$+ [\psi''_{0x}\psi'_{0y}x'y' + \psi''_{0y}\psi'_{0z}y'z' + \psi''_{0x}\psi'_{0z}x'z'] + \ldots,$$

where

$$\psi'_{0x} = (\partial\psi/\partial x')_{x'=y'=z'=0},$$

$$\psi''_{0x} = (\partial^2\psi/\partial x'^2)_{x'=y'=z'=0},$$

$$\psi'''_{0x} = (\partial^3\psi/\partial x'^3)_{x'=y'=z'=0}$$

(and analogous expressions for the derivatives with respect to y' and z').

Now we can write the integral in the right-hand part of expression (4) as

$$\int\frac{\partial\psi(x', y', z')}{\partial x'}dx'dy'dz' = \int_{-a}^{a} S(x')dx',$$
$$S(x') = \int_{C_r}\frac{\partial\psi(x', y', z')}{\partial x'}dy'dz', \quad (6)$$

where the integration domain C_r represents a circle of radius $r = b\sqrt{1 - x'^2/a^2}$ centered at the point $(x', 0, 0)$. By substituting (5) into (6), we check that only terms of expansion (5) which contain no odd powers of x', y, and z' are retained upon integration. Taking into account the relation $\psi''_{0y} + \psi''_{0z} = -\psi''_{0x}$, we eventually obtain

$$\int\frac{\partial\psi(x', y', z')}{\partial x'}dx'dy'dz'$$

$$= \frac{4}{3}\pi ab^2\left\{\frac{1}{2}\left[\frac{\psi(a, 0, 0) - \psi(-a, 0, 0)}{2a} + \psi'_{0x}\right]\right. \quad (7)$$

$$\left. - \frac{3}{20}a^2\left(1 + \frac{32}{9\pi}\frac{b^2}{a^2}\right)\psi'''_{0x}\right\} + \left(\frac{b}{a}\right)^4 O(\psi^V),$$

where $O(\psi^V)$ is a sum of terms proportional to the fifth derivatives of the magnetic potential. As can be seen from expression (7), an increase in the distance between granules naturally leads to

$$\int\frac{\partial\psi(x', y', z')}{\partial x'}dx'dy'dz' \longrightarrow VH_x(x_0, y_0, z_0),$$

$$V = \frac{4}{3}\pi ab^2.$$

Numerical calculations showed that a contribution of the $O(\psi^V)$ to the representation (7) in all cases of practical interest is negligibly small (below 1%). Thus, using expression (7) for the magnetic energy and taking into account only the explicit terms will ensure an errors not exceeding 1%.

Based on the above qualitative considerations, we may suggest that a relative contribution of the antiferromagnetic interaction would decrease for the ellipsoidal granules approaching each other, and the system of such granules would exhibit a tendency to transition from the antiferromagnetic state (characteristic of the distant granules) to a ferromagnetic state, and eventually this transition will take place.

We have used the obtained expression for the magnetic energy for checking this hypothesis. The energy of a certain magnetic state was characterized by the total energy of interaction between a granule situated at the center of the system and the surrounding granules, the number of which was selected sufficiently large to make the result independent of this number (an example illustrating the dependence of the results of calculations on the number of neighbors taken into consideration is presented in Fig. 5 below). Of course, using this approach, we can only characterize the magnetic state energy of a system in which all granules are equivalent from the standpoint of their environment (i.e., a system of infinitely large size). Thus, the results presented below refer to an "infinite" system at zero temperature.[2]

3. THE GROUND STATE OF A 2D LATTICE OF NONSPHERICAL MAGNETIC GRANULES

Here, we will consider a 2D lattice of elongated (ellipsoidal) magnetic granules with parallel magnetic moments lying in the plane of the lattice.[3] In a square lattice of such nonspherical granules, antiferromagnetic states correspond to the three basic configurations of magnetic moments: S_{10} structure, representing alternating (with respect to the magnetization direction) chains of magnetic moments, which are aligned parallel to the direction of moments; S_{01} structure, comprising alternating chains perpendicular to the direction of moments; and S_{11} structure, representing the chains parallel to diagonals of the lattice unit cell (Fig. 2). Calculations show that the S_{10} configuration possesses a minimum energy among the antiferromagnetic structures. For this reason, we will restrict the consideration to a square 2D lattice of the S_{10} type.

According to the results of numerical calculations, the energy w_{AFM} of the antiferromagnetic S_{10} state of a square lattice composed of ellipsoidal granules is always lower than the energy w_{FM} of the ferromagnetic

[2] A nonmonotonic dependence of the energy of an antiferromagnetic state of the system on the system dimensions (see Fig. 5) is related to the influence of boundaries. A system can be considered as virtually "infinite" if the boundary contribution to the total energy is, for example, below 1%. For a situation illustrated in Fig. 5, the corresponding effective thickness of a boundary layer of granules is on the order of 100 lattice periods. Therefore, to within a 1% accuracy, a system with linear dimensions on the order of 1000 lattice periods can be considered as infinite.

[3] The ground state of a 2D lattice of granules with the magnetic moments perpendicular to the plane of the lattice is always antiferromagnetic.

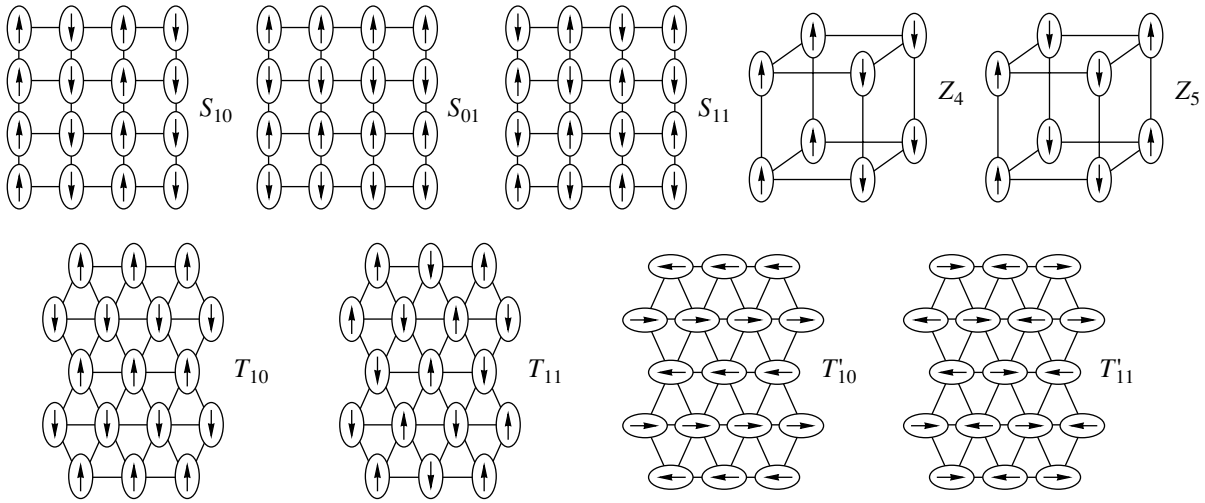

Fig. 2. Schematic diagrams of antiferromagnetic configurations of the square (S_{10}, S_{01}, S_{11}), triangular (T_{10}, T_{11}, T'_{10}, T'_{11}), and simple cubic (Z_4, Z_5) lattices of magnetic moments.

state, irrespective of the eccentricity and the ratio of the lattice period to the major axis. For example, Fig. 3 shows the difference $w_{FM} - w_{AFM}$ of the magnetic energies of a granule with $e = 0.95$ ($b/a = 0.31$) in a square lattice of such granules with ferromagnetic and antiferromagnetic ordering plotted versus period of the lattice. The result obtained by numerical calculations using relation (7) confirms the tendency of the two energies to leveling. However, the energy difference remains positive when the lattice period decreases (up to the state of granules touching one another), which implies that the antiferromagnetic state is energetically favorable.

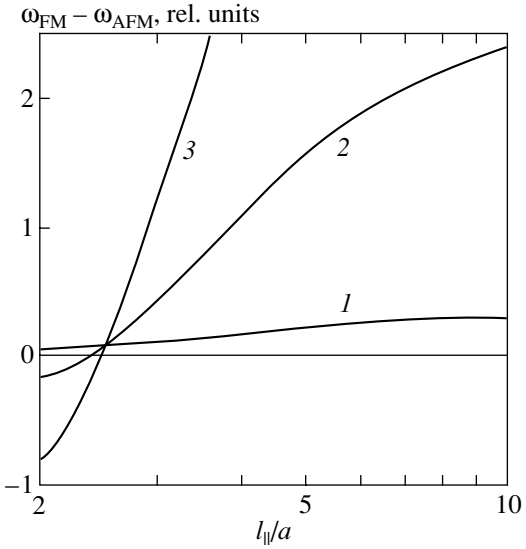

Fig. 3. Plots of the magnetic energy difference $w_{FM} - w_{AFM}$ versus longitudinal lattice period for the ellipsoidal granules with the eccentricity $e = 0.95$ forming rectangular lattices compressed to various degrees as determined by the ratio of longitudinal and transverse lattice periods $l_\parallel/l_\perp = 1$ (*1*), 2 (*2*), and 3 (*3*). The curves were calculated numerically using relation (7) for the square lattice samples composed of 50×50, 50×75, and 50×100 granules.

Nevertheless, we may "help" the system of magnetic moments to change the magnetic state. One possible way is to provide for a uniaxial compression of the lattice in the direction perpendicular to the magnetic moments of granules, whereby we essentially pass to considering a rectangular lattice rather than the square one.[4]

Figure 3 (curves *2* and *3*) illustrates the behavior of the magnetic energy difference $w_{FM} - w_{AFM}$ for such compressed rectangular lattices with the transverse period 1.5 and 2 times smaller than the longitudinal period. As can be seen, lattices featuring the ground ferromagnetic state under these conditions can exist.

The character of the ground magnetic state of the system under consideration is determined by three geometric parameters: (i) the eccentricity e of the granules; (ii) the ratio $l_\parallel/a$ of a longitudinal (parallel to the magnetic moment) lattice period to the longitudinal granule size; and (iii) the ratio of periods $l_\perp/l_\parallel$ of the rectangular lattice.

Figure 4 shows the phase diagrams of the system under consideration constructed in the space of these parameters. The ferromagnetic state of a rectangular lattice of ellipsoidal granules appears to be the ground state inside a quasi-rectangular region on the plane of parameters ($l_\perp/l_\parallel$, $l_\parallel/a$). The position and area of this region depend on the eccentricity e of the granules. The

[4] The transition to a rectangular lattice by no means implies a change in the sample shape, which is known to affect (via the so-called demagnetization factor) the magnetic state of a sample [4]. The results presented below refer, in fact, to a spherical sample, the magnetic state of which can be determined without taking into account the demagnetization factor.

upper and right-hand boundaries of the region of ferromagnetic states are determined from energy considerations, while the left-hand and bottom boundaries are determined from the condition of granules touching each other. It appears that, in particular, any rectangular lattice composed of granules with the eccentricity $e < 0.85$ is always antiferromagnetic.

Using an example of the lattice type under consideration, one can readily see how important it is to take into account a sufficiently large number of intergranular contacts in calculating the ground state energy. Figure 5 shows the plots of the energies of ferromagnetic (w_{FM}) and antiferromagnetic (w_{AFM}) states of a rectangular lattice versus the number of contacts involved in the calculation. As can be seen, a correct conclusion concerning the type of magnetic ordering in the given system is ensured if not less than 10^4 contacts are taken into account. This kind of check for the adequacy of numerical calculations should be performed in each particular case.

The curves presented in Fig. 5 are also well illustrative in another respect, showing that the interaction of each granule with its surroundings including about 100 nearest neighbors always favors establishment of an antiferromagnetic order of the magnetic moments (in this range, $w_{AFM} < w_{FM}$). It is only the interaction with the distant environment (another 10^3–10^4 granules) that can eventually make the ferromagnetic order energetically favorable ($w_{FM} < w_{AFM}$). A similar situation takes place in the case of a 3D lattice (see below). This circumstance implies that, despite being energetically favorable, the ferromagnetic state cannot be obtained by means of cooling. At a high temperature, the system possesses no magnetic order; for an ordered state to appear, the order must be established immediately within rather large regions (nuclei) containing about 100 granules, which is unlikely. This feature basically distinguishes the long-range dipole–dipole magnetism considered in our paper from the short-range exchange magnetism. A real way to obtaining a ferromagnetic state is offered by applying a sufficiently strong magnetic field, followed by slowly decreasing the field strength.

The approach described above can also be used to study the other 2D lattices, in particular, of a triangular configuration. This case is of interest by offering an example of a frustrated system degenerate with respect to energy. Here, the energy of antiferromagnetic states T_{11} and T'_{11} (see Fig. 2) coincides with the energy of a ferromagnetic state and is lower than the energies of frustrated structures T_{10} and T_{01}. However, this statement is only valid for a triangular lattice of spherical granules. In a system of ellipsoidal granules possessing a sufficiently large eccentricity ($e > 0.7$), the ground state corresponds to an antiferromagnetic structure (T_{10} or T_{01}). The corresponding phase diagrams are depicted in Figs. 6a (for T_{10} and T_{11}) and 6b (for T_{01} and T'_{11}). However, strongly extended lattices (with the extension coefficient $\beta \geq 3.6$ and 2.7, respectively) remain frustrated for any eccentricity of the magnetic granules.

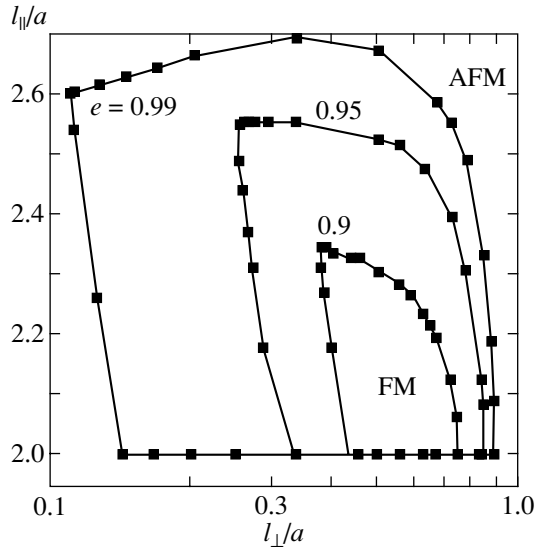

Fig. 4. Magnetic phase diagrams for a rectangular lattice of ellipsoidal granules. The ferromagnetic (FM) state is energetically favorable inside a closed quasi-rectangular region, the particular size of which depends on the eccentricity e of granules. Outside these regions (and anywhere for $e < 0.85$), the system is antiferromagnetic (AFM).

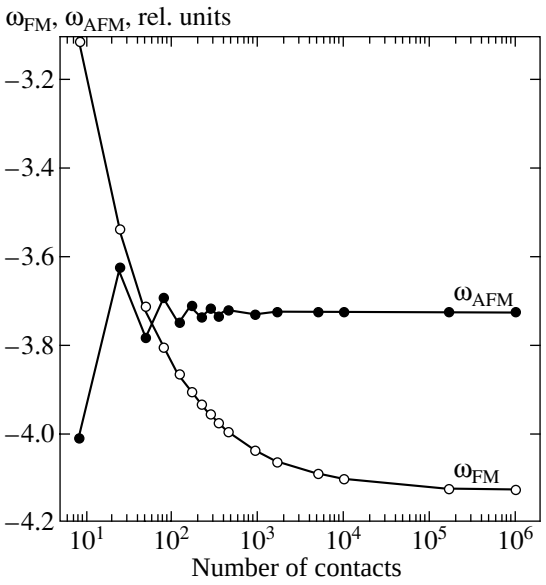

Fig. 5. Plots of the energies of ferromagnetic (w_{FM}) and antiferromagnetic (w_{AFM}) states of a rectangular lattice versus the number of contacts involved in the calculation performed for the following parameters: $l_\parallel/a = 2.6$; $l_\perp/a = 2.6/3$ (compression factor $1/\beta = 3$); $e = 0.99$.

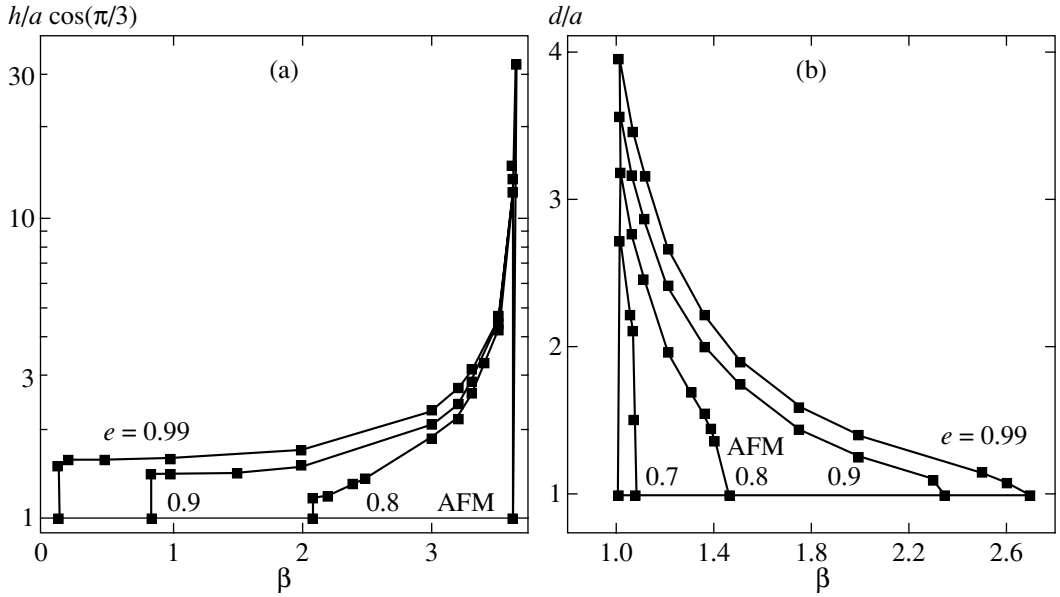

Fig. 6. Magnetic phase diagrams for a deformed triangular lattice of ellipsoidal granules with the magnetic moments (a) perpendicular and (b) parallel to the cell side edge aligned in the direction of extension. The antiferromagnetic (AFM) states (a) T_{10} and (b) T_{01} are energetically favorable inside quasi-triangular regions (bounded by the corresponding curves) outside which the system is always frustrated (h and d are the cell height and width, respectively).

4. THE GROUND STATE OF A 3D LATTICE OF NONSPHERICAL MAGNETIC GRANULES

Here, we will consider a tetragonal lattice of elongated (ellipsoidal) magnetic granules with magnetic moments parallel to the side edges of a unit cell obtained by uniaxially extending (or compressing) the initial simple cubic lattice. Even in this simple case, to which the consideration below is restricted, the system offers a large number of various antiferromagnetic configurations.

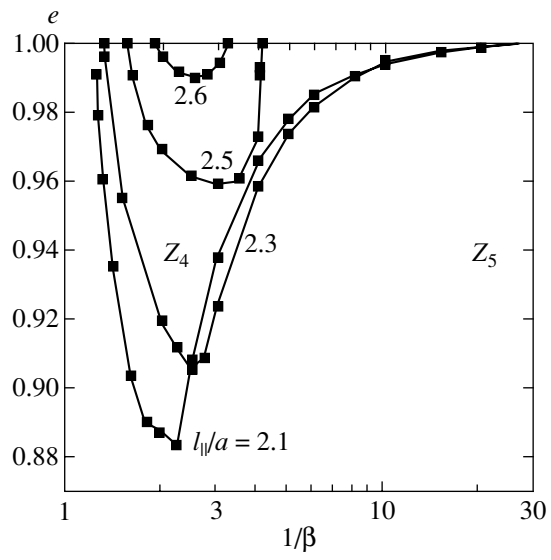

Fig. 7. Magnetic phase diagrams for a simple cubic lattice of ellipsoidal granules. Regions bounded by the curves correspond to the antiferromagnetic phase Z_4; regions outside the curves represent the antiferromagnetic phase Z_5.

As is known, a minimum energy in the case of a simple cubic lattice of point (or spherical) granules is offered by an antiferromagnetic configuration Z_5 (the result and notation from [5]). This structure comprises parallel chains of unidirectional magnetic moments, each being surrounded by the chains in which the moments are aligned in the opposite direction (Fig. 2). The same ground state is retained in a cubic lattice composed of ellipsoidal granules. If this structure is replaced by a tetragonal lattice with decreased period in the directions perpendicular to the magnetic moments of granules (with the cubic shape of the sample retained!), a sufficiently large eccentricity will inspire a magnetic structural transition to another ground state characterized by the configuration Z_4, in which every chain has only two (of the four) neighboring chains with the opposite direction of moments (Fig. 2). The resulting phase diagram of this system is depicted in Fig. 7.

It should be emphasized that the last result refers to the sample of cubic shape with a modified lattice. An alternative possibility is to retain the simple cubic lattice and change the sample shape. For a lattice of point granules, this possibility was considered in [4]. It was found that the energy of any antiferromagnetic state remains unchanged, while the energy of a ferromagnetic configuration depends on the sample shape in a simple manner: $w_{FM} \propto (4\pi/3 - N)$, where N is the demagnetization factor in the direction of the magnetic

moment. As the N value increases (which corresponds to elongation of the sample in this direction), the energy w_{FM} decreases and the ferromagnetic state may become energetically favorable. For a simple cubic lattice of point dipoles however, this decrease in w_{FM} is insufficient even in strongly elongated samples and the ground state of such lattices remains antiferromagnetic [4].

However, nonsphericity of the magnetic granules again basically changes the situation: calculations show that sufficiently elongated samples with nonspherical granules may become ferromagnetic. For example, the lattice of granules with $e = 0.99$ and $l/a = 2.1$ becomes ferromagnetic in a sample with the length exceeding thickness by a factor greater than eight. Here, we again (albeit for different reasons) encounter a situation when a ferromagnetic structure becomes energetically favorable due to the distant environment. In this case, the principally possible ferromagnetic state can also be attained only upon removal of the field applied initially in order to magnetize the sample to saturation.

Using the same approach, we can study more complicated 3D structures, including face- and body-centered cubic lattices, the energies of which in the case of point magnetic moments were determined in [4, 5].

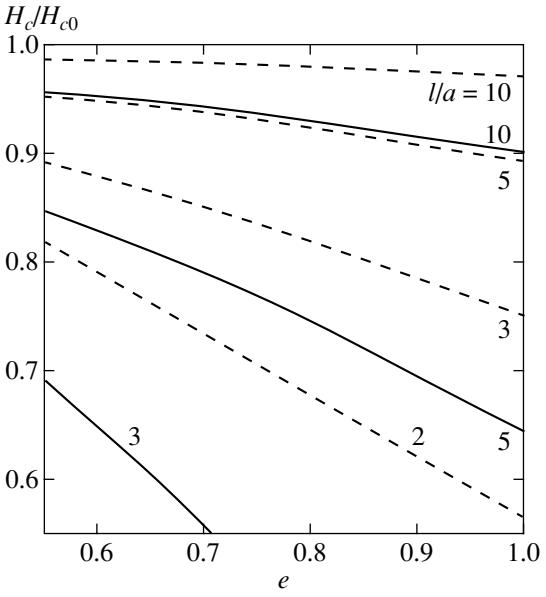

Fig. 8. Plots of the critical field strength versus eccentricity for the magnetic phase transition in a linear chain of ellipsoidal granules with constant magnetic moment and various chain l periods. Solid and dashed curves refer to 1D chains and 2D lattices, respectively.

5. FIELD-INDUCED MAGNETIC PHASE TRANSITION IN LATTICES OF NONSPHERICAL MAGNETIC GRANULES

The above considerations referred to the cases when the magnetic state of a lattice was changed either by increasing the eccentricity of the magnetic granules or by straining the lattice. However, there is one more way to "help" the system in the passage from antiferromagnetic to ferromagnetic state, which consist in applying an external magnetic field H_{ext}. If an applied field is parallel to the magnetic moments of the granules, the energy of each granule acquires the Zeeman increment [5]

$$w_H = -\mathbf{M} \cdot \mathbf{H}_{ext}.$$

The ferromagnetic state becomes favorable for an applied field strength of

$$H_e > H_c = (w_{FM} - w_{AFM})/M.$$

In order to study how the critical field strength H_c required for the phase transition depends on the eccentricity of granules (for the same magnetic moment), it is sufficient to calculate the difference $w_{FM} - w_{AFM}$ of energies of the ferro- and antiferromagnetic states of the system as a function of e. Figure 8 shows the results of such calculations in the form of the ratio H_c/H_{c0} (H_{c0} is the critical field for the phase transition in a system of point dipoles with the same moment) plotted versus eccentricity for ellipsoidal granules forming a linear chain and a square lattice. In the former case, the major axes (and, hence, the magnetic moments) of the granules are perpendicular to the chain direction; in the latter case, the moments are oriented along one side edge of the unit cell. As can be seen, the critical field strength significantly decreases with increasing eccentricity of the granules with dimensions (major axis $2a$) comparable with the lattice period l.

Owing to the aforementioned special character of the dipolar ferromagnetism, it is practically impossible to observe a sharp phase transition from antiferromagnetic to ferromagnetic state in a gradually increasing magnetic field. However, it is possible to monitor the reverse transition from ferromagnetic state (obtained upon application of a significantly stronger field ($H \gg H_c$) to antiferromagnetic state with decreasing applied field strength.

6. CONCLUSION

Thus, we have demonstrated that properties of the systems featuring dipole–dipole interaction significantly change on the passage from point (or spherical) dipoles to homogeneously magnetized ellipsoidal granules with significant eccentricity.

An important example of the 2D systems of this kind is offered by planar regular structures of nonspherical magnetic granules, now extensively studied as potential media for extremely high density data recording (patterned media) [8]. Typical structures comprise planar rectangular lattices formed by single-domain (almost homogeneously magnetized) granules of elongated shape possessing a uniaxial anisotropic geometry. The shape of these granules is close to an ellipsoid of revolution with an axial ratio of 3–5 ($e = 0.95$–0.98),

while the linear dimensions of granules are comparable with the lattice period. One bit of data is recorded by orienting the magnetic moment of a single granule. For this reason, the parameters of both the granules and the lattice must be selected so as to ensure that the magnetic moment could not change orientation spontaneously (under the action of thermal fluctuations and the magnetic field of surrounding granules).

The system in the initial state (upon data recording) is metastable, and, sooner or later, it will pass to the ground state. Under usual conditions, the rate of this transition is very small (the data can be stored for years!), but an increase in the temperature would allow the system to relax more rapidly into the ground state. The results of our investigation allow the nature of the final state to be predicted.

Another example relevant to the problem under consideration is offered by magnetic dielectric nanocomposites. The electric conductivity of such three-dimensional systems is related to the electron tunneling transitions between granules [9], the probabilities of which are determined by the mutual orientations of magnetic moments of the adjacent granules. Therefore, the resistance of this conducting medium directly depends on the magnetic state. The same is valid for the (giant) magnetoresistance of the system.

It must be noted, however, that the latter dielectric nanocomposites differ from the handmade magnetic recording structures of the patterned media type by the presence of a certain scatter in the parameters of magnetic granules (size, eccentricity, axis orientation, etc.). In connection with this, there arises a problem of stability of the ground state of lattices composed of the ferromagnetic granules with respect to such a scatter. We can briefly mention the following results obtained by model calculations for 2D rectangular lattices of ellipsoidal granules. The ground state of a rectangular lattice with a long period of $l_{\|} = 2.5a$, composed of granules with $e = 0.95$, remains ferromagnetic despite (i) a uniform distribution of the orientation of granules within ±15° (for the ratio of the lattice periods $l_{\perp}/l_{\|} = 0.5$ or 2) and (ii) a uniform distribution of granule dimensions within ±10% (for the ratio of the lattice periods $l_{\perp}/l_{\|} = 0.3$). At the same time, the ground state of a square lattice with a long period of $l = 3a$ remains antiferromagnetic despite (i) a uniform distribution of the orientation of granules within ±7° (for granules with $e = 0.95$) and (ii) a uniform distribution of the eccentricity of granules within $e = 0.90$–0.99. These results are indicative of a certain stability of the ground state of the magnetic systems under consideration.

Finally, our results provide an answer to the question as to whether ferromagnetism can exist in a system of particles linked only by dipolar interactions. We have principally demonstrated that such systems can exist and determined particular parameters of some possible variants.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 00-02-17191, 02-02-16974, and 01-02-16420) and by the joint Russian-French PICS-RFBR Program (project no. 01-02-22004).

REFERENCES

1. R. Brout, in *Magnetism*, Ed. by G. T. Rado and H. Suhl (Academic, New York, 1965), Vol. II, Part A.
2. P. I. Belobrov, R. S. Gekht, and V. A. Ignatchenko, Zh. Éksp. Teor. Fiz. **84**, 1097 (1983) [Sov. Phys. JETP **57**, 636 (1983)]; P. I. Belobrov, V. A. Voevodin, and V. A. Ignatchenko, Zh. Éksp. Teor. Fiz. **88**, 889 (1985) [Sov. Phys. JETP **61**, 522 (1985)].
3. M. D. Costa and Yu. G. Pogorelov, in *Proceedings of the Seeheim Conference on Magnetism, Seeheim, Germany, 2001*, to be published in Phys. Status Solidi B.
4. J. A. Sauer, Phys. Rev. **57**, 142 (1940).
5. J. M. Luttinger and L. Tisza, Phys. Rev. **70**, 954 (1946).
6. J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1999).
7. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 8: *Electrodynamics of Continuous Media* (Fizmatgiz, Moscow, 1959; Pergamon, New York, 1984).
8. R. L. White, J. Magn. Magn. Mater. **209**, 1 (2000).
9. J. S. Moodera and G. Mathon, J. Magn. Magn. Mater. **209**, 248 (2000).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 759–769.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 884–896.
Original Russian Text Copyright © 2002 by Zubova, Balabaev, Manevitch.

SOLIDS
Structure

Diffusion of Topological Solitons and Dielectric α_c Relaxation in a Polymeric Crystal

E. A. Zubova[a, *], N. K. Balabaev[b], and L. I. Manevitch[a]

[a]*Semenov Institute of Chemical Physics, Russian Academy of Sciences, ul. Kosygina 4, Moscow, 119991 Russia*
[b]*Institute of Mathematical Problems in Biology, Russian Academy of Sciences, Pushchino, Moscow oblast, 142292 Russia*
**e-mail: zubova@center.chph.ras.ru*

Received November 15, 2001

Abstract—A molecular dynamics simulation was performed to estimate the effective mass and the diffusion and friction coefficients of point defects in macromolecular chains of crystalline polyethylene. The results were compared with theoretical mass and kinetic coefficient predictions for topological solitons, with which these defects were identified. The results are used to discuss the soliton model of dielectric α_c relaxation in weakly oxidized polyethylene. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The relaxation properties and special features of phase transitions in solids are known to be determined by the energy characteristics and the type of dynamics of structural defects. In the past decades, starting with [1], these defects have been identified with topological solitons. More precisely, it was shown [1] that, in certain approximations for the simplest model of a crystal (for a linear chain of atoms linked by Hooke springs in a "substrate" potential), the free energy could be represented in the form of the sum of contributions of phonons and topological soliton-like excitations, which did not interact with each other. In the simplest situations (the sine-Gordon and ϕ^4 models), the anharmonic contribution to the heat capacity could be calculated exactly [2, 3]. At the same time, the problem of calculating the kinetic coefficients, which determine the diffusion mobility and friction of solitons, has been studied in much less detail even in such simplest model systems, and the available numerical results require further analysis and interpretation (see Section 4 for more details). However, even such limited information is absent for more realistic models of crystals.

Among crystalline substances, polymeric crystals hold a special position because of their strong anisotropy and flexibility of molecular chains. These special features manifest themselves in the mechanisms of various physical processes, unusual from the point of view of standard solid-state physics. Even the types of structural defects responsible for intracrystalline mobility with large lattice deformations are peculiar. If we restrict ourselves to point defects, we must, alongside vacancies caused by chain breakages, consider vacancies caused by large localized deformations of valence angles and combined defects of polymeric chain tension and torsion or contraction and torsion. In addition, in contrast to their known analogues, point defects in polymeric crystals can have soliton-type mobility, which results in special laws governing physical processes with their participation (such as dielectric relaxation, heat transfer, and premelting).

Dielectric measurements in polymers can be an effective tool for studying molecular mobility in these substances. In solid amorphous polymers, whose polar side groups are rigidly bound with the chain, two dielectric loss peaks are usually observed (in temperature scans at one frequency) [4]. The low-temperature β peak is usually related to excitation of local torsional modes, and the high-temperature α_a peak, to mobility of large chain segments. The α_a peak indicates vitrification [5]. Several amorphous-crystalline polymers of high crystallinity have one more peak at high temperatures near the melting point. This peak was for the first time observed by Mikhailov in weakly oxidized polyethylene [6]. This peak (α_c relaxation) is also observed in polytrifluorochloroethylene [7] and polyvinylidene fluoride [8].

As a result of analysis of subsequent experiments (see reviews [4, 9] and references therein), Mansfield [10] and Skinner and Wolynes [11] in 1980 independently advanced the suggestion that this peak could be caused by the diffusion in crystalline fraction chains of soliton-like excitations several dozen CH_2 groups wide, namely, kinks of chain torsion through 180° with chain half-period elongation or contraction (for crystallographic order to be preserved outside the kink region).

This hypothesis explains several observed process characteristics, which could not be described by the other theories (see Appendix A). The particular relaxation model [12] based on the kink hypothesis, however, required the introduction of several adjustment parameters. The value of one of these (the effective fric-

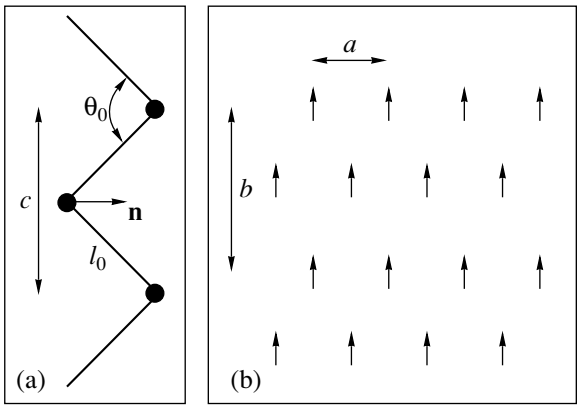

Fig. 1. Model polymeric crystal (polyethylene with united atoms): (a) plane zigzag chain parameters and (b) cross section plane of an equilibrium crystal (each small arrow indicates the direction from the atom of a molecule just below the plane to the nearest atom of a molecule above the plane).

tion coefficient of kinks, γ) calculated from the available experimental data was five orders of magnitude lower than its theoretical estimate (see Section 4) for the ϕ^4 model, which can be considered the simplest model of polymeric chains in crystals. A subsequent molecular dynamics simulation of the behavior of kinks in a simplified model (a chain of atoms bound by Hooke springs on a sine substrate [13]) substantiated that friction coefficient γ was too large for the kink model [12] to give at least the correct order for the central frequency of the process. Thus, in the absence of other hypotheses, theoretical explanations of dielectric α_c relaxation came to a deadlock in the late 1980s.

In this work, we suggest a solution to the problem. The diffusion of a kink of tension with torsion and (for comparison) a vacancy (kink of pure tension by a chain period) in a crystal chain is analyzed using a three-dimensional molecular dynamics model of a simple polymeric crystal containing zigzag chains. The mass, friction coefficient, and diffusion coefficient of a tension-with-torsion kink at temperature T = 300 K and pressure p = 3.5 kbar are estimated. The results are compared with theoretical models. It is shown that the kink hypothesis may be valid for dielectric α_c relaxation if one of the model [11, 12] assumptions is abandoned.

2. A NUMERICAL MODEL OF A POLYMERIC CRYSTAL (POLYETHYLENE WITH "UNITED" ATOMS)

For our purposes, it is sufficient to use the simplest three-dimensional dynamic model of a polymeric crystal comprising zigzag chains that can form kinks of torsion with tension (or contraction). Such a model was developed by us earlier [14] to study the dynamics of point structural defects [15, 16].

This is a model of polyethylene with united atoms (see Fig. 1). The chain is a planar *trans* zigzag; the bonds between atoms (point particles of mass m) are absolutely rigid and have length l_0; the energies of deformation of valence θ_n and conformational τ_n angles have the form

$$U_3(\theta_n) = \frac{1}{2}K_\theta(\theta_n - \theta_0)^2, \quad (1)$$

$$U_4(\tau_n) = K_0 + K_1\cos(\tau_n) + K_3\cos(3\tau_n), \quad (2)$$

and the atoms separated by more than two neighbors or situated in different chains interact by the law

$$U(r) = \begin{cases} U_{LJ}(r) - U_{LJ}(R), & r \le R, \\ 0, & r > R, \end{cases}$$

where $U_{LJ}(r) = 4\epsilon((\sigma/r)^{12} - (\sigma/r)^6)$ is the Lennard-Jones potential with a minimum at $r_0 = 2^{1/6}\sigma$. The numerical constant values that we used are summarized in the table. The crystal model had periodic boundary conditions in all three directions. The cell for calculations was a rectangular parallelepiped. The corresponding classical Lagrange equations of the first kind were solved using the Verlet leapfrog algorithm [19] taking into account the restrictions imposed by rigid bonds [20]. The periodic boundary conditions in the direction of the molecular axis allowed the dynamics of the defect to be examined indefinitely, and the periodic boundary conditions in the cross section plane obviated the necessity for introducing nonphysical boundary conditions of the type of a rigidly fixed second coordination sphere.

As the length of the projection of the molecule onto its cross section plane was $l_\perp \approx 0.843$ Å, and the van der

Model crystal parameters

Parameter	Value	Refs.	Parameter	Value	Refs.
m, au	14	–	K_1, kJ/mol	1.675	[17]
l_0, Å	1.53	[17]	K_3, kJ/mol	6.695	[17]
θ_0	113°	[17]	ϵ, kJ/mol	0.4937	[18]
K_θ, kJ/mol	331.37	[17]	σ, Å	3.8	[18]
K_0, kJ/mol	8.370	[17]	R	$2r_0$	–

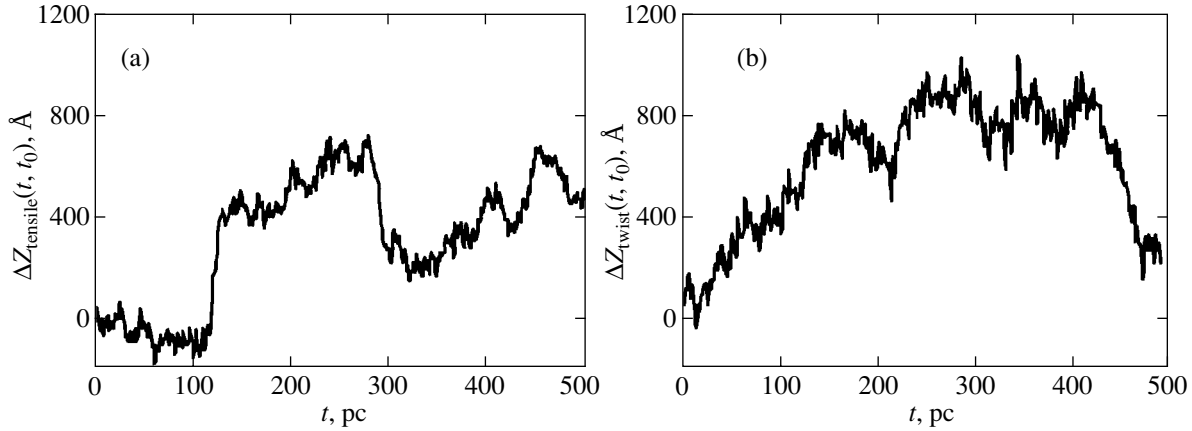

Fig. 2. Wandering of point defects over a polymeric crystal chain: characteristic time dependences of (a) vacancy (defect of tension by one chain unit) coordinate and (b) the coordinate of the defect of rotation through 180° with tension by half-chain unit. Each point in the plots was obtained by averaging over a 1-ps time interval.

Waals radii of the united atoms were $r_0 \approx 4.265$ Å $\approx 5l_\perp$, the packing of such plane zigzags in the crystal at $T = 0$ was close to the packing of cylinders (Fig. 1b).

3. DIFFUSION OF KINKS IN A CHAIN OF A HEATED POLYMERIC CRYSTAL: A MOLECULAR DYNAMICS EXPERIMENT

The cell for calculations contained 30 molecules, of which 29 were constructed of 400 CH_2 groups, and one molecule, of 398 groups (in studying the diffusion of a chain unit vacancy) or 399 groups with torsion through 180° (in studying the diffusion of a torsion-with-tension defect). Such a length of molecules is sufficient to prevent soliton action on itself when periodic boundary conditions are imposed along chains (the defect extension is of the order of 70 CH_2 groups).

The sample was heated to the required temperature with the use of a Berendsen thermostat [19] and relaxed at this temperature and a constant pressure to the equilibrium state, when the mean values of cell parameters ceased to change. The equilibrium cell parameters a, b, and c (see Fig. 1) at temperature $T = 300$ K and pressures $p = 0$ and $p = 3.5$ kbar were {4.24, 8.00, and 2.53 Å} and {4.12, 7.85, and 2.53 Å}, respectively. After relaxation, the external thermostat was switched off, and the volume of the isolated sample was fixed. Further, the diffusion of the point defect in a thermally excited chain was studied; a natural thermostat for the chain was the heated neighboring crystal chains. We found that the temperature of the sample remained constant to a satisfactory accuracy.

In the numerical experiment, we monitored changes in the position $Z_{cm}(t)$ of the center of mass of a periodic fragment (initially situated in the cell for calculations) of an infinite chain with a defect. This value can easily be recalculated to the displacement $\Delta Z(t;\ t_0)$ of the defect in time t counted from t_0,

$$\Delta Z(t;\ t_0) = -\frac{N}{\Delta N}(Z_{cm}(t + t_0) - Z_{cm}(t_0)), \qquad (3)$$

where N is the number of atoms in a chain without defects, ΔN is the number of particles removed from the chain in defect formation ($\Delta N = 1$ for a torsion-with-tension defect and $\Delta N = 2$ for a pure tension defect). The time t dependences of the displacement of simple tension defects $\Delta Z_{\text{tensile}}(t;\ t_0)$ and torsion-with-tension defects $\Delta Z_{\text{twist}}(t;\ t_0)$ are shown in Fig. 2.

The mean square displacement of the center of mass of a chain with a defect $\langle (Z_{cm}(t + t_0) - Z_{cm}(t_0))^2 \rangle_{t_0}$ is shown in Fig. 3 as a function of time t (averaging was performed along the trajectory over the t_0 starting point) for simple tension and torsion-with-tension defects.

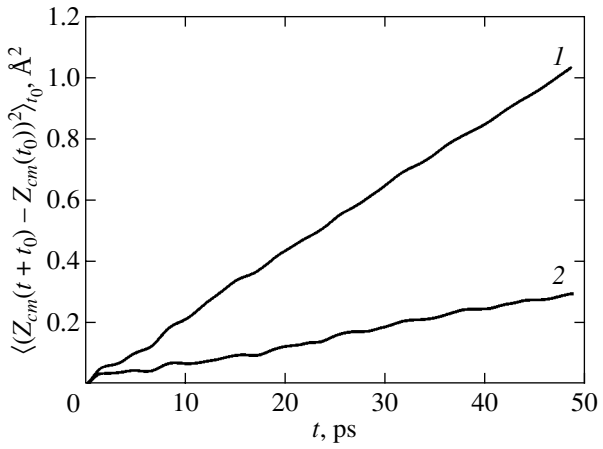

Fig. 3. Mean square displacement of the center of mass of a molecule with a defect as a function of time at pressure $p = 0$: (*1*) vacancy and (*2*) torsion-with-tension defect. Averaging was performed along a trajectory 500 ps long.

Figure 3 shows that the curves begin to oscillate about straight lines in time shorter than $\tau_1 \sim 2$ ps; this immediately gives a crude lower bound estimate for the kink friction coefficient, $\gamma_{\text{bottom}} \sim 1/\tau_1 \sim 0.5 \times 10^{12}$ s^{-1}. The diffusion coefficients of both defects calculated from the slopes of the lines shown in Fig. 3 coincide within the error of measurements, $D(p = 0) \approx 4 \times 10^{-2}$ cm^2/s. At a $p = 3.5$ kbar pressure, the diffusion coefficients estimated by the same procedure were $D(p = 3.5\ \text{kbar}) \approx 3 \times 10^{-2}$ cm^2/s. As expected, applying pressure to the sample decreased the diffusion coefficients.

A description of the dynamics of a simple tension defect in a cold crystal reduces (see [15] and the references therein) to the integrable sine-Gordon equation. The system of equations for two coupled fields of chain torsion with tension in the substrate potential generated by neighboring chains does not have soliton solutions in the classical sense. An approximate analytic solution for a soliton-like topological excitation, a torsion-with-tension defect at $T = 0$, was given in [16]. In that work, the spectrum of the velocities of such waves was also studied; the upper bound of this spectrum was found to be four times lower than the upper bound of the spectrum for a simple tension defect. Indeed, Fig. 2 shows that the character of simple tension defect wandering differs from the dynamics of a torsion-with-tension defect. In Fig. 2a, there are rare extended regions of rectilinear uniform motion of vacancies at comparatively high velocities; such regions are absent in the trajectory of motion of a torsion-with-tension defect (for instance, a vacancy can run over about 350 CH_2 groups at an almost constant velocity of about 4.5 km/s, whereas the longitudinal sound velocity along the chain in a crystal at $T = 0$ equals 14.7 km/s). Nevertheless, in spite of these differences in the dynamics, the diffusion coefficients of both defects were found to be virtually identical. The value obtained in this work gives the upper estimate of the diffusion coefficient in real polyethylene.

4. DIFFUSION OF KINKS IN A THERMAL BATH: THEORETICAL MODELS AND REAL PHYSICAL SYSTEMS

The diffusion of solitons is usually considered using two approaches (see review [21]). In the first approach (see [22, 23]), a model of the system (for instance, the sine-Gordon model) is augmented by a pair of Langevin terms, namely, additive white noise $\zeta_A(x, t)$ and effective friction with the corresponding coefficient γ_A,

$$\phi_{tt} - v_s^2\phi_{xx} + \omega_0^2 \sin\phi = -\gamma_A\phi_t + \zeta_A(x, t). \quad (4)$$

Here, v_s is the sound velocity in an isolated chain, and ω_0 is the width of the phonon spectrum gap, which appears when the chain is placed into a sine substrate field. In first-order perturbation theory, the velocity of a kink is obtained in the form of the Langevin equation with the same friction coefficient γ_A and white noise. Sometimes, spatially correlated noise (for instance, noise with an exponentially decreasing correlation function) is introduced into (4). The relation between kink diffusion and friction coefficients then differs from the Einstein equation by a factor that depends on the ratio between the noise correlation length and the kink width.

In a similar way, multiplicative white noise $\zeta_M(x, t)$ is also introduced:

$$\phi_{tt} - v_s^2\phi_{xx} + \omega_0^2 \sin\phi = -\gamma_M\phi_t + \zeta_M(x, t)\sin\phi. \quad (5)$$

In first-order perturbation theory, this equation also leads to the Langevin equation for the kink velocity [24, 25].

Equations (4) and (5) are the simplest models of thermostats in a chain with a kink. The first one presupposes that the role of the thermostat for a kink is played by (small-amplitude) vibrations of chain atoms (that is, phonons). Equation (5) assigns this role to atomic vibrations in neighboring chains that form the substrate. Not to mention that there is no single physical problem that leads to a sine-Gordon-type equation for which the approximations of both additive and multiplicative white noise are warranted, the main problem is nevertheless first-principles calculations of the friction coefficient.

The second approach to the diffusion of kinks just supposes a study of kink diffusion in interaction with "natural" thermostats, viz., a "bath" of phonons in the chain and in neighboring chains that form the substrate. Several steps in this direction were made. The interaction of kinks with a bath of phonons in a chain coupled (for instance, at chain ends) with a thermostat at temperature T was studied by perturbation theory in the approximation

$$\frac{k_B T}{E_s} \ll 1, \quad (6)$$

where k_B is the Boltzmann constant and E_s is the static kink energy. It was found that the diffusion of kinks in the integrable sine-Gordon model qualitatively differed from the diffusion in the nonintegrable ϕ^4 model. In the first model, kinks on average preserved their initial velocity v_0 and only the coordinates of kinks experienced smearing, because the interaction between a kink and a phonon was limited to a shift in the position of the kink caused by phonon passage. Such a diffusion is called anomalous. Diffusion coefficient D_A calculated from the mean square deviation from $v_0 t$ by perturbation theory was [26]

$$D_A = \frac{v_s^2}{\omega_0}\frac{2}{3\pi}\left(\frac{k_B T}{E_s}\right)^2 \quad (7)$$

(in earlier work [27], in which anomalous diffusion was considered on the assumption that kinks interacted with

nonoverlapping phonon packets, the diffusion coefficient was found to be eight times larger than in work [26] that we cite, where overlapping of phonon packets was taken into account).

At temperatures of about 320–420 K, phonon spectrum gap frequency $\omega_0 = 1.6 \times 10^{13}$ s^{-1}, sound velocity in an isolated chain $v_s = 5$ km/s, and effective kink mass $m^* = 5 \times 10^{-27}$ kg (these values are from [12]; all the further numerical estimates refer to these system parameters unless otherwise stated), the ratio between temperature in energy units and the static kink energy $E_s = m^* v_s^2 \approx 18$ kcal/mol is

$$\frac{k_B T}{E_s} \sim (3.4\text{–}4.5) \times 10^{-2}. \tag{8}$$

Provided this value is sufficient for condition (6) to be satisfied, (7) predicts the value $D_A \sim (2.1\text{–}4.2) \times 10^{-4} v_s^2/\omega_0 \sim (3.4\text{–}6.7) \times 10^{-5}$ cm^2/s (compare this with the diffusion coefficient of sodium chloride in water at room temperature, 1.1×10^{-5} cm^2/s, and the diffusion coefficient of hydrogen in oxygen at 0°C, 0.7 cm^2/s).

The diffusion coefficient for the ϕ^4 model calculated in the same approximation of linear interactions between solitons and phonons differs from (7) by a small multiplier [28, 29]. It, however, turns out that, in contrast to the integrable sine-Gordon model, in which there is no diffusion other than anomalous, nonlinear interactions between solitons and phonons in the ϕ^4 model result in exchange of momenta between them. In the first nonvanishing (fourth) order of perturbation theory, there arises viscous friction with coefficient γ_V. In time $t \gg \tau_V$,

$$\begin{aligned} \tau_V &= \frac{1}{\omega_0} \frac{1}{8 \times 4.56 \times 10^{-3}} \left(\frac{E_s}{k_B T} \right)^2 \\ &\approx (2.4\text{–}1.4) \times 10^4 \frac{1}{\omega_0} \approx (1.55\text{–}0.88) \times 10^{-9} \text{ s}, \end{aligned} \tag{9}$$

a kink loses memory of the initial velocity and performs diffusion motion with the coefficient [30]

$$\begin{aligned} D_V &= \frac{v_s^2}{\omega_0} \frac{1}{8 \times 4.56 \times 10^{-3}} \frac{E_s}{k_B T} \\ &\approx (8.1\text{–}6.1) \times 10^2 \frac{v_s^2}{\omega_0} \approx (13\text{–}10) \text{ cm}^2/\text{s}, \end{aligned} \tag{10}$$

where the transition from friction coefficient γ to diffusion coefficient D is performed according to the Einstein equation

$$\frac{D m^* \gamma}{k_B T} = 1. \tag{11}$$

A comparison of the orders of magnitude of (7) and (10) shows that anomalous diffusion can be ignored for the ϕ^4 model at times $t \gg \tau_V$. Unfortunately, the question of possible contributions of higher order perturbation theory terms to the viscosity of a kink was not considered (although the question of the ratio between D_A and D_V was discussed [31, 32]).

The above comparison of the diffusion coefficients D_A and D_V for the integrable and nonintegrable models, respectively, leads us to suggest that the main source of viscosity in real physical systems should be their deviations from the model integrable system. This suggestion is substantiated by the numerical analysis of anomalous diffusion in the sine-Gordon model [33]. The authors had to resort to artificial quenching of low-frequency short-wave phonons of a discrete chain, interactions with which resulted in strong normal diffusion (because of viscosity). The fine anomalous diffusion effect of the continuous integrable model was simply indiscernible against the background of normal diffusion. According to [15, 34], a vacancy in a chain of a cold polymeric crystal differs from the soliton solution to the continuous sine-Gordon equation precisely by the type of its dynamics; namely, a vacancy emits phonons whose phase velocity coincides with the kink velocity (such phonon modes exist because the dispersion curves of phonons in a polymeric crystal and the dispersion curves calculated for the Frenkel–Kontorova model do not coincide). As already mentioned, the upper boundary of the spectrum of the velocities of torsion-with-tension kinks, which can move along chains virtually without friction, is four times lower that the corresponding boundary for simple tension kinks. It follows that exchange of momenta between kinks and torsional phonons is still stronger. It appears that precisely this, the strongest, viscosity mechanism determines the diffusion coefficient of point defects in polymeric crystal chains. The anomalous diffusion coefficient of the sine-Gordon model therefore has no bearing on this value, and the normal diffusion coefficient in the ϕ^4 model can only serve as a very crude upper estimate.

Indeed, the diffusion coefficient $D(p = 0)$ value obtained in Section 3 is almost three orders of magnitude smaller than the normal diffusion coefficient of the ϕ^4 model [Eq. (10)]. Accordingly, the curve of the mean square defect displacement tends to a straight line in times three order of magnitude shorter than the characteristic time τ_V (9) of establishing equilibrium in the process of normal diffusion in the ϕ^4 model. It follows that friction responsible for viscosity in our model is indeed much stronger than friction caused by the nonintegrability of the ϕ^4 model.

5. THE AUTOCORRELATION DIPOLE MOMENT FUNCTION: A MOLECULAR DYNAMICS EXPERIMENT

The complex permittivity of a "rarefied gas" of C=O dipoles fixed in chains is proportional to the Fourier

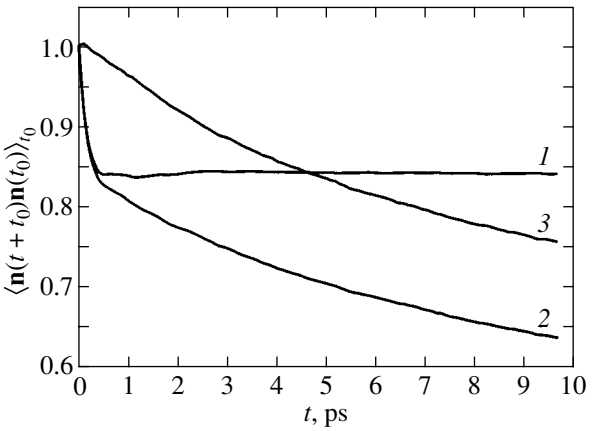

Fig. 4. Normalized dipole moment autocorrelation functions according to the molecular dynamics experiment: (*1*) for a dipole in a chain without a kink and (*2*) for a dipole in a chain with a kink. Each curve was obtained by averaging over 16 trajectories 500 ps long each. (*3*) The result of dividing curve (*2*) by curve (*1*): the contribution of kinks to the autocorrelation function.

transform of the time derivative of normalized autocorrelation function $\Phi(t)$ of one dipole moment $\mathbf{P}(t)$,

$$\Phi(t) = \frac{\langle \mathbf{P}(t)\cdot\mathbf{P}(0)\rangle}{\langle \mathbf{P}(0)\cdot\mathbf{P}(0)\rangle}.$$

In the model of dielectric relaxation caused by the diffusion of defects that invert dipoles and within the approach that ignores phonons in the chain, the physical meaning of the $\Phi(t)$ function is the probability that, in time t, the dipole remains uninverted by a kink that approaches it. If the linear kink density in the chain (the number of kinks per unit chain length) is n_0 and their diffusion coefficient equals D, then the asymptotic behavior of the $\Phi(t)$ function at long times (obtained fairly long ago, in [35]) has the form

$$\Phi_\infty(t) = \exp(-\sqrt{t/\tau_\infty}), \tag{12}$$

where

$$\tau_\infty = \frac{\pi}{16}\frac{1}{Dn_0^2} \tag{13}$$

is the mean dipole expectation time for the arrival of a kink [indeed, the mean distance to the nearest kink is about $1/(2n_0)$, and the kink path for time t is of the order of $\sqrt{2Dt}$; equating these values yields (13) accurate to a factor of $\pi/2$].

Skinner and Park [12] used another, more complex, expression for the autocorrelation function, which remained valid at much smaller $dt\gamma$ and $t\gamma$ values and therefore included kink effective mass m^*:

$$\Phi(t) = \exp\left(-\sqrt{\frac{16}{\pi}}n_0\sqrt{\frac{k_BT}{m^*}}\frac{\sqrt{\gamma t - 1 + \exp(-\gamma t)}}{\gamma}\right). \tag{14}$$

Our three-dimensional molecular-dynamics model of a polymeric crystal allows the validity of (14) to be directly checked. Indeed, a comparison of the asymptotic behavior of (14) at short times,

$$\Phi_0(t) = \exp(-t/\tau_0), \tag{15}$$

where

$$\tau_0 = \sqrt{\frac{\pi}{8}}\frac{1}{n_0}\sqrt{\frac{m^*}{k_BT}} = \sqrt{\tau_\infty\frac{2}{\gamma}}, \tag{16}$$

with its asymptotic behavior at long times (12) immediately shows that, at low densities n_0, time τ_∞ (13) is the upper estimate of the characteristic time of the process. This time decreases as $1/n_0^2$ as the density of kinks n_0 grows. At a density of one kink per 400 CH_2 groups, it reaches a value of approximately 130 ps, already observable in our molecular dynamics experiment.

The numerical experiment was performed at temperature T = 300 K and pressure p = 3.5 kbar. Time t averaging of $\mathbf{n}(t)\mathbf{n}(0)$, where $\mathbf{n}(t)$ is the unit vector in the direction of the bisector of one of the chain valence angles (see Fig. 1a), gave the normalized dipole moment autocorrelation function. The results of such a molecular dynamics experiment for two chains, one with a single torsion-with-tension kink per 400 CH_2 groups with a dipole and the other without a kink (recall that periodic boundary conditions were imposed along the chain axis; our experiment therefore corresponded to 1.9×10^5 cm^{-1} and zero kink densities), are shown in Fig. 4. For zero kink density, the decrease in the autocorrelation function with a characteristic time shorter than 1 ps corresponded to a fast β process, which caused torsional disordering of chain atoms; the mean angle between bisectors of the valence angles in the molecule cross section plane approximately equaled 33°, that is, the arc cosine of the constant value that the curve reached. In the presence of a kink in the chain, the experimental curve was formed as a superposition of the slow α_c and fast β processes,

$$\langle \cos(\phi_\beta + \phi_{\alpha_c})\rangle$$
$$= \langle \cos(\phi_\beta)\cos(\phi_{\alpha_c})\rangle - \langle \sin(\phi_\beta)\sin(\phi_{\alpha_c})\rangle.$$

On the assumption that these processes are mutually independent and as $\langle\sin(\phi_\beta)\rangle = 0$, we find that, to determine the $\langle\cos(\phi_{\alpha_c})\rangle$ contribution of kinks (for which we have theoretical prediction (14)) to the autocorrelation function, we must pointwise divide sum curve *2* by curve *1*. The result of this division is also shown in Fig. 4 (curve *3*).

The asymptotic behavior of the autocorrelation function at short times given by (15) only depends on the kink mass at a given density of kinks and is independent of the kink friction coefficient. The initial curve portion (more precisely, the $[\ln(\Phi(t))/t]|_{t=0}$ value) can therefore be used to estimate the mass of the

kink and then determine the second parameter, the friction coefficient. However, just in the initial curve portion $0 < t < 2$ ps, the fast β and slow α_c processes are not independent, and, for this reason, the mass of the kink can only be determined from the experimental curve with a very large error, $m^*_{\text{exp}} = (1.5 \pm 0.3) \times 10^{-27}$ kg. The theoretical estimate of this value from the form of the approximate analytic solution for a soliton-like topological excitation (a torsion-with-tension defect) at $T = 0$ (see Appendix B), $m^*_{\text{theor}} \approx 1.2 \times 10^{-27}$ kg, is at the lower boundary of the experimental confidence interval. Accordingly, the friction coefficient can be determined from the available portion of the curve and m^*_{exp} as a value in the range $(0.45–0.73) \times 10^{12}$ s^{-1}. For reasons that are clear from the discussion in Section 4, theoretical estimates of this value have not been obtained yet. For our model, this value is of the order that follows from the lower estimate $\gamma_{\text{bottom}} \sim 0.5 \times 10^{12}$ s^{-1} obtained in Section 3 in analyzing the form of the $\langle (Z_{cm}(t + t_0) - Z_{cm}(t_0))^2 \rangle_{t_0}$ function.

The "kink part" of the dipole moment autocorrelation function obtained in the molecular dynamics experiment is shown in Fig. 5 in comparison with theoretical curve (14) constructed for the parameter values $m^*_{\text{best}} = 1.5 \times 10^{-27}$ kg and $\gamma_{\text{best}} \sim 0.56 \times 10^{12}$ s^{-1} and its two asymptotic expressions (15) and (12). We see that, at $t > 3$ ps, the experimental curve is very closely described by theoretical dependence (14). Substituting the obtained m^*_{best} and γ_{best} values and the $D(p = 3.5$ kbar) diffusion coefficient independently determined in Section 3 gives the ratio $Dm^*\gamma/k_BT = 0.6 \pm 0.5$.

6. THE THEORY OF DIELECTRIC α_c RELAXATION: WHAT SHOULD BE CHANGED

We showed in the preceding section that the dipole moment autocorrelation function at a fixed density of kinks is very closely described by theoretical formula (14) with realistic kink mass $m^*_{\text{best}} = 1.5 \times 10^{-27}$ kg and friction coefficient $\gamma_{\text{best}} = 0.56 \times 10^{12}$ s^{-1} values. However, Skinner and Park [12] used friction coefficient values of $(5 \times 10^2–1 \times 10^5)$ s^{-1} to describe the available experimental data on dielectric α_c relaxation. This was several orders of magnitude below our estimate. The diffusion coefficient corresponding to the friction coefficient $\gamma_{SP} \sim 5 \times 10^3$ s^{-1} is $D_{SP} \sim 2 \times 10^6$ cm^2/s, which is five orders of magnitude larger than the normal diffusion coefficient D_V [Eq. (10)] and eight orders of magnitude larger than the coefficient obtained in our molecular dynamics experiment.

The reason why the diffusion coefficient was assigned such a large value can be understood from physical considerations. Let kinks experience free Brownian movement described by the Fokker–Planck equation (we study slow processes which correspond to the roughest time scale $dt\gamma$, $t\gamma \gg 1$). The system then only involves two physical parameters in addition to the linear density of dipoles (n_d), namely, the density of kinks (n_0) and their diffusion coefficient (D). The dielectric relaxation rate should increase as either the density of kinks or their diffusion coefficient increases. For this reason, the results only of dielectric measurements cannot be used to independently determine both parameters. An accurate theoretical consideration substantiates this conclusion. Indeed, the density of kinks and the diffusion coefficient appear in Eqs. (12) and (13), which correspond to $dt\gamma$, $t\gamma \gg 1$, only as the Dn_0^2 combination. The shape of curve (14) in principle allows the parameter

$$B = \sqrt{\frac{16}{\pi}} n_0 \sqrt{\frac{k_BT}{m^*}} = \sqrt{2}/\tau_0$$

to be determined from the initial portion of curve (15), after which the γ value can be obtained. Precisely in this way, we estimated m^* and then γ at known n_0 and T in Section 5. The experimental data [36, 37], however, cannot be used to construct the initial curve portion both because of the presence of β and α_c processes and because of technical frequency limitations, 10–10^5 Hz. It may well be that, for this reason, the formula of the equilibrium theory of fluctuations as applied to a chain described by the continuous sine-Gordon equation [38] was used in [12] for the density of kinks,

$$n_0^{(fl)} = 2\sqrt{\frac{2}{\pi}} \frac{\omega_0}{v_s} \sqrt{\frac{E_s}{k_BT}} \exp\left(-\frac{E_s}{k_BT}\right). \qquad (17)$$

Substituting (17) into (13) yields an experimentally observed qualitative dependence of the type $\sim\exp(-W/k_BT)$ for the central process frequency. For our numerical

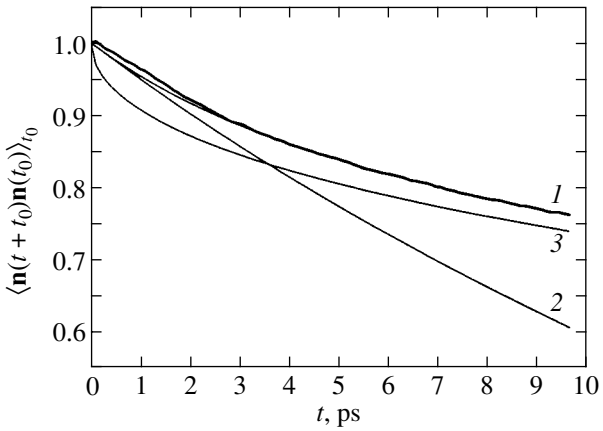

Fig. 5. Comparison of the "kink part" of the normalized dipole moment autocorrelation function obtained in the molecular dynamics experiment (uneven curve) with (*1*) theoretical curve calculated by (14) and its asymptotic expression (*2*) at short times (15) and (*3*) at long times (12).

values, (17), however, gives $n_0^{fl} \sim (4.2 \times 10^{-5}–4.6 \times 10^{-2})$ cm^{-1} (for one of the temperature and volume values, Skinner and Park obtained the estimate $n_{0SP} \sim 3.4 \times 10^{-3}$ cm^{-1}). At such a value, there are $5.4 \times 10^{-10}–6 \times 10^{-7}$ kinks per C=O dipole (1000 CH_2 groups)!

In such an anomalously "rarefied gas" of kinks, dielectric α_c relaxation is not sufficiently slow to ensure a decrease in the dipole autocorrelation function in the necessary time without using an anomalously large kink diffusion coefficient (and, accordingly, an anomalously small kink friction coefficient).

As mentioned, the diffusion coefficient of kinks in real polyethylene cannot be larger than in our model polymeric crystal with united atoms. It follows that the population of kinks in a polymeric crystal should be substantially larger than that given by (17). For the diffusion coefficient to be smaller by eight orders of magnitude, it is sufficient to increase the density of kinks by four orders of magnitude with respect to the value calculated by (17). This still gives a fairly low density, $n_0 \sim (4.2 \times 10^{-1}–4.6 \times 10^{2})$ cm^{-1}, which corresponds to $5.4 \times 10^{-6}–6 \times 10^{-3}$ kinks per C=O dipole.

The following suggestions can be made to explain the inapplicability of (17) to the density of kinks in polyethylene. This formula was obtained in the approximation of "low temperatures" (6), which is fulfilled even at temperatures exceeding the melting point [see (8)]. It is, however, known, even from the early experiments on dielectric relaxation in polyethylene [39], that, in the temperature interval corresponding to the dielectric loss peak, strong asymmetric broadening of the crystal lattice of polyethylene is observed. More recent studies [40, 41] substantiated the hypothesis [42] that polyethylene melting occurred in two stages: first, the substance experienced the transition to the "rotation" phase (orientation disordering of atoms of all chains while chain axes remained parallel to each other) and the lattice changed to hexagonal, and then, melting proper occurred. If we assume that the first phase transition is caused by the appearance of pairs of torsional kinks of opposite topological signs, then the dielectric α_c relaxation process is an indicator of this phase transition, just as α_a relaxation is an indicator of vitrification. This hypothesis is substantiated by the observation that the experimental dielectric intensity of the process at a constant volume [37] decreases as temperature increases much more sharply than $1/T$, whereas the Kubo relaxation theory [43] predicts the dependence $\delta\varepsilon = \varepsilon_0 - \varepsilon_\infty \sim n_d\mu^2/k_BT$ (here, $n_d = N_d/V$ is the ratio between the number of dipoles in the crystalline fraction to the sample volume and μ is the absolute electric dipole moment value). Hence, it follows that a part of the dipoles "leaves" the crystalline phase, in which only kinks can exist, as temperature increases. This presupposes at least crystalline fraction disordering caused by torsion, that is, the phase transition from the orthorhombic structure to hexagonal. Clearly, the density of kinks should then be much higher than that predicted by (17).

7. CONCLUSION: ANSWERS AND QUESTIONS

In this work, we, for the first time, determined the diffusion coefficients of point structural defects by statistical measurements in a three-dimensional molecular dynamics model of a heated polymeric crystal. We found that, within a 25% error of measurements, the diffusion coefficients of a simple tension defect (the dynamics of which reduces to the integrable sine-Gordon equation in the continuous model in the approximation of immobile neighboring chains) and a torsion-with-tension defect (for which there is no soliton solution in the same approximation) were equal. A comparison of the obtained value with theoretical estimates for the integrable sine-Gordon model and for the nonintegrable (but admitting soliton-like solutions) ϕ^4 model and other numerical simulation results showed that friction of solitons in the system under consideration (as, it appears, in all real physical systems) is determined by its difference from the integrable model system (in the sine-Gordon model, there is no viscous friction at all).

The determination of the dipole moment autocorrelation function in the molecular dynamics experiment allowed us to directly check the validity of the theory of α_c relaxation in weakly oxidized polyethylene suggested in [12]. In the absence of torsion-with-tension defects, the dipole moment autocorrelation function was only indicative of the occurrence of one fast β process. In the presence of a kink in the chain, we observed an additional slow decrease in this function. Its kink part was well described by the formula obtained in [11] at a given kink density without any adjustment of the m^* and γ parameters. Namely, we theoretically estimated mass m^* for our model, and the γ friction coefficient was calculated by Einstein equation (11) from the m^* mass value and the diffusion coefficient determined in an independent molecular dynamics experiment.

We showed that the available physical experimental data [36, 37] were insufficient for independently determining kink density n_0 and kink diffusion coefficient D and could only be used to calculate the Dn_0^2 product. In [12], the *a priori* use of Eq. (17), which predicted an anomalously small value for the density of kinks, required assuming an anomalously large D value. The process under consideration, however, occurred close to the melting point, where the "approximation of low temperatures" used to obtain (17) was inapplicable. Obtaining realistic kink diffusion coefficients required increasing the density of kinks by four orders of magnitude in comparison with the value predicted by (17). It follows that there is good reason to consider the hypothesis suggested in [11] valid for dielectric α_c relaxation in weakly oxidized polyethylene provided assumption (17) is abandoned.

To completely substantiate the model suggested in [11], one must theoretically describe the experimental dependences of the central frequency of the process, the shape of the dielectric loss curve, and the temperature T and volume V dependences of dielectric intensity. This would require constructing the statistical thermodynamics of a polymeric crystal, that is, deriving the equation of state and calculating the $n_0(V, T)$, $m^*(V, T)$, and $\gamma(V, T)$ dependences.

APPENDICES

A. The Soliton Model of α_c Relaxation

After the α_c peak in polyethylene was detected, several molecular models of the process were suggested and its more detailed experimental study was undertaken (see reviews [4, 9] and the references therein). It was found that, in samples with a high degree of orientation of chains, the α_c peak was absent if the vector of the applied electric field was parallel to chains and appeared when the field was rotated through 90°. It was noted in studying α_c relaxation in samples with different chain lengths in crystallites that, starting with chain lengths of approximately 100 Å, the relaxation time (the central dielectric loss peak frequency) was virtually independent of the length of chains [36].

The mechanism of α_c relaxation was therefore associated with rotation of dipoles localized in the crystalline polymer fraction as a result of rotation of small (of the order of 80 CH_2 groups) chain segments about their axes. The mechanism of relaxation based on the hypothesis of diffusion of kinks in chains of the crystalline polymer fraction was independently, suggested in [10, 11].

This hypothesis explains certain special features of the process which are not described by the other theories. Indeed, kinks do not exist in crystals with short chains (the square of their width is proportional to the ratio between intra- and interchain rigidities), and the chain should rotate as a whole. Hence, it directly follows that, in paraffins, the rate of the process should decrease as chain length increases in short-chain compounds, whereas the central peak frequency should be independent of the chain length if this length exceeds a certain value (kink width). Precisely this is observed experimentally. A sharp decrease in the dielectric intensity (in the difference of the static and high-frequency permittivities $\delta\varepsilon = \varepsilon_0 - \varepsilon_\infty$) as temperature increases [37] can be explained by the onset of crystalline phase disordering, outside which such kinks also do not exist. This hypothesis, in contrast to the other, can explain [44] the observation of the α_c peak during dielectric relaxation of weakly oxidized polyethylene and isotactic polypropylene and its absence in syndiotactic polypropylene and isotactic polystyrene.

An attempt was made [12] to construct a quantitative model of dielectric α_c relaxation in polyethylene with the use of this hypothesis. The model was based on the assumption [11] that a decrease in the dipole moment autocorrelation function with time (the Fourier transform of its derivative was proportional to complex permittivity $\varepsilon^*(\omega)$ [43]) takes place at temperature T due to Brownian movement of kinks with mass m^*, friction coefficient γ, and linear density (the number of kinks per unit chain length) n_0. Kinks were treated as point objects which did not interact with each other (in conformity to the phenomenological model of an "ideal gas of phonons and kinks" suggested in [1]). When a kink approached a dipole, instantaneous dipole inversion occurred.

Under these assumptions, an analytic equation for the dipole moment autocorrelation function was obtained [11]. The equation contained three adjustment parameters, $\delta\varepsilon$, $B = 4n_0\sqrt{k_B T/\pi m^*}$, and γ, that were not determined by the suggested theory and depended on volume V and temperature T of the sample.

B. Theoretical Estimation of the Mass of a Torsion-with-Tension Kink

The sine-Gordon equation is well known to be invariant with respect to Lorentz transformations, and the energy of its kink solutions as a function of velocity has the form of the energy of a relativistic particle. Accordingly, the mass of a kink is understood to be the energy of a resting kink divided by the square of the sound velocity in an isolated chain. The total energy of one chain of our crystal in the continuous approximation of the model of immobile neighboring chains is [16]

$$H = \int \frac{dx}{l_0 s_0}\left[I_\phi \frac{\phi_t^2}{2} + I_\psi \frac{\psi_t^2}{2} + K_\phi \frac{\phi_x^2}{2} + K_\psi \frac{\psi_x^2}{2} + A(1 - \cos\phi\cos\psi) + B(1 - \cos 2\phi)\right], \tag{18}$$

where ψ is the displacement of the atom along the chain axis from its equilibrium position divided by the period of the chain; ϕ is the angle of the atom in the cylindrical system of coordinates with the axis that coincides with the chain axis; and I_ϕ, I_ψ and K_ϕ, K_ψ are the inertial and rigidity chain parameters, respectively, "torsional" (with respect to ϕ) and "longitudinal" (with respect to ψ). The relation between these constants and crystal numerical model parameters was obtained in [16],

$$I_\phi = m\left(\frac{l_0 c_0}{2}\right)^2, \quad I_\psi = m\left(\frac{l_0 s_0}{\pi}\right)^2 \tag{19}$$

$(s_0 = \sin(\theta_0/2), c_0 = \cos(\theta_0/2))$;

$$K_\phi = l_0^2(K_1 + 9K_3), \tag{20}$$

$$K_\psi = \left(\frac{2 l_0 s_0^2}{\pi c_0}\right)^2 K_\theta. \tag{21}$$

In the same work, the $A \approx 6.54 \times 10^{-2}$ kcal/mol and $B \approx 2.07 \times 10^{-1}$ kcal/mol values were estimated.

It is easy to see that the ϕ and ψ variables are separable in the linearized system equations. Two dispersion curve branches with two sound velocities correspond to these variables: "torsional" sound velocity $v_\phi = \sqrt{K_\phi/I_\phi} \approx 7.63$ km/s and "longitudinal" sound velocity $v_\psi = \sqrt{K_\phi/I_\phi} \approx 14.70$ km/s, which are maximum velocities of propagation of small-amplitude perturbations along ϕ and ψ, respectively.

Substituting the exact solution for the static simple tension kink into (18),

$$\psi(t) = 4\arctan\exp(x/L_\psi)$$

($L_\psi = \sqrt{K_\psi/A} \approx 33.5(c/2)$ is the half-width of a static kink), we easily obtain the static kink energy

$$H_{\psi^0} = \frac{8}{(2\pi)^2} 2m \frac{2l_0 s_0}{L_\psi} v_\psi^2.$$

For the numerical values of our model, we have energy and mass estimates $H_{\psi^0} \sim 17.5$ kcal/mol and $m^*_{\psi^0} = H_{\psi^0}/v_\psi^2 \approx 0.56 \times 10^{-27}$ kg (which is approximately two times smaller than the atomic mass unit), respectively.

Substituting the approximate solution for a static torsion-with-tension kink obtained in [16] into (18),

$$\phi_1 = 2\arctan\exp\left(\frac{x}{L_\phi}\right), \qquad (22)$$

$$L_\phi = \sqrt{K_\phi/4B} \approx 5.08(c/2);$$

$$\psi_1 \qquad (23)$$

$$= \begin{cases} 4\arctan\exp(x/L_\psi - \ln\tan(3\pi/8)), & x - vt < 0, \\ 4\arctan\exp(x/L_\psi + \ln\tan(3\pi/8)) - \pi, & x - vt \geq 0, \end{cases}$$

we obtain the kink energy

$$H_{\psi+\phi} = H_\psi - H_\phi,$$

$$H_\psi = H_{\psi^0}\left(1 - \tanh\ln\tan\frac{3\pi}{8}\right) \approx 5.13 \text{ kcal/mol},$$

$$H_\phi = \frac{8}{(8s_0/c_0)^2} 2m \frac{2l_0 s_0}{L_\phi} v_\phi^2 \approx 8.39 \text{ kcal/mol},$$

and the kink mass

$$m^* = \frac{H_\psi}{v_\psi^2} + \frac{H_\phi}{v_\phi^2}$$

$$\approx (0.164 + 1.002) \times 10^{-27} \text{ kg} \approx 1.166 \times 10^{-27} \text{ kg}.$$

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (projects nos. 01-03-33122 and, in part, 00-15-97431); E.A.Z. thanks the Russian Academy of Sciences Commission for Work with Young Scientists (grant 123 of the 6th Expertise Competition in 1999). The authors are grateful to L.S. Zarkhin, A.I. Musienko, and S.N. Chvalun for valuable discussions. Computational facilities (MBC–100M supercomputer) were provided by the Interinstitution Supercomputer Center of the Russian Academy of Sciences.

REFERENCES

1. J. A. Krumhansl and J. R. Schriffer, Phys. Rev. B **11**, 3535 (1975).
2. J. F. Currie, J. A. Krumhansl, A. R. Bishop, and S. E. Trullinger, Phys. Rev. B **22**, 477 (1980).
3. T. Schneider and E. Stoll, Phys. Rev. B **22**, 5317 (1980); T. Schneider and E. Stoll, Phys. Rev. B **23**, 4631 (1981).
4. Y. Ishida, J. Polym. Sci., Part A-2 **7**, 1835 (1969).
5. G. M. Bartenev, Vysokomol. Soedin., Ser. A **41**, 789 (1999).
6. G. P. Mikhaĭlov, S. P. Kabin, and B. I. Sazhin, Zh. Tekh. Fiz. **27**, 2050 (1957) [Sov. Phys. Tech. Phys. **2**, 1899 (1958)].
7. A. H. Scott, D. J. Scheiber, A. J. Curtis, *et al.*, J. Res. Natl. Bur. Stand., Sect. A **66**, 269 (1962).
8. H. Sasabe and S. Saito, Rep. Prog. Polym. Phys. Jpn. **11**, 379 (1968).
9. R. H. Boyd, Polymer **26**, 323, 1123 (1985).
10. M. S. Mansfield, Chem. Phys. Lett. **69**, 383 (1980).
11. J. L. Skinner and P. G. Wolynes, J. Chem. Phys. **73**, 4015 (1980); J. L. Skinner and P. G. Wolynes, J. Chem. Phys. **73**, 4022 (1980).
12. J. L. Skinner and Y. H. Park, Macromolecules **17**, 1735 (1984).
13. K. J. Wahlstrand, J. Chem. Phys. **82**, 5247 (1985); K. J. Wahlstrand and P. G. Wolynes, J. Chem. Phys. **82**, 5259 (1985); K. J. Wahlstrand, Polymer **29**, 256, 263 (1988).
14. N. K. Balabaev, O. V. Gendel'man, M. A. Mazo, and L. I. Manevich, Zh. Fiz. Khim. **69**, 24 (1995).
15. E. A. Zubova, N. K. Balabaev, and L. I. Manevich, Zh. Éksp. Teor. Fiz. **115**, 1063 (1999) [JETP **88**, 586 (1999)].
16. E. A. Zubova, N. K. Balabaev, L. I. Manevich, and A. A. Tsygurov, Zh. Éksp. Teor. Fiz. **118**, 592 (2000) [JETP **91**, 515 (2000)].
17. D. W. Noid, B. G. Sumpter, and B. Wunderlich, Macromolecules **23**, 664 (1990).
18. D. Rigby and R. J. Roe, Macromolecules **22**, 2259 (1989).
19. M. P. Allen and P. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).
20. P. G. Khalatur, N. K. Balabaev, and A. S. Pavlov, Mol. Phys. **59**, 753 (1986).
21. O. M. Braun and Yu. S. Kivshar, Phys. Rep. **306**, 1 (1998).

22. F. Marchesoni, Phys. Lett. A **115**, 29 (1986).
23. N. R. Quintero, A. Sanchez, and F. G. Mertens, Eur. Phys. J. B **16**, 361 (2000).
24. P. Pasqual and L. Vazquez, Phys. Rev. B **32**, 8305 (1985).
25. P. Biller and F. Petruccione, Phys. Rev. B **41**, 2139, 2145 (1990).
26. Y. Wada and H. Ishiuchi, J. Phys. Soc. Jpn. **51**, 1372 (1982).
27. K. Fesser, Z. Phys. B **39**, 47 (1980).
28. Y. Wada and J. R. Schriffer, Phys. Rev. B **18**, 3897 (1978).
29. N. Theodoracopoulos, Z. Phys. B **33**, 385 (1979).
30. M. Ogata and Y. Wada, J. Phys. Soc. Jpn. **54**, 3425 (1985).
31. M. Ogata and Y. Wada, J. Phys. Soc. Jpn. **55**, 1252 (1986).
32. F. Marchesoni and C. R. Willis, Europhys. Lett. **12**, 491 (1990).
33. N. Theodoracopoulos and E. W. Weller, Phys. Rev. B **38**, 2749 (1988).
34. E. A. Zubova, Zh. Éksp. Teor. Fiz. **120**, 1027 (2001) [JETP **93**, 895 (2001)].
35. P. Bordewijk, Chem. Phys. Lett. **32**, 592 (1975).
36. C. R. Ashcraft and C. R. Boyd, J. Polymer Sci.: Polym. Phys. Ed. **14**, 2153 (1976).
37. J. A. Sayre, St. R. Swanson, and R. R. Boyd, J. Polym. Sci.: Polym. Phys. Ed. **16**, 1739 (1978).
38. M. Buttiker and R. Landauer, Phys. Rev. A **23**, 1397 (1981).
39. M. Takayanagi, T. Aramaki, M. Yoshino, and K. Hoashi, J. Polym. Sci. **46**, 531 (1960).
40. M. de Langen and K. O. Prins, Chem. Phys. Lett. **299**, 195 (1999).
41. M. A. Shcherbina, S. N. Chvalun, V. I. Selikhova, *et al.*, Vysokomol. Soedin., Ser. A **41**, 1768 (1999).
42. J. I. Frenkel, *Kinetic Theory of Liquids* (Nauka, Moscow, 1945; Clarendon, Oxford, 1946).
43. R. Kubo, J. Phys. Soc. Jpn. **12**, 570 (1957).
44. J.-L. Syi and M. L. Mansfield, Polymer **29**, 987 (1988).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 770–778.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 897–907.
Original Russian Text Copyright © 2002 by Solodovnikov.

SOLIDS
Structure

Spin-Lattice Relaxation in Solid ^{3}He in Weak Magnetic Fields

I. S. Solodovnikov

Kapitsa Institute for Physical Problems, Russian Academy of Sciences, ul. Kosygina 2, Moscow, 113307 Russia
e-mail: solodov@kapitza.ras.ru

Received December 7, 2001

Abstract—Spin–lattice relaxation times T_1 were measured for solid ^{3}He at temperatures of 0.22 to 0.73 K in a 44-Oe magnetic field. An increase in T_1 at temperatures higher than approximately 0.4 K was related to switching on the vacancy mechanism of atomic mobility in the crystal. At a melting curve minimum, in the region of predominance of exchange motions of atoms in the crystal, measurements of T_1 were performed in magnetic fields of 2 to 71 Oe. The data obtained in fields higher than 5 Oe were in agreement with the Cowan–Fardis theory. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The problem of vacancies in ^{3}He crystals has been studied in many works [1–5]. The data on the concentration of vacancies and their energy of formation obtained by various methods for ^{3}He crystals with high melting points, higher than approximately 1 K, agree with each other [6]. For crystals of a lower density, direct studies of vacancies by X-ray diffraction [1] are impeded by their low concentration. The energies of formation of vacancies obtained in [1] for crystals with melting points below approximately 0.6 K appear to be underestimated compared with the data obtained more recently [2]. In [3], vacancies were studied by measuring the pressure increase caused by heating a crystal of a constant density. The latest literature data were obtained by measuring the rate of porous membrane motion [4] and the mobility of negative charges [5]. A model that treated motion of charges in solid ^{3}He from the point of view of the energy band structure of vacancies was suggested in [7]. There is a large spread of data reported by various authors for the mobility of vacancies in the crystal and for the vacanson band width [8].

NMR measurements of crystals allow vacancies to be recorded by their mobility [9]. The special feature of solid ^{3}He is the presence of exchange atomic motions whose rate is independent of temperature and rapidly increases as the density of the crystal decreases. Measurements of spin–lattice relaxation time T_1 in solid ^{3}H give information about the mobility of atoms in the crystal lattice and allow vacancies to be observed either under "high-temperature" or "low-temperature" conditions [2]. Under low-temperature conditions, at $T \simeq 0.2$ K, vacancies determine the rate of spin relaxation in fairly high fields and interrelate the exchange energy reservoir and the lattice [2, 6]. An analysis of the data on T_1 obtained under low-temperature conditions led the authors of [10] to conclude that the energy of formation of vacancies decreased to zero in the limit of the lowest crystal density; this conclusion was not substantiated in more recent works. It appears that low-temperature T_1 data processing for crystals of a low density [10] is impeded because of a decrease in temperature to the melting point and, probably, because of a nonexponential (two-time) character of spin–lattice relaxation under these conditions.

At higher temperatures, in the "intermediate" and "high-temperature" regions, spin relaxation is due to modulation of dipole–dipole interactions by motions of atoms in the crystal [6]. At temperatures below approximately 0.6 K, the mobility of solid ^{3}He atoms is largely determined by exchange processes, which are quantum-mechanical tunneling of two or more atoms through crystal lattice energy barriers. For this reason, at intermediate temperatures, approximately from 0.2 to 0.6 K, the spin relaxation time is independent of temperature, and an "exchange plateau" is observed. Under high-temperature conditions, the rate of atomic motions in the crystal caused by the mobility of vacancies exceeds the rate of exchange processes. Spin–lattice and spin–spin relaxation and spin diffusion are then determined by vacancies; Reich [11] has experimentally measured the corresponding characteristics. Information about vacancies is usually obtained from spin–lattice relaxation times T_1 measured in strong fields [2]. It appears that an analysis of the T_1 data obtained under high-temperature conditions in strong magnetic fields involves difficulties caused by a complex, nonmonotonic temperature dependence of T_1 [12], which first decreases and then begins to increase as temperature rises above the plateau level. So far as we know, the experimental data on the energy of formation of vacancies obtained under high-temperature conditions are limited to crystals with melting points above 1.2 K [13, 14].

For less dense crystals, the region of purely vacancy high-temperature conditions gradually becomes very narrow and then completely disappears. For this reason, we deemed it interesting to perform spin–lattice relaxation time measurements in the region of lower temperatures corresponding to the onset of the transition from the exchange plateau to the high-temperature vacancy conditions. In this work, we measured T_1 in this transition temperature region in a low magnetic field for ^{3}He crystals with very low densities.

2. RELAXATION MODEL

The spin relaxation rate in solids with fast molecular motion can be described by the Bloembergen–Purcell–Pound theory equation [2, 6, 15]

$$T_1^{-1} = \frac{2M_2\tau_c}{3}\left(\frac{1}{1+\omega_0^2\tau_c^2} + \frac{4}{1+4\omega_0^2\tau_c^2}\right) \quad \text{at } \tau_c\sqrt{M_2} \ll 1, \tag{1}$$

where $\omega_0 = \gamma H_0$ is the Larmor frequency (γ is the gyromagnetic ratio), M_2 is the second van Vlek moment of the rigid lattice determined by dipole fields, and τ_c^{-1} is the frequency of modulation of dipole–dipole interactions by atomic motion. For a polycrystalline body-centered cubic phase and spin 1/2, we have $M_2 = 3.27\gamma^4\hbar^2 n^2$, where n is the concentration of spins. In weak magnetic fields H_0 at $\omega_0\tau_c \ll 1$, we have

$$T_1 = 0.3\frac{1}{\tau_c M_2}. \tag{2}$$

In weak magnetic fields, a decrease in the characteristic time of atomic motion τ_c caused by heating the crystal increases the spin-lattice relaxation time. The Bloembergen–Purcell–Pound model is usually applied to solid ^{3}He to describe spin relaxation under high-temperature vacancy conditions [2]. We assume that equation (2) in the limit of weak magnetic fields is also valid in the region of lower temperatures where exchange processes occur. Atomic motion in crystalline ^{3}He is caused by quantum-mechanical exchange processes and the mobility of vacancies,

$$\tau_c^{-1} = \omega_e + ax\omega_v, \tag{3}$$

where ω_e describes the frequency of exchange motions in the absence of vacancies (for simplicity, we omit the numerical coefficient of ω_e, which is of the order of one), x is the relative concentration of vacancies, and ω_v is the frequency of vacancy jumps into one of the neighboring lattice nodes. These frequencies are of the order of magnitude of $\omega_e \sim 10^8$ s^{-1} and $\omega_v \sim 10^{10}$ s^{-1} [6]. Coefficient α can be determined within the framework of a particular model of vacancy motions; according to [6, 13], $\alpha \approx 10$. We assume that all vacancies randomly move over the crystal; this assumption ignores the possibility that low-energy vacancies near the band bottom can be localized in the lattice to form a magnetic polaron [8, 16]. As in [6], we assume that the motion of vacancies in the cubic phase of solid ^{3}He occurs by tunneling without thermally activated overcoming of the barrier. The ω_v frequency is then independent of temperature. We also do not consider possible influence of vacancies that appear when the crystal is heated on the rate of exchange atomic motion processes. The deceleration of exchange processes, which might be caused by an effective increase in crystal density accompanying the appearance of vacancies in it, is likely to be proportional to the concentration of vacancies and can be included in the last term in (3). The magnetic field can be considered weak if $H_0 \ll \omega_e/\gamma \sim 5$ kOe. In low magnetic fields, because of the smallness of the heat capacity of the Zeeman energy reservoir compared with the exchange reservoir, the low-temperature vacancy mode is not observed in the $T_1(T)$ dependence, and the exchange plateau persists even to the lowest temperatures, to the melting point.

We will follow the model of a narrow energy band of vacancy states, which is usually [13, 14] applied to consider experimental data obtained under high-temperature vacancy conditions. The width of the energy band of vacansons expressed in degrees is then assumed to be much smaller than the temperature, and the temperature dependence of the concentration of vacancies is described by the Arrhenius law $x = \exp(-\Phi/T)$, where activation energy Φ equals the energy of formation of a vacancy in the center of the narrow band. In this approximation, the spin relaxation time is written as

$$T_1 = T_e + T_0\exp(-\Phi/T). \tag{4}$$

The first term in (4) is independent of temperature and corresponds to the region of the low-temperature exchange plateau. The second term describes an increase in T_1 as temperature increases up to the melting point. The temperature dependences of time T_1 obtained in this work are used to determine the activation energy of vacancies Φ and preexponential factor T_0, which has the dimension of time. Using T_0 in the equation $T_0 = 0.3\alpha\omega_v/M_2$ allows the frequency of vacancy motions and the width of the energy band corresponding to vacancies to be estimated. The T_e time in (4) characterizes exchange processes and only depends on the density of crystals. According to [10, p. 66; 17, p. 216], the exchange interaction frequency depends on the density as $\omega_e \propto V_m^{18}$, where V_m is the molar volume of the crystal. The strong power dependence is related to a rapid deceleration of exchange processes as the density of solid ^{3}He increases. As $M_2 \propto V_m^{-2}$, we have $T_e \propto V_m^{20}$. According to the measurements performed by Devoret *et al*. [18], $T_e \approx T_1 \propto V_m^{19\pm1}$ in the region of the exchange plateau in a 0.9-kOe magnetic field. In this

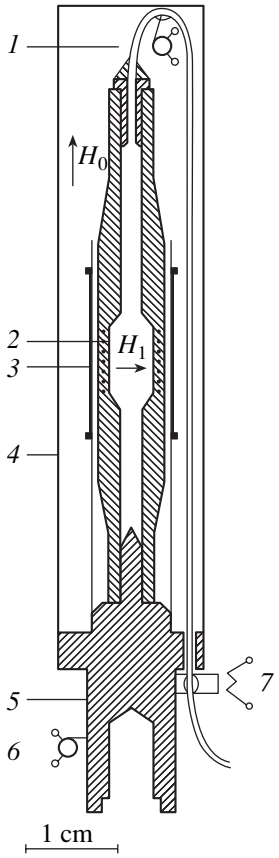

Fig. 1. Scheme of experimental cell: *1*, thermometer on a capillary for filling; *2*, receiving coil of flow transformer; *3*, RF field coil; *4*, cell screen; *5* copper cold conductor; *6*, thermometer; and *7*, heater on a thermal bridge.

work, we check these data on the crystal density dependence of T_e.

In Section 6, we describe measurements of the magnetic field dependence of T_1 in the region of the predominance of exchange processes, at a minimum of the ^{3}He melting curve. The results of these measurements allow us to check the conclusion drawn in theoretical work [19] on the presence of a root singularity in the $T_1(H_0)$ dependence in weak magnetic fields, which differs from dependence (1) predicted by the Bloembergen–Purcell–Pound model.

3. EXPERIMENTAL

The cell used in experiments with solid ^{3}He is shown in Fig. 1. The high-pressure chamber vessel was made of paper-reinforced Stycast-1266 epoxide resin. The inner cell volume was about 0.3 cm^3. A copper bushing was glued into the top part of the cell; into this bushing, a capillary for filling was soldered. A copper cold conductor needle was glued into the bottom part; the conductor was in thermal contact, provided by a threaded connection, with the solution chamber of a ^{3}He–^{4}He refrigerator. The temperature of the cell was controlled by a resistance thermometer made of a 70-Ω, 0.125-W Matsushita carbon resistor mounted on the cold conductor of the cell. This thermometer was used to stabilize temperature of the solution chamber. The thermometer was calibrated against ^{3}He condensation thermometer readings and the magnetic susceptibility of cerium–magnesium nitrate. The calibration was checked by the spin signal of solid ^{3}He. The deviations of the melting points of crystals measured in this work from the values corresponding to the melting curve of ^{3}He did not exceed 1% of temperature values. The thermometer mounted in the top part of the cell on the capillary for filling was used to determine the beginning and end of crystal growth. The resistance thermometers were connected to bridge measuring schemes (cryobridges), the output from which could be written to a computer. The capillary for filling, 0.14-mm inside diameter, was made of a copper–nickel alloy and was in thermal contact with the cell cold conductor through a thermal bridge, which was a brass plate with a 0.3-mΩ residual resistance. To prevent plugging of the capillary by crystals, an electric heater was mounted on the plate. Further, the filling line had a similar thermal bridge at the site of its contact with the evaporation chamber of the solution refrigerator, whose temperature was about 0.5 K. In some experiments, there was one more thermal bridge at the contact with the solution chamber.

The ^{3}He gas used for growing crystals contained not more than 0.02% ^{4}He. Pressure was controlled by an electronic pressure gauge and visually by an arrow manometer. The noise level of the electronic pressure gauge did not exceed 1 mbar. The absolute accuracy of determining pressure, which could be checked by the pressure at a melting curve minimum, was about 40 mbar.

The longitudinal magnetization of ^{3}He nuclei was detected by a magnetometer based on an RF SQUID [20, 21]. Stationary magnetic field H_0 was trapped by a niobium cell screen tube during cell cooling. Radiofrequency (RF) circularly polarized field H_1 used for NMR excitation was created by two crossed saddle superconducting coils. Field H_1 was calibrated by the shift of the resonance line frequency under the conditions of applying a strong continuous RF field detuned from the resonance; this shift was similar to the Bloch–Siegert shift [22, ch. II, §1]. An $H_1 = 0.02$ Oe continuous RF field at the NMR frequency of protons was constantly applied during the whole NMR experiment. This field saturated the proton signal of cell walls. The output from the SQUID was written to a computer, which also performed commutation of the RF field and controlled the RF field frequency sweep near the ^{3}He NMR frequency. The ^{3}He resonance NMR line was usually sequentially passed many times with signal accumulation and averaging on a computer.

The receiving SQUID magnetometer coil was glued directly into a cell wall. For this reason, apart from the magnetic signal caused by changes in the receiving coil contour area, the magnetometer was also sensitive to pressure changes in the cell. This allowed us to control the stability of pressure inside the cell filled by a crystal. The melting temperature can be determined from a sharp increase in pressure when the crystal was heated.

4. GROWING CRYSTALS

A 34-bar pressure necessary for growing crystals was created using a high-pressure chamber (gasifier) of volume 4 cm^3. The volume of the high-pressure gas system, which had room temperature and included an arrow manometer, an electronic pressure gauge, and a valve, was about 14 cm^3. First, ^{3}He (0.08 mol) was condensed into the cell and gasifier. The gasifier was then heated to 30–40 K to create the required pressure. Gasifier heating could be controlled automatically by a computer; the control parameters were electronic pressure gauge output and readings of the resistance thermometer on the capillary for filling the cell. Pressure could be stabilized or varied at the required rate.

Prior to growing crystals and while pressure was increased to the required value, cell cold conductor temperature (further, this temperature point will be referred to as "*A*") was stabilized somewhat above the T_m melting temperature corresponding to the selected pressure; the heaters of thermal bridges at thermal contact sites of the capillary for filling the cell were switched on. Next, pressure was stabilized, and temperature *A* controller was switched to a value about T_m – 60 mK. The onset of crystal growth was determined as the moment at which the temperature measured by the thermometer on the capillary for cell filling began to increase (further, this thermometer is called "*B*") because of inflow of a more heated liquid into the cell. Usually, at $T_m < 0.6$ K, a crystal began to grow within 4–10 min after cooling the cold conductor of the cell and stabilizing the specified T_A temperature. Apparent "entry" of liquid ^{3}He into the region of the existence of crystals in the phase diagram with a $P - p_m(T_A) \le 0.5$ bar value (P is the pressure in the cell, and $p_m(T_A)$ is the pressure of crystal melting at temperature T_A) cannot be strictly interpreted as a metastable state of the liquid because of the complexity of heat transfer in liquid ^{3}He close to the cold conductor needle with possible arising of convection. It appears that, in separate experiments, crystals began to grow in the capillary for cell filling somewhere in the vicinity of thermometer *B* and plugged the capillary for a short time. This manifested itself by T_B and SQUID output jumps caused by pressure oscillations in the cell. At low melting temperatures, $T_m \le 0.55$ K, crystals usually began to grow not on the cold conductor needle at the bottom of the cell. The moment when the growing crystal reached the needle was detected by a short-term cold conductor temperature T_A rise. After this, the crystal gradually filled the cell from bottom to top. Clearly, the rate of crystal growth was then determined by the removal of the heat of crystallization from the growth front through the column of the already grown crystal to the cold conductor needle. Crystal growth could be monitored by measuring the intensity of the NMR signal from the solid phase (Fig. 2). The duration of crystal growth from the onset of the process on the needle was 10 min at T_m = 0.5 K to 50 min at T_m = 0.74 K. When crystal growth stopped, a sharp decrease in T_B was observed. An increase in the NMR signal at this moment was caused by crystal cooling to cold conductor temperature T_A.

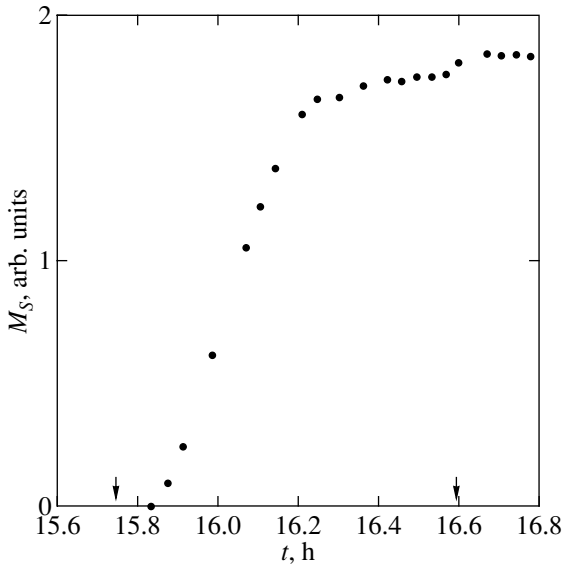

Fig. 2. Solid phase signal during crystal growth. Pressure P = 34 bar ($T_m \approx 0.74$ K), T_A = 0.66 K, and arrows indicate crystal growth beginning and termination.

In the early series of experiments, the capillary for cell filling had three thermal contacts (three bridges and heaters on them), with the cell cold conductor, with the solution chamber, and with the evaporation chamber of the solution refrigerator. The electric heaters connected to a common circuit were simultaneously switched off approximately 10 min after crystal growth in the cell was completed. We found that the pressure in the cell began to decrease soon afterward. Clearly, this was caused by plugging of the capillary for cell filling by crystals; the plug was formed at the contact with the evaporation chamber. Because the crystals were highly plastic, a decrease in pressure in the top part of the cell was transferred to the measuring volume in the center of the cell. In these experiments, the crystal was first held at T_m – 70 mK for 1 h to attain pressure stability and then "annealed" at T_m – 10 mK for one more hour, which allowed NMR measurements at various temperatures to be performed under constant density conditions. The total decrease in pressure in the cell after crystal growth amounted to 50–70 mbar. Further, dur-

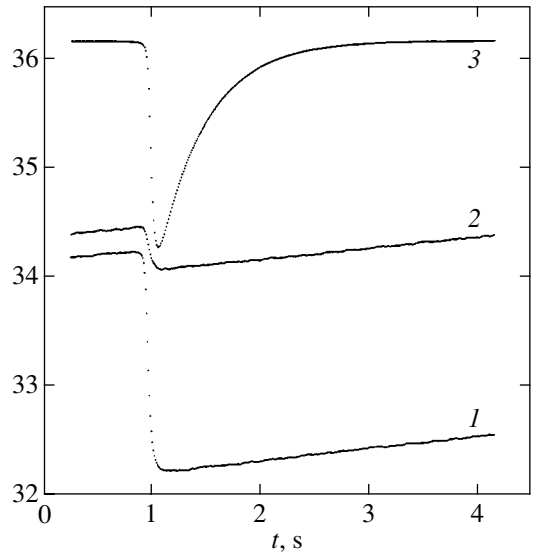

Fig. 3. SQUID output records for resonance passage. H_0 = 44 Oe, H_1 = 16 mOe, RF field frequency sweep rate 3.27 kHz/s. Curves *1* and *2* correspond to liquid ^{3}He at T = 0.53 K, P = 30.3 bar, dP/dt = 1 mbar/s; curve rises are caused by an increase in pressure within the cell. Record *2* was made 1.5 min after record *1*, and a decrease in signal intensity was caused by incomplete relaxation of spins after passage *1*. Curve *3* corresponds to solid ^{3}He at T = 0.50 K and P = 30.5 bar; this curve was obtained by averaging over 20 passages through the resonance line. Time T_1 was determined from the curve portion to the right of vertical lines.

ing NMR measurements, temperature was decreased in steps to its lowest value (about 0.25 K). The stability of the output from the SQUID was evidence of a constant crystal density in the measuring volume of the cell. In these experiments, pressure usually again began to decrease when crystals were heated after cooling.

In a later series of experiments, controlled crystal growth conditions were created using only two thermal contacts in the filling line, between the capillary for filling and the cell and the evaporation chamber. The heater at the heating place of the contact with the evaporation chamber was constantly switched on. About 0.5 h after crystal growth was completed, at $T_A \approx T_m$ – 55 mK, current through the electric heater at the contact between the filling line and the cell cold conductor was smoothly decreased from the initial power of 5 μW with the use of a capacitor bank with a 1-h time constant. In 1 h, the heater was switched off, the cell was heated to T_m – 10 mK, and the crystal was held at this temperature ("annealed") for 0.5 h. After such a procedure for crystal preparation, pressure remained stable during NMR measurements in cooling to the lowest temperature and subsequent heating to T_m. The T_m melting point was determined in all experiments from the onset of a sharp increase in pressure in the cell when the crystal was heated.

Experiments with solid ^{3}He at a melting curve minimum in the phase diagram are described in Section 6. In these experiments, cold conductor temperature T_A was stabilized at a minimum point temperature (0.315 K) prior to growing crystals, and pressure was increased at a rate of 0.14 mbar/s. Crystals began to grow at $P - p_m(T_A)$ = 0.10–0.15 bar. Crystals filled the experimental cell in about 3 min (from the beginning of growth on the cold conductor needle to a decrease in T_B); during growth, pressure was noticeably higher than at the minimum point. Pressure was then decreased, and crystals melted almost fully. The absence of the solid phase in the central measuring cell part was controlled by measuring the NMR signal. A small crystal still remained on the cell bottom on the cold conductor needle. Further, the temperature of the gasifier began to increase under computer control; used as the control parameter was the T_B temperature. The crystal began to grow and filled the whole cell in about 20 min; pressure during growth did not exceed minimum point pressure by more than 1–2 mbar.

5. T_1 MEASUREMENTS. THE CONTRIBUTION OF VACANCIES

The excitation of the spin system after which spin-lattice relaxation was observed was performed by adiabatic fast passage of the NMR line. RF field H_1 and the rate of frequency sweep were selected to fulfill the adiabatic passage condition $(\gamma H_1)^2 \gg d\omega/dt$, which ensured magnetization inversion (spin flip) during the passage [22, Ch. II]. The time of passage of the order of $\gamma H_1/(d\omega/dt)$ was much shorter than T_1. It follows that fast adiabatic inversion of spins occurred during line passage. This inversion was followed by comparatively slow relaxation of spin magnetization to its equilibrium value. The computer-controlled measuring cycle involved switching on circularly polarized RF field H_1 of 11 to 16 mOe at a frequency about 2 kHz below the resonance frequency, sweeping the RF field frequency through the resonance line at an about 3-kHz/s frequency during a period of about 4 s, and, finally, switching the RF field off. The SQUID output records describing the evolution of longitudinal ^{3}He spin magnetization during one cycle are shown in Fig. 3. In experiments with growing crystals, when the liquid and solid phases coexisted in the cell, fast spin relaxation of solid ^{3}He allowed us to identify the signal from ^{3}He nuclei in the solid phase. The T_1 values in solid ^{3}He under our experimental conditions equaled 0.3 to 0.5 s, and the spin system returned to the equilibrium state by the end of the measuring cycle. When T_1 was measured for solid ^{3}He, this cycle was repeated many times with computer signal accumulation; usually, cycles were repeated every 6 s. T_1 measurements described in this section were performed in a stationary magnetic field H_0 = 44 Oe (the resonance frequency equaled 142 kHz).

In this field, the magnetic moments of nuclei were fairly large, and, at a 10-min time of signal accumulation, we were able to measure T_1 with an accuracy of 0.1% or higher. Prior to T_1 measurements, crystals were held at the required temperature for 4 min or longer.

The T_1 times were determined by approximating $S(t)$ SQUID outputs recorded about 0.2 s after inversion time t_i with the use of the equation

$$S(t) = \text{const} - M_s \exp\frac{t_i - t}{T_1}. \quad (5)$$

The M_s value corresponds to twice the magnetic moment of solid ^{3}He nuclei equilibrium at a given temperature (in arbitrary units of SQUID outputs). The subtraction of adjustment curve (5) from the experimental time dependences of SQUID outputs allowed us to check whether or not spin–lattice relaxation was exponential.

The T_1 times obtained for the crystal with $T_m = 0.51$ K at various temperatures are shown in Fig. 4. At temperatures below approximately 0.4 K, times T_1 did not depend on temperature. In this temperature region, relaxation was determined by exchange processes of atomic motions in the crystal. We described the experimental temperature dependences of T_1 by (4). The T_e time, which characterized exchange processes, only depended on the density of the crystal. The obtained density dependence of time T_e, $T_e \propto V_m^{20}$, was in agreement with the density dependence of exchange interaction frequency $\omega_e \propto V_m^{18}$ [10, 17] and the results of T_1 measurements in field $H_0 = 0.9$ kOe [18].

An increase in time T_1 observed when crystals were heated above 0.4 K was evidence of acceleration of atomic motions in crystals. It was interpreted as switching on the vacancy mechanism of motion. The adjustment parameters of (4) that describe the contribution of vacancies are preexponential factor T_0 and activation energy Φ. Time T_0 was independent of crystal density within the accuracy of our measurements (Fig. 5a); its mean value for eight samples with melting points of 0.51 to 0.735 K was $T_0 = 59 \pm 12$ s. The densities of the crystals with melting points of 0.51 and 0.735 K differed by about 2%. In what follows, we neglect a possible dependence of time T_0 on crystal density. To determine the activation energy, the experimental $T_1(T)$ dependences were again treated by Eq. (4), in which the T_0 parameter was fixed at the mean value given above. This procedure allowed us to decrease the statistical spread of activation energy values and reveal its dependence on the density of crystals. The obtained Φ values are shown in Fig. 5b. The activation energies of vacancies increased as the density of crystals grew.

The $\Delta T_1 = T_1(T) - T_e$ values for the crystal with the highest density, $T_m = 0.735$ K, are shown in Fig. 6 in semilogarithmic coordinates as functions of reciprocal temperature. The T_e value for this crystal (0.30 s) was determined by averaging T_1 times measured at $T < 0.4$ K. The slope of this dependence gives the activation energy of vacancies Φ for the crystal with a given den-

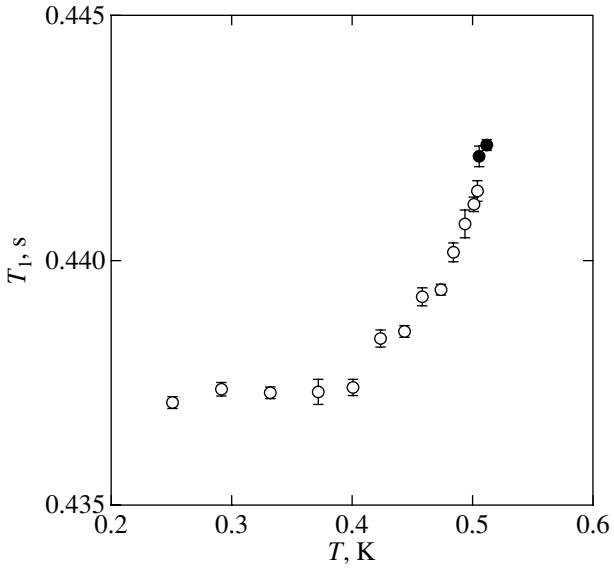

Fig. 4. Temperature dependence of spin relaxation time T_1 for a crystal with $T_m = 0.51$ K (crystal growth pressure was 30.5 bar); (○) cooling and (●) crystal heating at the final stage of measurements.

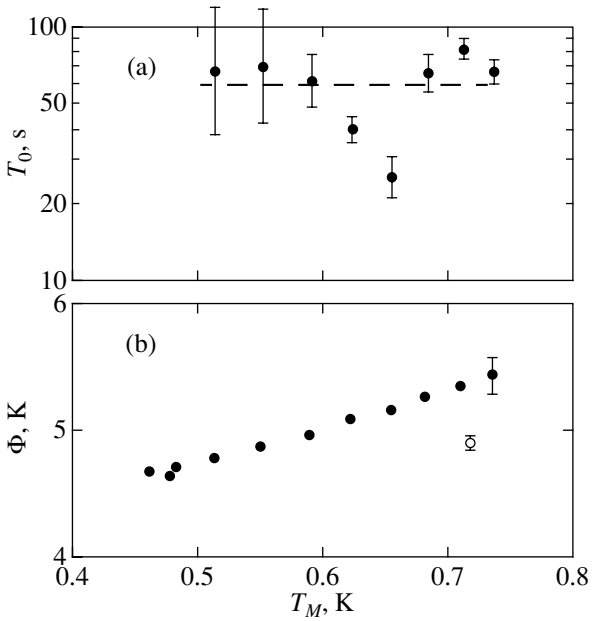

Fig. 5. (a) Preexponential factor obtained in treating $T_1(T)$ dependences by (4) as a function of crystal melting temperature. (b) Activation energies of vacancies. These values were determined using the mean preexponential factor value shown by dashes in Fig. 5a. For the $T_m = 0.735$ K point, the error caused by inaccuracy of determining T_0 is shown; (○) the data [3] obtained from the temperature dependence of pressure.

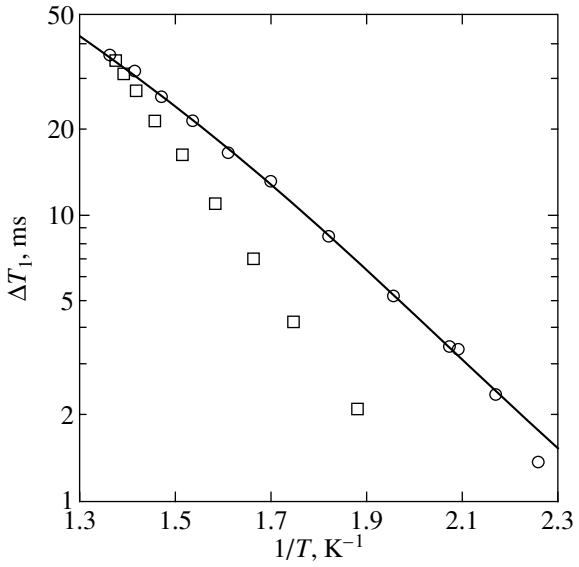

Fig. 6. (□) $T_1 - T_e$ values for the densest crystal with T_m = 0.735 K studied in this work; (○) increase in T_1 from the exchange plateau region to the melting point as a function of reciprocal melting temperature. The solid curve was obtained using the mean T_0 value and the $\Phi(T_m)$ dependence shown in Fig. 5b.

sity. Open circles in Fig. 6 are the $T_1(T_m) - T_e$ values as functions of the reciprocal melting temperature for each separate crystal for which the temperature dependence of T_1 was measured. The $T_1(T_m)$ value was found by extrapolating the $T_1(T)$ dependence for each crystal to its melting temperature. This curve describes an increase in T_1 from the exchange plateau region to the melting point and shows how the vacancy contribution to T_1 at this point decreases as the melting temperature lowers. The smaller slope of this dependence is caused by a decrease in the activation energy with decreasing crystal density. The solid curve in Fig. 6 describes the $T_0\exp(-\Phi/T_m)$ value; it was obtained using the smoothed dependence of the activation energy on the melting temperature of crystals shown in Fig. 5b and the mean preexponential factor T_0 value. The $d\Phi/dT_m$ derivative value can be used to calculate the derivative of the activation energy of vacancies with respect to crystal density. According to our measurements, $d\Phi/dV_m = -1.4$ K/(cm^3/mol).

These results can be compared with the data on the activation energy of vacancies obtained by various authors taking into account that equations of type (4) within the framework of band models of vacansons often contain power preexponential temperature factors. The activation energies obtained in this work are higher than those reported in [2–4] for $V_m > 24$ cm^3/mol by about 0.4 K. Our Φ values are approximately 0.5 K lower than the values extrapolated to larger molar volumes from the data on crystals with $T_m > 1.2$ K (see [6, Fig. 3.13] and [13, Fig. 2]).

The obtained T_0 [preexponential factor in (4)] value can be used to estimate the characteristic frequency of vacancy movements and the width of the energy band of vacancies. Set coefficient α in (3) equal to 10 [6, 13]. Calculations of the band width of vacansons by [2, 13]

$$\Delta = 4(z-1)^{1/2}\hbar\omega_v/k_B \qquad (6)$$

(here, z is the number of nearest neighbors and k_B is the Boltzmann constant) then give $\Delta = 0.6$ K. The authors of [13] analyzed the data on T_1 that they obtained under high-temperature vacancy conditions to determine the frequency of vacancy motions, $\omega_v \simeq 0.5 \times 10^{10}$ s^{-1}; according to (6), this corresponded to the vacanson band width $\Delta \approx 0.4$ K. In [4], the vacanson band width in solid ^{3}He was found to be $\Delta = 3.5$–4 K at the lowest densities; this result was obtained from the data on porous membrane motions. Estimates of Δ show that, under our experimental conditions, $\Delta \gtrsim T$; it appears that our analysis of the $T_1(T)$ dependences within the framework of the narrow band model is not absolutely correct and should be considered a phenomenological approach. At the same time, the ambiguous situation with the development of a general band model of vacancies in solid ^{3}He prevents us from performing a rigorous analysis of the experimental data.

Note that time T_0 (which may be treated as the spin–lattice relaxation time of the hypothetical crystal with the concentration of vacancies of the order of one) is close to the volume spin relaxation time in liquid ^{3}He in the vicinity of the melting curve; this value is about 100 s [23]. Similar closeness was noted by Reich [11] for the preexponential factor in the equation for the spin diffusion coefficient in solid ^{3}He in the vacancy mode.

6. THE MAGNETIC FIELD DEPENDENCE OF T_1

The Bloembergen–Purcell–Pound model [15] is usually applied to describe spin relaxation in liquids and solids with fast molecular motions. For low magnetic fields, this model predicts a quadratic increase in T_1 as the field grows from the T_1 value in zero magnetic field [see (1)]. Under the conditions of our experiments, at $\omega_0\tau_c \approx \omega_0/\omega_e \lesssim 10^{-2}$, this increase in T_1 by (1) does not exceed a value of the order of 10^{-4} T_1, which is within the error of measurements. The Bloembergen–Purcell–Pound model, which considers spin relaxation in terms of fluctuations of dipole fields, actually takes into account mutual displacements only of the nearest neighbors and describes these displacements by one correlation time. Comparatively recently, Cowan and Fardis [19] considered the region of exchange atomic motion processes to analyze the behavior of dipole correlation functions at long times. An asymptotic behavior proportional to $t^{-3/2}$ was caused by dipole interac-

tions between spins spaced $\sim(Dt)^{1/2}$ apart, where D is the spin diffusion coefficient. The spectral density functions contained terms proportional to $\omega^{1/2}$, and the T_1 time in low magnetic fields contained a root term additional to its value in zero magnetic field [19],

$$T_1(\omega_0) = T_1(0) + 0.70\frac{n\hbar^2\gamma^4}{D^{3/2}}(T_1(0))^2\omega_0^{1/2}. \quad (7)$$

The available experimental data [19, 24] substantiate the existence of such a dependence. The corresponding measurements were, however, performed in fairly high magnetic fields, higher than 150 Oe, and the additional term in (7) exceeded 10% of $T_1(0)$. In our view, testing the model suggested in [19] required $T_1(H_0)$ dependence measurements in the weakest magnetic fields possible.

Under our experimental conditions, changing magnetic field H_0 required heating the experimental cell and the destruction of crystals. For this reason, the $T_1(H_0)$ dependences were measured with different crystals grown at the same phase diagram point, namely, at a melting curve minimum. Our goal was to obtain the most accurate absolute T_1 values in these experiments. Numerical simulations of resonance passage by the Bloch equations with $T_1 = T_2$ were performed to determine the shift of the experimental T_1 value related to the "wing" of the resonance line at a given RF field H_1 value. We also took into account the speed of SQUID magnetometer operation. The obtained correction was taken into account in determining T_1; its value in $H_0 \geq 5$ Oe fields did not exceed 0.9 ms. The T_1 values found for H_0 fields of 2 to 71 Oe are shown in Fig. 7 as functions of the square root of the magnetic field. In fields $H_0 > 5$ Oe, in which T_1 could be measured fairly accurately, a linear dependence of T_1 on $H_0^{1/2}$, which satisfied (7), was observed. The spin diffusion coefficient calculated from the slope of the dependence by (7) was $D \approx 1.3 \times 10^{-7}$ cm^2/s. This was close to the D value that could be obtained by extrapolating the experimental D values measured for low-density ^{3}He crystals [6, 19] to the density at the melting curve minimum. Such an extrapolation gives approximately (1.5–2) 10^{-7} cm^2/s.

Consider the question of a possible influence of the magnetic field dependence of T_1 on the vacancy contribution to T_1. As $D \approx 1/(n^{2/3}\tau_c)$, the second term in (7), which describes an increase in T_1 in field H_0, can be written as

$$\Delta T_1(\omega_0) \approx T_1(0)(\omega_0\tau_c)^{1/2} \approx (1/M_2)(\omega_0/\tau_c)^{1/2}.$$

This value increases when vacancies appear in the crystal. The ratio between $\Delta T_1(\omega_0)$ changes and changes in $T_1(0)$ given by (2) as a function of τ_c is about 2%. This allows us to ignore the dependence of T_1 on H_0 in considering the vacancy contribution to T_1 under the conditions of our experiments.

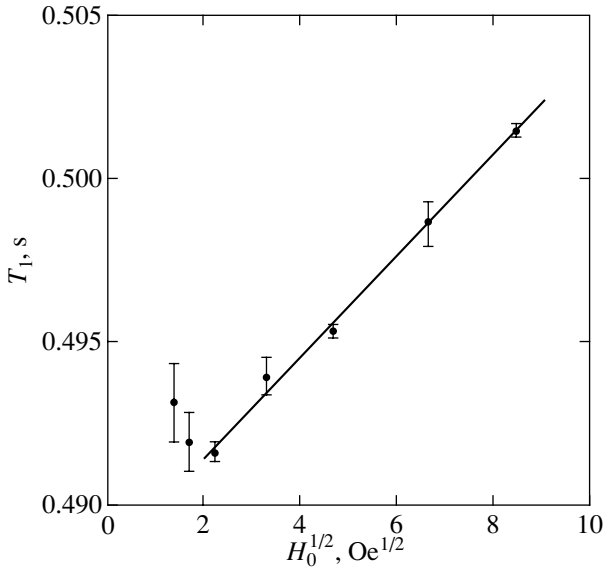

Fig. 7. Dependence of spin relaxation time on the square root of magnetic field for solid ^{3}He at a melting curve minimum. The slope of the dependence at $H_0 > 5$ Oe allows D to be estimated.

7. CONCLUSIONS

In this work, we measured spin–lattice relaxation times T_1 of nuclei in ^{3}He crystals with melting temperatures below 0.735 K in low magnetic fields. At temperatures above ≈0.4 K, T_1 values increased, which was explained by switching on the vacancy mechanism of atomic motions in crystals. The measurement results were treated using the activation law. The preexponential factors and activation energies of vacancies were obtained, and the derivative of the activation energy with respect to the density of crystals was estimated. The activation energy found within the framework of the narrow vacanson band model equaled 5.0 ± 0.15 K for the crystal with a 0.6 K melting point and was approximately 0.4 K higher than the most recent literature activation energy values [2, 4]. The obtained dependence of T_1 on magnetic field H_0 in the range from about 5 to 71 Oe, which was the highest magnetic field used in our experiments, was in agreement with the conclusion [19] of a root singularity in the $T_1(H_0)$ dependence in low magnetic fields. Predictions of this theory were in close quantitative agreement with our experimental data.

ACKNOWLEDGMENTS

The author thanks V.I. Marchenko and A.Ya. Parshin for discussions.

REFERENCES

1. S. M. Heald, D. R. Baer, and R. O. Simmons, Phys. Rev. B **30**, 2531 (1984).
2. M. E. R. Bernier and J. H. Hetherington, Phys. Rev. B **39**, 11285 (1989).
3. I. Iwasa and H. Suzuki, J. Low Temp. Phys. **62**, 1 (1986).
4. N. V. Zuev, V. V. Boiko, N. E. Dyumin, and V. N. Grigor'ev, J. Low Temp. Phys. **111**, 597 (1998).
5. Ya. E. Volokitin, A. I. Golov, and L. P. Mezhov-Deglin, Pis'ma Zh. Éksp. Teor. Fiz. **56**, 595 (1992) [JETP Lett. **56**, 578 (1992)].
6. A. Abragam and M. Goldman, *Nuclear Magnetism: Order and Disorder* (Clarendon, Oxford, 1982; Mir, Moscow, 1984), Vol. 1, Chap. 3.
7. A. F. Andreev and A. D. Savishchev, Zh. Éksp. Teor. Fiz. **119**, 403 (2001) [JETP **92**, 352 (2001)].
8. G. Montambaux, M. Heritier, and P. Lederer, J. Low Temp. Phys. **47**, 39 (1982).
9. M. Goldman, *Spin Temperature and Nuclear Magnetic Resonance in Solids* (Clarendon, Oxford, 1970; Mir, Moscow, 1972), Chap. 3.
10. M. Chapellier, M. Bassou, M. Devoret, *et al.*, J. Low Temp. Phys. **59**, 45 (1985).
11. H. A. Reich, Phys. Rev. **129**, 630 (1963).
12. S. R. Hartman, Phys. Rev. A **133**, 17 (1964).
13. N. Sullivan, G. Deville, and A. Landesman, Phys. Rev. B **11**, 1858 (1975).
14. M. E. R. Bernier and G. Guerrier, Physica B (Amsterdam) **121**, 202 (1983).
15. N. Bloembergen, E. M. Purcell, and R. V. Pound, Phys. Rev. **73**, 679 (1948).
16. A. F. Andreev, Pis'ma Zh. Éksp. Teor. Fiz. **24**, 608 (1976) [JETP Lett. **24**, 564 (1976)].
17. M. E. R. Bernier, J. Low Temp. Phys. **56**, 205 (1984).
18. M. Devoret, A. S. Greenberg, D. Esteve, *et al.*, J. Low Temp. Phys. **48**, 495 (1982).
19. B. Cowan and M. Fardis, Phys. Rev. B **44**, 4304 (1991).
20. O. V. Lounasmaa, *Experimental Principles and Methods Below One Degree Kelvin* (Academic, New York, 1974; Mir, Moscow, 1977), Chap. 7.
21. I. S. Solodovnikov and N. V. Zavaritskiĭ, Zh. Éksp. Teor. Fiz. **110**, 2047 (1996) [JETP **83**, 1127 (1996)].
22. A. Abragam, *The Principles of Nuclear Magnetism* (Clarendon, Oxford, 1961; Inostrannaya Literatura, Moscow, 1963).
23. B. T. Beal and J. Hatton, Phys. Rev. A **139**, 1751 (1965).
24. B. Cowan and M. Fardis, Fiz. Nizk. Temp. **23**, 598 (1997) [Low Temp. Phys. **23**, 448 (1997)].

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 779–784.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 908–914.
Original Russian Text Copyright © 2002 by Rylkov, Aronzon, Davydov, Kovalev, Meilikhov.

SOLIDS
Electronic Properties

Long-Term Relaxation of Magnetoresistance in a Granular Ferromagnet

V. V. Rylkov[a, b, *], B. A. Aronzon[a, b], A. B. Davydov[b], D. Yu. Kovalev[a, b], and E. Z. Meilikhov[b]

[a]*Institute of Theoretical and Applied Electrodynamics, IVTAN (Institute of High Temperatures) Scientific Association, Russian Academy of Sciences, Moscow, 127412 Russia*

[b]*Russian Research Centre Kurchatov Institute, pr. Kurchatova 1, Moscow, 123182 Russia*

**e-mail: rylkov@imp.kiae.ru*

Received September 7, 2001

Abstract—Long-term relaxations (of the logarithmic type) are revealed in the tunnel magnetoresistance of Fe/SiO_2 nanocomposites, which are due to variation of the magnetization of the nanocomposites. Good qualitative agreement between experimental results and the recently developed concepts of the behavior of magnetization of granular ferromagnets [7] proves that the revealed relaxations are associated with the spin-glass nature of the magnetic state of such systems. It is further demonstrated that it is, in principle, impossible to observe such relaxations using the anomalous Hall effect (proportional to magnetization) because of physical reasons, i.e., mesoscopic fluctuations of the Hall voltage as a result of the magnetic field effect and variation of magnetization. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The fundamental nature of physical phenomena and the possibility of important practical applications promoted significant progress in the investigation and understanding of the physical nature of the effect of giant magnetoresistance in most diverse systems, among which one of the first places is taken by disordered magnetic nanocomposites (granular ferromagnets). They consist of small (1 to 10 nm) ferromagnetic particles located in a dielectric matrix. With the metal content x below some critical value $x_c = 0.5$–0.6, such systems develop a percolation transition from metallic conductivity to conductivity of the tunneling type, under conditions of which the maximal ("giant") magnetoresistance defined by the magnetization of the system is attained [1]. Also unusual in such systems proved to be the behavior of the nondiagonal component of magnetoresistance (under conditions of the Hall effect). As was found by Pakhomov and Yan [2], of basic importance in magnetic nanocomposites (including the mode of tunneling conduction [3]) is, as in the case of homogeneous ferromagnets, the so-called anomalous Hall effect under conditions of which the Hall resistance R_H is proportional to the system magnetization M rather than to the magnetic induction. Note that, in the vicinity of the threshold ($x \approx x_c$), the value of R_H may exceed that in the homogeneous case ($x = 1$) by four orders of magnitude. For this reason, the Hall effect was also referred to as giant [2].

On the other hand, even before the giant magnetoresistance and Hall effect were revealed, systems of small ferromagnetic particles in a dielectric matrix were known as objects with unusual kinetics of relaxation of magnetization (see [4] and the references cited therein). In particular, repeated reports were made of observations of the long-term relaxation of the magnetization of such systems, described by a logarithmic law ($M \propto \text{const} - \ln t$, where t is the time) which is usually attributed to the spin-glass nature of these objects [4, 5].

Therefore, the systems being treated are "bearers" of two out-of-the-ordinary physical phenomena, each of which is associated with the magnetic properties of these systems. It is therefore of interest to investigate the simultaneous manifestation and interference of the respective processes. Such processes include the manifestation of magnetic relaxations in the electrical properties of systems with giant magnetic reluctance. In addition, this approach is important because, in a number of existing models of electrical conductivity of nanocomposites [6], their characteristic magnetic features are, as a rule, of no significance. Moreover, some of those models are based on the concepts of the significant part played by a strong fluctuation potential [6], when the relaxation of resistance is largely associated with Coulomb, rather than magnetic, effects.

In this paper, experimental proof is given of the existence of long-term (logarithmic-type) relaxations of the longitudinal resistance of Fe/SiO_2 nanocomposites, which are due to changes in the magnetization of the nanocomposites. In addition, it was found that the potential possibility of observation of such relaxations using the Hall effect could not be realized in practice for a number of fundamental physical reasons. As was revealed, an important part in this case is played by the

fluctuations of the longitudinal resistance between Hall probes, which arise due to the asymmetry of their effective position; these fluctuations are the stronger, the greater the correlation radius L_c of the percolation cluster. The part played by the magnetic field consists in that it brings about a perturbation of the percolation net. Note that the possibility of "rearrangement" of the percolation net under the effect of a magnetic field is usually ignored, because the magnetoresistance of a nanocomposite amounts usually to just a small fraction of its total resistance.

To the best of our knowledge, no one has previously analyzed the manifestation of magnetic relaxations in the conductivity of magnetic nanocomposites, as was done by us in this study. This may be due to several reasons. First, the variations of magnetoresistance observed in the time range of interest from the experimental standpoint (usually, from several seconds to several tens of minutes) are much less than the respective variations of magnetization. In this case, it is hard to reveal the law of relaxation and, therefore, to judge the nature of the state of a magnet [4]. Second, no simple and adequate models of relaxation of magnetization of the systems being treated were available that would be suitable for processing the results of concrete experiments [4]. Such a model describing the spin-glass behavior of magnetic nanocomposites was developed only recently [7]. This model predicts the range of experimental conditions in which it is possible to observe the logarithmic relaxation in the given materials and enables one to determine a number of characteristic parameters of such systems. It was the notion of nanocomposite as glass, on which this model was based, that defined our approach to formulating and performing relaxation experiments.

2. NANOCOMPOSITE AS SPIN GLASS

Different glasses are often described within a model of two-level systems, i.e., a set of microscopic subsystems with two energy states each, the transitions between which (activated and/or tunneling) are controlled by the energy barrier. Usually, the times of transitions between these states are distributed randomly in a fairly wide range, which is the main reason for the long-term (not exponential) relaxation of the respective physical parameter [8].

A one-domain granule of nanocomposite is also a two-level system, because, by virtue of magnetic anisotropy, it may have two stable states of magnetic moment relative to its easy magnetization axis. The magnetic anisotropy may be due both to crystalline anisotropy and to geometric (associated with the asymmetric shape of granules) anisotropy; in the case of iron, preference must be given to the latter [7]. Indeed, the energy W required for the reorientation of the magnetic moment of an iron granule in the presence of crystalline or geometric anisotropy is

$$W_a^{(c)} \approx \frac{1}{3}K_1 V \quad \text{or} \quad W_a^{(g)} = \frac{1}{2}I_s^2 V\nu,$$

respectively, where $K_1 = 5 \times 10^5$ erg/cm^3 is the crystalline anisotropy constant, V is the granule volume, $I_s \approx 1700$ G is the saturation magnetization, and ν is the form factor dependent on the nonsphericity of the granules. Note that, even with a small nonsphericity of a granule, when the ratio of its axes is just 1.25, the parameter ν is equal to approximately unity (the maximal value of ν for highly extended ellipsoidal granules is about six). Then, for granules 5 nm in size, we find $W_a^{(c)}/k \approx 80$ K and $W_a^{(g)}/k \approx 700$ K at $\nu = 1$ (k is the Boltzmann constant). In other words, the magnetic anisotropy of iron granules is fully defined by the nonsphericity of their shape. Note further that the times of transfer of magnetic moment that are convenient for measurements, $\tau > 1$ s, are usually attained with the measurement temperature $T < W/20k$. The foregoing estimate indicates that, for nonspherical granules, this condition may be realized at nitrogen temperatures.

When the crystalline anisotropy is ignored, the problem concerning magnetization of granular ferromagnets is simplified [7]. We assume that the granules have the shape of ellipsoids of revolution with semiaxes $a > b = c$ (Fig. 1a) and derive, for the magnetic energy W of the granule in the external magnetic field H,

$$\frac{W}{V} = \frac{1}{2}I_s^2 \nu \sin^2(\beta - \gamma) - HI_s \cos\gamma,$$

where γ is the angle between the magnetic field and the magnetic moment of the granule, β is the angle between the magnetic field and the major axis of the ellipsoid, and ν in this case is the difference between the ellipsoid demagnetization coefficients along the b and a axes. The dependence of the energy W on the angle γ of orientation of the magnetic moment relative to the magnetic field for the granules whose long axes form the angles $\beta = \pi/4$ and $\beta = 3\pi/4$ with the magnetic field is given in Fig. 1b for different values of the reduced field $h_\nu = 2H/I_s\nu$. One can see that, in the region of weak fields ($h_\nu \ll 1$), the $W(\gamma)$ dependence has two energy minima; in strong fields ($h_\nu > 1$), only one minimum remains. If the initial (at $H = 0$) magnetic moment is directed at an acute angle to the magnetic field (in this case, $\pi/4$), then (at $T = 0$ K) it will always remain at the right-hand energy minimum (Fig. 1b). The equilibrium angle γ of the magnetic moment orientation tends monotonically to zero with increasing H. If the initial angle is obtuse ($3\pi/4$), the angle γ varies in jumps with the sign variation, which corresponds to the jump of magnetic moment from the left-hand energy minimum to the right-hand minimum. Hence follows a simple method of observation of the maximal variation of

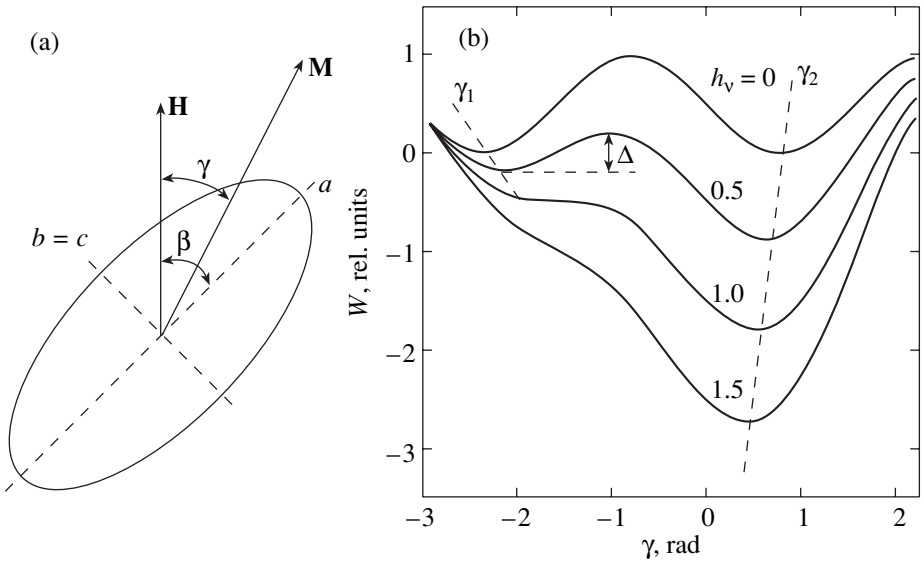

Fig. 1. (a) An ellipsoidal granule and (b) the dependence of its magnetic energy W on the angle γ of orientation of the magnetic moment relative to the magnetic field at $\beta = 3\pi/4$.

magnetization in a granular ferromagnet in relaxation experiments.

We will place a sample at a finite temperature into a fairly high field $H = H_0$, which corresponds to $h_\nu > 1$ (for iron granules, this condition is valid at $H_0 > I_s\nu/2$, i.e., in fields of 5 to 10 kOe). The magnetic moments of the majority of granules will find themselves at one and the same energy minimum (for definiteness, at the right-hand minimum) and will be oriented mainly along the field. We will then switch over the field (with a change of its direction) to a small value of $H = -H_f$ at which the angular dependence of energy $W(\gamma)$ has two minima. It is apparent that, at the initial moment of time, the magnetic moments of granules will be located at the right-hand minimum, which is positioned energetically above the left-hand minimum and separated from the latter by the energy barrier Δ. Such a state is nonequilibrium, and the respective transition to equilibrium will be accompanied by the relaxation of the magnetic moment with the characteristic time

$$\tau = \tau_0 \exp\left[\frac{\Delta(H, \beta)}{kT}\right],$$

where τ_0 is the period of magnetic moment precession (for iron granules, $\tau_0 \sim 10^{-10}$ s [9]). Note that the height of the barrier Δ depends on the angle β (initial orientation of the magnetic moment of granules relative to the field); in particular, the maximal value of Δ is attained at $\beta = \pi/2$ and π [7]. In other words, even in a system of randomly oriented but identical granules, the parameter τ may exhibit a fairly strong spread, and the relaxation of magnetization may be a strong nonexponential function of time. However, detailed analysis reveals that the spread of the shape of granules must be additionally taken into account in order to obtain the relaxation of the logarithmic type in the actual range of time (10 to 10^3 s) [7]. Figure 2 gives the results of calculation of the dependence $M(t)$ for granules 5 nm in size, the values of whose form factor ν are uniformly distributed in the range from 0.5 to 3 (which corresponds to ellipsoidal granules with the axis ratio a/b ranging from 1.1 to 2). It follows from Fig. 2 that $M(t)$ may vary logarithmically in a wide range of time; with the observation times of up to 10^2–10^3 s ($t/\tau_0 = 10^{12}$–10^{13}), the relaxation terminates under conditions when the parameter $T_1 = kT/0.5I_s^2 V \approx 0.1$ and $h_1 = 0.25$, which corresponds to the temperature $T \approx 70$ K and magnetic field $H = 210$ Oe.

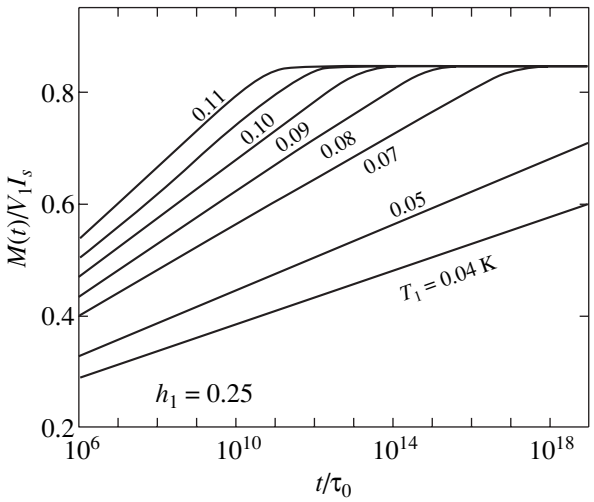

Fig. 2. Calculated curves of magnetization relaxation of a system of randomly oriented ellipsoidal granules with a random form factor for different temperatures.

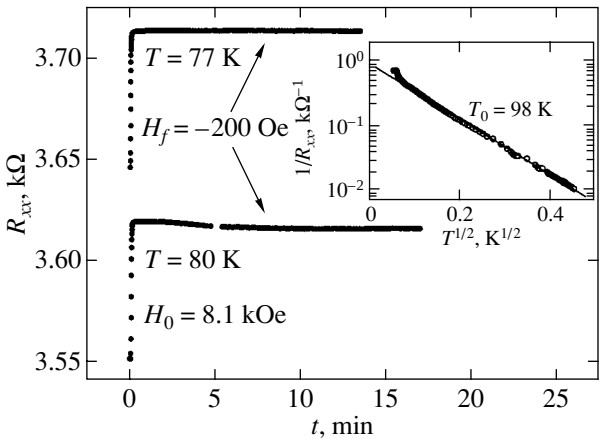

Fig. 3. The relaxation of magnetoresistance at T = 77 and 80 K. The inset gives the temperature dependence of the sample conductance.

3. MAGNETORESISTANCE RELAXATION

We investigated Fe/SiO_2 samples prepared by combined ion-beam sputtering of Fe and SiO_2 from a composite target, which made it possible to vary their volume ratio. The film thickness was varied in the range from 0.2 to 0.6 μm. The samples had the shape of a double Hall "cross" with the conducting channel width w = 2 mm and length L = 7 mm. The precision of alignment of Hall probes was approximately 10 μm. The electron-microscopic investigations revealed that the characteristic size of granules (with a relatively low iron content $x \approx 0.1$) was 3–5 nm [10]. The percolation transition in these objects was observed at $x_c \approx 0.6$, where the maximal value of the effect of giant magnetoresistance was attained (3–4% at T = 77 K) [11]. Note that it is difficult to analyze the shape of granules under these conditions using an electron microscope, because the granules "shield" one another in the case of small distance between them. It is clear, however, that in this case extended metallic formations of two or more granules are very likely to emerge, which are often used to explain the high coercive force in systems of small one-domain nanoparticles [12]. In addition, the results of our investigations of "well"-conducting samples ($x \approx 0.8$) using scanning tunneling microscopy also confirm the presence of extended (ellipsoidal) formations of granules.

The initial magnetic field H_0 was developed by an electromagnet which made possible the switching off of the field of 10^4 Oe during a period of time of the order of 1 s; this was accomplished by deenergizing the electromagnet, after which it was "discharged" via a back-biased diode (relative to the polarity of the supply) connected to the electromagnet winding. The oppositely directed residual field H_f was preinduced by a bias winding and varied within 100–400 Oe. For the field within the film to coincide in magnitude with the external field, a Hall probe was used to orient the sample plane normally to the direction of the electromagnet field (precision of orientation of 1° or better).

Preliminary experiments revealed that the results of measurements of long-term relaxation of magnetoresistance could be significantly distorted by the temperature drift. This is due to a fairly strong temperature dependence of the longitudinal resistance R_{xx} of the objects being investigated (in the region of nitrogen temperatures, the temperature sensitivity of nanocomposites is usually higher than, for example, in the case of frequently used carbon temperature sensors). Figure 3 illustrates the effect of temperature during investigations of the magnetoresistance relaxation.[1] One can see that, as the temperature varies by 4%, the variation of the resistance R_{xx} is approximately two orders of magnitude higher than the amplitude of its relaxation $\Delta R_{xx} = R_{xx}(t = 0) - R_{xx}(t = \infty)$ after the magnetic field is switched over. Therefore, the basic measurements were performed at liquid nitrogen temperature. In order to preclude temperature drifts, the chamber in which the sample was placed was filled with helium gas.

The data on the resistance relaxation after the switching over of the magnetic field are given in Fig. 4a. The value of the final field H_f was selected such that the relaxations would decay with the observation times ranging from several minutes to approximately 20 min. One can see in Fig. 4a that (in accordance with [7]) this occurs in fields of about 200 Oe. The results of [7] relate to the relaxation of magnetization $M(t)$; therefore, the experimentally measured variation of resistance $\Delta R_{xx}(t)$ must be recalculated to the variation of magnetization with due regard for the fact that $\Delta R_{xx}(t) \propto M^2(t)$. The recalculation results are given in Fig. 4b. Note good qualitative agreement between the curves in Fig. 4b and the calculated curves in Fig. 2 (see also the calculated curve in Fig. 4b shown by the dashed curve). Some difference between the temperatures of observation of logarithmic relaxation (up to the moment of decay) may be attributed to a number of factors which were not taken into account. On the one hand, it is the model character of the calculations in [7], which, in particular, fails to account for the possible correlation of the magnetic moments of granules; on the other hand, it is the inaccuracy of orientation of the sample plane relative to the magnetic field and the possible effect of its longitudinal component on the relaxation of magnetoresistance (according to the estimation made in [7], the magnetic field must be perpendicular to the sample plane with an accuracy of much better than 1°).

Nevertheless, adequate qualitative agreement between experiment and the calculation results clearly indicates that the revealed relaxations of resistance are

[1] The inset to Fig. 3 demonstrates the temperature dependence $G(T)$ of the conductance of the given sample in semilog coordinates. It is described well by the known 1/2 law [13]: $G(T) \propto \exp[-(T_0/T)^{1/2}]$ (in this case, T_0 = 98 K).

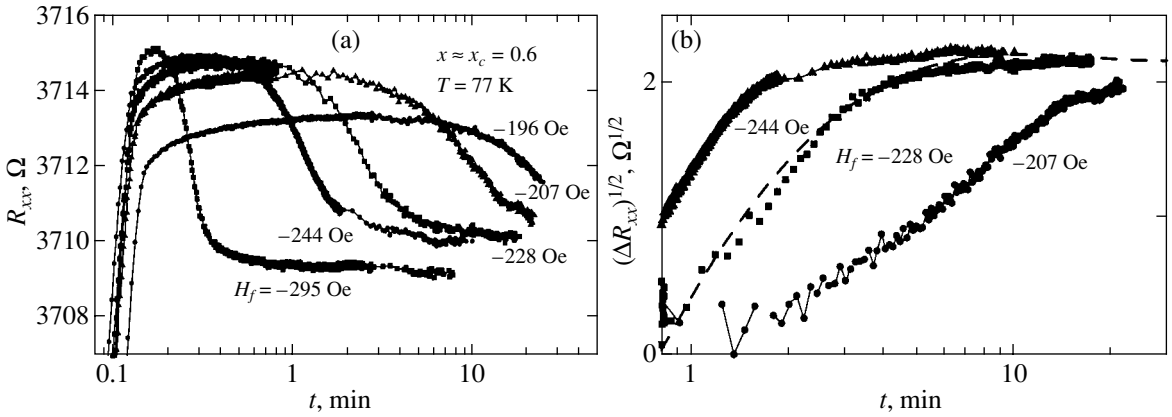

Fig. 4. The relaxation of magnetoresistance (a) after the switching over of the field with a variation of its direction from the initial value $H_0 = 8.1$ kOe to the final value H_f, and (b) recalculated to magnetization. The dashed curve is calculated at $T_1 = 0.1$ ($T \approx 70$ K) and $h_1 = 0.25$ ($H_f = -210$ Oe) with due regard for the scatter of the shape of granules (ν = 0.5 to 3).

associated with the spin-glass nature of the magnetic state of nanocomposite.

4. FLUCTUATIONS DURING MEASUREMENTS OF THE HALL EFFECT

Regular variations or fluctuations of the magnetization δM show up during measurements of the Hall effect as well. First of all, the Hall effect in a magnetic nanocomposite is anomalous; therefore, the potential difference V_y between the Hall probes contains a component proportional to δM. Secondly, in an inhomogeneous system such as the percolation medium being treated, an appreciable voltage is always present between the Hall probes even in the absence of the field. It is proportional to the asymmetry resistance R_a (the resistance arising due to asymmetry of the percolation net) and to the current I_x through the sample, whose fluctuations are proportional to δM^2. In order to determine the contribution made by the Hall resistance R_H, the measurements are performed for two opposite directions of the magnetic field. In this case, $R_H = (R_{xy}^+ - R_{xy}^-)/2$, and $R_a = (R_{xy}^+ + R_{xy}^-)/2$, where R_{xy}^+ and R_{xy}^- are the transverse resistances $R_{xy} = V_y/I_x$ corresponding to the positive and negative directions of the magnetic field, respectively.

Here, it is implicitly assumed that the behavior of the asymmetry resistance in the magnetic field is the same as that of the total sample resistance R_{xx}. However, the results of our experiments demonstrate that this is not the case: the quantity R_a experiences unusual fluctuations.

Figure 5 gives the dependences $R_a(H)$ for a dielectric sample (see Fig. 3) at temperatures $T = 77$ and 300 K. Given in the same figure for comparison are the dependences $R_{xx}(H)$ of magnetoresistance between potential probes located at a distance $l_p = 2.5$ mm from one another. This sample was characterized by a fairly low asymmetry resistance: $R_a \approx 16\ \Omega$ with the resistance $R_{xx} \approx 3700\ \Omega$ (for $T = 77$ K). This corresponds to the effective distance between the Hall probes $l_a \approx l_p R_a/R_{xx} \approx 11$ μm. Note the essentially nonmonotonic pattern of the dependences $R_a(H)$ and their difference from the field dependences of magnetoresistance $R_{xx}(H)$. It is important that the pattern of fluctuations recurs; at $T = 300$ K, the fluctuations are much lower than at $T = 77$ K (cf. curves *3* and *4*). The observed deviations of the dependences $R_a(H)$ from $R_{xx}(H)$ could be interpreted as the effective shift of the Hall probes through the distance

$$\Delta l_a \sim \frac{l_p R_a(0)}{R_{xx}(0)}\left[\frac{R_a(H)R_{xx}(0)}{R_a(0)R_{xx}(H)} - 1\right].$$

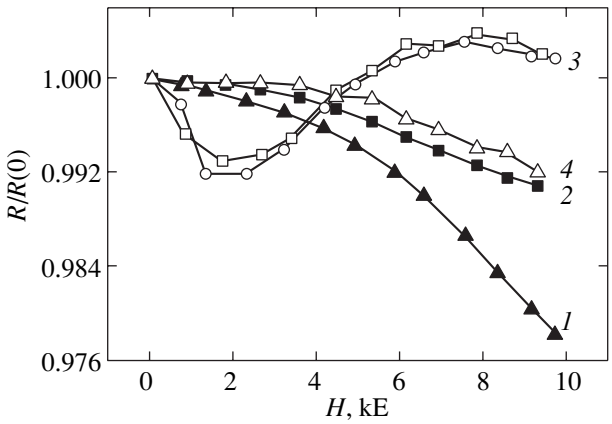

Fig. 5. The magnetic-field dependences of the longitudinal resistance R_{xx} (curves *1* and *2*) and of the asymmetry resistance R_a (curves *3* and *4*) at $T =$ (*1*, *3*) 77 and (*2*, *4*) 300 K. The symbols □ and ○ in the case of curve *3* relate to different series of measurements and demonstrate the reproducibility of this curve.

We recalculate the fluctuations $R_a(H)$ to fluctuations Δl_a in accordance with the latter relation to find that Δl_a amounts to 200 and 30 nm at $T = 77$ and 300 K, respectively. It must be emphasized that we refer to the variation of the effective, rather than real, distance between the Hall probes.

We observed such fluctuations previously during investigation of quasi-two-dimensional semiconductor objects with hopping conductivity, in which the percolation cluster structure varied under conditions of the field effect. It was demonstrated that the characteristic scale Δl_a of the effective "shift" of the Hall probes was defined by the correlation radius (cell size) L_c of the percolation cluster [14].

We believe that in the case being treated, the differences in the behavior of $R_a(H)$ and magnetoresistance $R_{xx}(H)$ are likewise associated with the perturbations of the percolation cluster net. However, the mechanism of this strong effect of the magnetic field on the current paths is not fully clear, because the magnetic field causes the sample resistance to vary by only several percent. However, one can assume that, on mesoscopically small scales, this effect may be much stronger because of the presence of "weak" (sensitive to the magnetic field) spots in the percolation cluster, such as, in particular, micronarrowings with ballistic transport, where the magnetoresistance may reach hundreds of percent [15] (note that the sample is close to percolation transition).

It is natural to attribute the decrease in fluctuations of $R_a(H)$ with increasing temperature to the fact that, as the temperature rises, the cell size of the percolation cluster in a nanocomposite decreases. If we proceed from the 1/2 law for the temperature dependence of conductivity [13], the quantity L_c must vary approximately as $1/T$; i.e., it must decrease by a factor of four as the temperature increases from nitrogen to room temperature.

Therefore, the magnetic disorder of nanocomposites and the spin-glass pattern of their behavior may manifest themselves in two ways, namely, in terms of the variation of the conductivity of the percolation cluster (longitudinal magnetoresistance) and in terms of the variation of the topology of the current paths in this cluster (fluctuations of Hall resistance).

ACKNOWLEDGMENTS

This study received support from the Russian Foundation for Basic Research—RFBR (projects nos. 02-02-16974, 00-02-17191, and 01-02-16420) and from the PICS-RFBR (project no. 01-02-22004).

REFERENCES

1. A. Milner, A. Gerber, B. Groisman, *et al.*, Phys. Rev. Lett. **76**, 475 (1996).
2. A. B. Pakhomov and X. Yan, Solid State Commun. **99**, 139 (1996).
3. B. A. Aronzon, D. Yu. Kovalev, A. N. Lagar'kov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 87 (1999) [JETP Lett. **70**, 90 (1999)].
4. A. Aharoni, in *Magnetic Properties of Fine Particles*, Ed. by J. L. Dorman and D. Fiorani (North–Holland, Amsterdam, 1992), p. 3.
5. Sh. Kogan, *Electronic Noise and Fluctuation in Solids* (Cambridge Univ. Press, Cambridge, 1996), p. 255.
6. C. J. Adkins, in *Metal–Insulator Transitions Revisited*, Ed. by P. P. Edwards and C. N. R. Rao (Taylor & Francis, London, 1995).
7. E. Z. Meilikhov, Zh. Éksp. Teor. Fiz. **116**, 2182 (1999) [JETP **89**, 1184 (1999)]; Zh. Éksp. Teor. Fiz. **117**, 1136 (2000) [JETP **90**, 987 (2000)].
8. W. A. Phillips, Rep. Prog. Phys. **50**, 1657 (1987).
9. F. Bodker, S. Morup, M. S. Pedersen, *et al.*, J. Magn. Magn. Mater. **177–181**, 925 (1998).
10. B. Raquet, M. Goiran, N. Negre, *et al.*, Phys. Rev. B **62**, 17144 (2000).
11. B. A. Aronzon, A. E. Varfolomeev, D. Yu. Kovalev, *et al.*, Fiz. Tverd. Tela (St. Petersburg) **41**, 944 (1999) [Phys. Solid State **41**, 857 (1999)].
12. Yu. I. Petrov, *Physics of Small Particles* (Nauka, Moscow, 1982).
13. E. Z. Meilikhov, Zh. Éksp. Teor. Fiz. **115**, 1484 (1999) [JETP **88**, 819 (1999)].
14. B. A. Aronzon, V. V. Rylkov, A. S. Vedeneev, and J. Leotin, Physica A (Amsterdam) **241**, 259 (1997).
15. L. R. Tagirov, B. P. Vodopyanov, and K. B. Efetov, Phys. Rev. B **63**, 104428 (2001).

Translated by H. Bronstein

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 785–793.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 915–924.
Original Russian Text Copyright © 2002 by Kulakovskiĭ, Gubarev, Lozovik.

SOLIDS
Electronic Properties

Properties of Exciton States in GaAs/AlGaAs Quantum Wells in the Presence of a Quasi-Two-Dimensional Electron Gas

D. V. Kulakovskiĭ[a, *], S. I. Gubarev[a], and Yu. E. Lozovik[b, **]

[a]*Institute of Solid State Physics, Russian Academy of Sciences, Chernogolovka, Moscow oblast, 142432 Russia*
[b]*Institute of Spectroscopy, Russian Academy of Sciences, Troitsk, Moscow Oblast, 142092 Russia*
** e-mail: kulakovd@issp.ac.ru*
*** e-mail: lozovik@isan.troitsk.ru*
Received October 30, 2001

Abstract—The changes in binding energy and oscillator strength of the exciton state due to the screening by a quasi-two-dimensional electron gas are calculated self-consistently in the approximation of noninteracting electrons and in the local field approximation. It is shown that the collapse of the bound state occurs at very low concentrations, $N_s \approx 5 \times 10^9$ cm^{-2}, which is a consequence of the inclusion of the nonlinearity of the response of the system to a Coulomb perturbation. The temperature dependence of the exciton collapse is investigated. The phase diagram of the dissociation of the given bound state is constructed, and the region in which it is possible to observe experimentally the temperature dependence of the exciton collapse is indicated. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Mobile electrons in selectively doped GaAs/AlGaAs quantum wells directly participate in the screening of the Coulomb interaction in 2D structures. Thus, such electrons considerably affect both the stability of the Coulomb centers and the nature of large-scale fluctuations in quantum wells. In spite of the fact that the spectrum of electrons in quantum wells, for which the experiments were made, is indeed size-quantized and may be treated as a purely 2D spectrum, these structures are quasi-two-dimensional as applied to the problem of the screening of Coulomb interaction since the width of typical quantum wells, which ranges from 200 to 300 Å, exceeds the exciton Bohr radius. For this reason, the screening of the Coulomb interaction in real quantum wells is of a mixed type, varying from purely 2D screening at large distances (much longer than the quantum well width) to virtually 3D screening at small distances. The screening effects can be observed in the luminescence and reflection spectra [1–4].

It was shown experimentally in [3, 4] that the threshold concentration for which the rearrangement of exciton states takes place strongly depends on the quality of the structure and is observed for most perfect structures at extremely low concentrations $N_s = 5 \times 10^9$ cm^{-2}. This corresponds to the dimensionless parameter r_s describing the mean distance between electrons in the gas in the units of Bohr radius a_B, $r_s = 1/(a_B\sqrt{2\pi N_s}) \approx 8$. This value is several times higher than that observed earlier in experiments on structures having a worse quality (see, for example, [2]), in which the rearrangement of exciton states was observed for order-of-magnitude higher electron gas concentrations. A theoretical analysis of the Coulomb interaction screening by a 2D electron gas was carried out by Bauer [5] in the framework of the theory of dielectric screening and by Kleinman [6] in the approximation of a linear dielectric response for a purely 2D electron gas. However, both these approaches lead to considerably higher values of the threshold concentration as compared to those observed in perfect GaAs/AlGaAs structures.

In the present work, a method of self-consistent calculation of the screening of the Coulomb interaction by a quasi-two-dimensional electron gas is developed both in the approximation of noninteracting electrons and in the local field approximation. This method allowed us to take into account, to a certain extent, the nonlinearity of screening; as a result, the threshold values of concentration shifted from $r_s \sim 3$ to the range of values of $r_s \sim 8$, which is in qualitative agreement with the results of recent experiments. We also analyzed the screening at a nonzero temperature. In this case, the "blurring" of the Fermi step lowers the efficiency of screening of a quasi-two-dimensional electron gas, leading to an increase in the threshold value of concentration. We also calculated the thermal dissociation of the exciton state. The corresponding phase diagram is presented in the last section of this research.

Henceforth, we will be interested in the binding energy of the exciton state in the presence of a quasi-two-dimensional electron gas with concentration N_s. It is well known [7, 8] that the problem of determining the binding energy of a large-radius exciton can be reduced to the problem of a Coulomb center with a particle mass equal to the reduced mass of the exciton, μ =

$m_e m_h/(m_e + m_h)$, where m_e and m_h are the planar masses of the electron and hole in the quantum well, respectively.

2. FORMULATION OF THE PROBLEM

Let us consider a quantum well of width l_0 with infinitely high walls, so that the electron wave function in the z direction is strictly limited by its size. It was stated above that, as a model of an exciton, we can consider a positively charged Coulomb center located at the middle of a quantum well for $z = 0$ and an electron having the mass equal to the reduced mass μ of an exciton and bound to this center. The Hamiltonian of such a center in the cylindrical system of coordinates has the form

$$\hat{H} = -\frac{\hbar^2}{2\mu}\Delta + U(\rho, z), \tag{1}$$

where

$$U(\rho, z) = -\frac{e^2}{\epsilon\rho} + F(z). \tag{2}$$

Here, $\rho = \sqrt{r^2 + z^2}$, $F(z) = 0$, $z \le |l_0/2|$, and $F(z) = \infty$, $z > |l_0/2|$.

We will seek the ground-state energy by using the Ritz variational method with a test wave function of the bound state in the form

$$\Psi(r, z) = N\cos\left(\frac{\pi z}{l_0}\right)\exp\left(-\frac{\sqrt{r^2 + \gamma^2 z^2}}{r_0}\right). \tag{3}$$

In this function, there are two variable parameters: the effective radius r_0 of the Coulomb center (exciton) in the (x, y) plane and the parameter γ taking into account the anisotropy associated with the restriction of the motion in the z direction. Such a wave function correctly describes the behavior of the system in narrow quantum wells with $l_0 \ll r_0$ (in this case, $\gamma \longrightarrow 0$ and the function coincides with a purely 2D function) as well as in wide quantum wells with $l_0 \ge r_0$. In the latter case, $\gamma \sim 1$ and the function is spherically symmetric as for 3D systems.

For variational calculations, it is convenient to introduce the effective 2D potential $U_{\text{eff}}(r)$ which can be written in the adiabatic approximation[1] in the form

$$U_{\text{eff}}(r) = \int |\Psi(r, z)|^2 U(r, z)\,dz. \tag{4}$$

The Fourier transform of this potential is

$$\begin{aligned} U_{\text{eff}}(q) &= \iint e^{i\mathbf{qr}} U_{\text{eff}}(r)\,d\mathbf{r} \\ &= 2\pi\int J_0(qr) U_{\text{eff}}(r) r\,dr, \end{aligned} \tag{5}$$

where $J_0(x)$ is a Bessel function of the first kind.

The ground-state energy of a quasi-two-dimensional exciton can be determined from the minimum of the functional F:

$$F = \langle\Psi(r, z)| -\frac{\hbar^2}{2\mu}\Delta|\Psi(r, z)|\rangle + \langle|U_{\text{eff}}(r)|\rangle \tag{6}$$

in the parameters r_0 and γ.

We will take into account the screening of the effective potential $U_{\text{eff}}(r)$ by a quasi-two-dimensional electron gas through the dielectric function $\epsilon(q)$:

$$U_{\text{eff}}^{\text{scr}}(r) = \int J_0(qr)(U_{\text{eff}}(q)/\epsilon(q))q\,dq. \tag{7}$$

3. LINDHARD THEORY OF SCREENING

The response function, or the susceptibility $\chi(q, \omega)$ of the system, of a quasi-two-dimensional electron gas with potential $U_{\text{eff}}(q)$ for the external perturbation $V^{\text{ext}}(q, \omega)$ can written, by definition, as

$$\delta n(q, \omega) = \chi(q, \omega) V^{\text{ext}}(q, \omega), \tag{8}$$

where $\delta n(q, \omega)$ is the change in the density due to the interaction with the external perturbation.

The polarization operator $\Pi(q, \omega)$ is defined as the susceptibility of the system to the already induced potential $V^{\text{ind}}(q, \omega)$ which is the sum of the external potential and the effective interaction potential, $V^{\text{ind}}(q, \omega) = V^{\text{ext}}(q, \omega) + U_{\text{eff}}(q)$:

$$\delta n(q, \omega) = \Pi(q, \omega) V^{\text{ind}}(q, \omega). \tag{9}$$

Formulas (8) and (9) lead to the following dependence between the susceptibility and polarization of the system:

$$\chi(q, \omega) = \frac{\Pi(q, \omega)}{1 - U_{\text{eff}}(q)\Pi(q, \omega)}. \tag{10}$$

By definition, the dielectric function is the ratio of the external perturbation to the induced potential and, hence, can be written in the form

$$\begin{aligned} \frac{1}{\epsilon(q, \omega)} &= 1 + \frac{U_{\text{eff}}(q)\delta n(q, \omega)}{V^{\text{ext}}(q, \omega)} \\ &\equiv 1 + U_{\text{eff}}(q)\chi(q, \omega) \end{aligned} \tag{11}$$

or

$$\epsilon(q, \omega) = 1 - U_{\text{eff}}(q)\Pi(q, \omega). \tag{12}$$

In the Hartree–Fock approximation (HFA), electrons "respond" to the external field as free particles;

[1] The condition that the separation between the exciton levels is much smaller than the characteristic size-quantization energy of e^- (which is equal approximately to $\pi^2\hbar^2/2m_e l_0^2$) in the quantum well is sufficient for the applicability of the adiabatic approximation in the given problem.

consequently, the quantity $\chi(q, \omega)$ for a homogeneous system can be approximated by the polarization of free electrons (Lindhard approximation of noninteracting electrons) $\Pi^0(q, \omega)$; i.e.,

$$\chi^{\mathrm{HFA}}(q, \omega) = \Pi^0(q, \omega),$$
$$\frac{1}{\epsilon^{\mathrm{HFA}}(q, \omega)} = 1 + U_{\mathrm{eff}}(q)\Pi^0(q, \omega). \quad (13)$$

The expression for $\Pi^0(q, \omega)$ was derived using the first order of perturbation theory in the external potential [9]:

$$\Pi^0(q, \omega) = \frac{1}{L^2}\lim_{\alpha \to 0}\sum \frac{f_0(E_k) - f_0(E_{k+q})}{E_{k+q} - E_k - \hbar\omega - i\hbar\alpha}, \quad (14)$$

where f_0 is the Fermi–Dirac distribution function and L^2 is the area of the system.

In the random-phase approximation (RPA), electrons respond to the induced field as free electrons, so that

$$\Pi^{\mathrm{RPA}}(q, \omega) = \Pi^0(q, \omega),$$
$$\epsilon^{\mathrm{RPA}}(q, \omega) = 1 - U_{\mathrm{eff}}(q)\Pi^0(q, \omega). \quad (15)$$

For $T = 0$ and for the Fermi wave vector $k_F = \sqrt{2\pi N_s}$, the static dielectric function in the RPA can be written (see, for example, [10]) as

$$\epsilon^{\mathrm{RPA}}(q) = 1 + U_{\mathrm{eff}}(q)\frac{m_e}{\pi\hbar^2}$$
$$\times [1 - \Theta(q - 2k_F)\sqrt{1 - (2k_F/q)^2}], \quad (16)$$

where $U_{\mathrm{eff}}(q)$ is defined by formula (5).

The ground-state energy of an exciton in the effective screened potential $U_{\mathrm{eff}}^{\mathrm{scr}}(r)$ obtained through the numerical solution of the integral equation (7) was calculated by varying functional (6) in the parameters r_0 and γ. Since potential $U_{\mathrm{eff}}^{\mathrm{scr}}(r)$ is itself a function of the wave function parameters r_0 and γ, successive iterations led to self-consistent values of r_0 and γ and of the potential $U_{\mathrm{eff}}^{\mathrm{scr}}(r)$ depending on them. It should be noted that such a procedure of self-consistent calculations makes it possible to go beyond the linear response of the electron subsystem and to take into account, to a certain extent, the nonlinear nature of screening of the 3D Coulomb potential by a 2D gas.

In all our calculations, we operate with quantum wells based on GaAs/AlGaAs. This means that we will use the values of the planar masses of electron $m_e = 0.067m_0$, hole $m_h = 0.26m_0$, and the static permittivity $\epsilon = 12.8$. Figure 1 shows the results of numerical calculations of the dependence of the exciton binding energy on the dimensionless parameter r_s. It can be seen that,

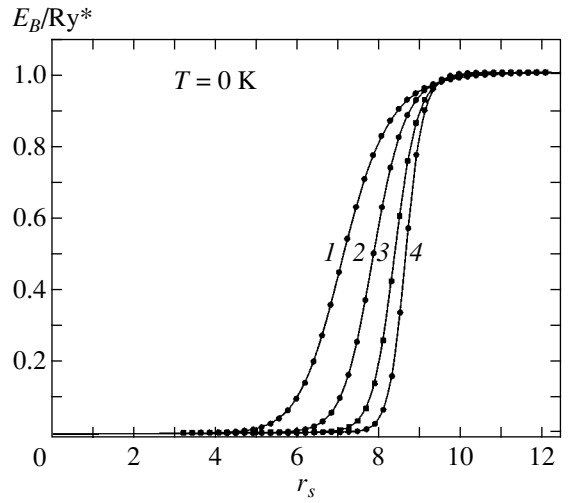

Fig. 1. Binding energy E_B of the exciton state in GaAs/AlGaAs quantum wells of widths 50 (*1*), 100 (*2*), 200 (*3*), and 300 Å (*4*) as a function of the dimensionless parameter r_s. Calculations were made using the values of planar masses of electron ($m_e = 0.067m_0$) and hole ($m_h = 0.26m_0$) and the static permittivity $\epsilon = 12.8$.

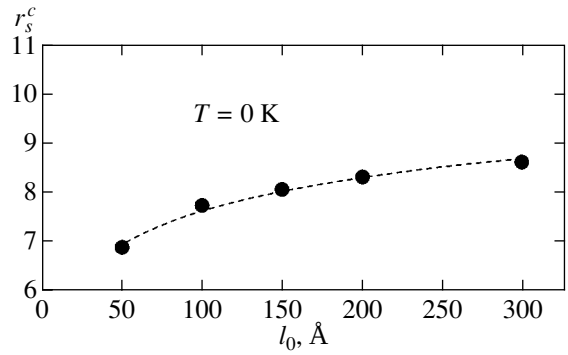

Fig. 2. Dependence of the critical parameter r_s^c at which the exciton states collapse on the width of a GaAs/AlGaAs quantum well.

as the concentration of a quasi-two-dimensional electron gas increases, the exciton binding energy decreases abruptly, in a threshold manner. For a quantum well of width $l_0 = 300$ Å, the sharp decrease (rearrangement) of the binding energy occurs in the region of $r_s \approx 8$. As the well width decreases, the value of the threshold concentration for which the screening of the exciton states increases strongly is shifted towards the region of smaller values of r_s (higher concentrations) (curves *1–4* in Fig. 1). Assuming, for the sake of definiteness, that the threshold concentration is that for which the binding energy decreases by a factor of *e*, we can plot the dependence of the critical parameter r_s^c on the quantum well width (Fig. 2).

The results of our calculations show that the rearrangement of exciton states emerges for much lower

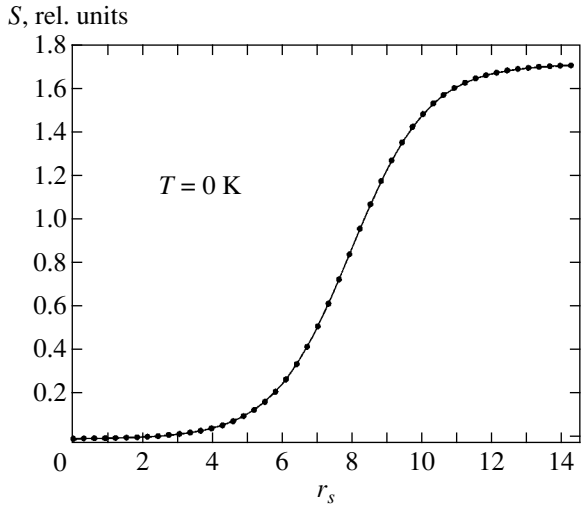

Fig. 3. Oscillator strength S of the exciton transition as a function of the dimensionless parameter r_s calculated for a GaAs/AlGaAs quantum well of width 200 Å.

concentrations of the electron gas ($r_s \approx 8$) than in the previous calculations made by Bauer [5] ($r_s = 1.8$), who considered the dielectric screening of the Coulomb interaction, or by Kleinman [6] ($r_s = 2.8$), who made calculations for a purely 2D case in the linear screening approximation.

Several aspects which are important in this problem are worth noting. First, we must take into account the dependence of the dielectric function on the parameters of effective interaction $U_{\text{eff}}(r)$, i.e., the nonlinear approximation for the system response. If we confine our analysis to the linear response approximation in the procedure of self-consistent calculation, the value of the threshold concentration for which the screening of exciton states is observed is shifted to the region of higher concentrations, corresponding to the parameter $r_s = 3.5$, which is close to the results obtained in [6, 11]. Second, the singularity in the dielectric function (10) for $q = 2k_F$ leads to Friedel oscillations of the quasi-two-dimensional electron gas concentration as in the 3D case. In contrast to 3D systems, the effect of these oscillations is significant since the asymptotic form of the screening potential in the 2D case is of the power type, and the contribution from the oscillatory behavior of local concentration in the vicinity of a Coulomb center is noticeable, leading to a more effective potential screening.

In experiments on the screening of exciton states by a quasi-two-dimensional electron gas, information on the change in the dependence of the binding energy of exciton states on the density of this gas is not available as a rule since the knowledge of the energy of uncorrelated electron and hole, which have no specific features in the luminescence and absorption spectra, is required in this case. At the same time, a threshold variation in the exciton-transition oscillator strength is clearly manifested in optical experiments as the density of the quasi-two-dimensional electron gas attains a certain threshold value. For convenience of comparison with the experiments, we calculated, apart from the binding energy, the behavior of the exciton-transition oscillator strength $S \propto |p_{cv}|^2 a_B^{-2} |\Psi(r = 0, z = 0)|^2$ [6] as a function of the electron gas concentration in quantum wells (Fig. 3). It can be seen that the exciton transition intensity decreases upon an increase in the density of the quasi-two-dimensional electron gas less sharply than the binding energy of the exciton state. As a result, the exciton absorption line can be observed in optical experiments even for relatively high electron concentrations, when the binding energy of excitons has already been decreased significantly as a result of screening by a quasi-two-dimensional electron gas.

It is also interesting to note that the rearrangement of exciton states is accompanied by a manifold increase in the effective Bohr radius of an exciton state. The effective radius of an exciton along the well is many times larger than the quantum well width even for broad quantum wells with $l_0 = 300$ Å for high densities of the electron gas; consequently, the exciton becomes virtually two-dimensional. At the same time, at low densities, the effective radius of an exciton is considerably smaller than the well width and the exciton wave function in the well differs insignificantly from a 3D wave function. Thus, the screening of the exciton state in broad quantum wells is accompanied, in addition to a decrease in the binding energy and the oscillator strength, by the exciton "crossover" 3D $\longrightarrow$ 2D, i.e., a transition of the exciton state from three to two dimensions. This may be an additional reason for the sharpness of the observed rearrangement of the exciton state in broad quantum wells.

It should also be noted that, since the problem has become purely two-dimensional after the introduction of the effective potential $U_{\text{eff}}(r)$ [4], the bound state always exists in this case (see, for example, [8]). However, the binding energy of this state (see Fig. 1) in the case of a high concentration (small r_s) is found to be exponentially small:

$$E_B \sim \frac{\hbar^2}{\mu r_0^2} \exp\left[-\frac{\hbar^2}{\mu P}\right],$$

where

$$P = \left|\int_0^\infty U_{\text{eff}}(r)dr\right|$$

is the power of the potential well.

4. LOCAL FIELD APPROXIMATION

In this procedure, we used the Lindhard susceptibility approximation corresponding to the situation of noninteracting electrons in the gas. To a certain extent,

the Coulomb interaction between electrons at small distances can be taken into account in the local field approximation [12] by replacing the polarization operator (9) with a more complex operator taking into account more adequately the interaction at small distances, which has the following form [13, 14]:

$$\Pi(q, \omega) = \frac{\Pi^0(q, \omega)}{1 - f_q(\omega)\Pi^0(q, \omega)}, \tag{17}$$

where $\Pi^0(q, \omega)$ is the polarization operator in the RPA, which is defined by formula (14), and $f_q(\omega)$ is the local field factor. Then, formulas (10) and (17) give

$$\chi(q, \omega)) = \frac{\Pi^0(q, \omega)}{1 - (U_{\text{eff}}(q) + f_q(\omega))\Pi^0(q, \omega)} = \frac{\Pi^0(q, \omega)}{1 - v_{\text{eff}}(q, \omega)\Pi^0(q, \omega)}, \tag{18}$$

where $v_{\text{eff}}(q, \omega) = U_{\text{eff}}(q) + f_q(\omega)$ is the effective interaction potential in the local field approximation. The dielectric function corresponding to this polarization operator has the form

$$\epsilon^{\text{STLS}}(q, \omega) = 1 - \frac{U_{\text{eff}}(q)\Pi^0(q, \omega)}{1 - f_q(\omega)\Pi^0(q, \omega)}. \tag{19}$$

The formula for the effective static potential $v_{\text{eff}}(r)$ can be presented in the form (see [15] and the Appendix)

$$v_{\text{eff}}(r) = -\int_r^\infty dr g(r) \frac{dU_{\text{eff}}(r)}{dr}, \tag{20}$$

where $g(r) = g_{\uparrow\uparrow}(r) + g_{\uparrow\downarrow}(r)$ is the pair correlation function. Obviously, if $g(r) = 1$ (this corresponds to noninteracting electrons), we have $v_{\text{eff}}(r) \equiv U_{\text{eff}}(r)$. This returns us again to the Lindhard limit. In analogy with formula (7), we can now obtain the following expression for the screened effective static interaction potential in the local field approximation:

$$v_{\text{eff}}^{\text{scr}}(r) = \int J_0(qr)(v_{\text{eff}}(q)/\epsilon^{\text{STLS}}(q))rdr, \tag{21}$$

where $\epsilon^{\text{STLS}}(q)$ is the static dielectric function calculated by formula (19) for $\omega = 0$.

Thus, we calculate the static form factor (A.4) and, hence, the pair correlation function (A.2) in the initial approximation of noninteracting electrons. Further, we assume that the following changes in $U_{\text{eff}}(r)$ do not lead to considerable changes in the static form factor and in the pair correlation function. In this case, the local field factor can be written in the form (see [13]) $f_q = G(q)U_{\text{eff}}(q)$, where $G(q) = 1 - g(0) + o(1/q^2)$. Consequently, in order to determine the effective static potential $v_{\text{eff}}(r)$, we must evaluate integral (20), and then the effective screened potential $v_{\text{eff}}^{scr}(r)$ of the interaction between a positively charged Coulomb center and an electron bound to it will be obtained from the numerical solution of the integral equation (21), the changes in the dielectric function $\epsilon^{\text{STLS}}(q)$ being determined only by the changes in the effective 2D potential $U_{\text{eff}}(q)$.

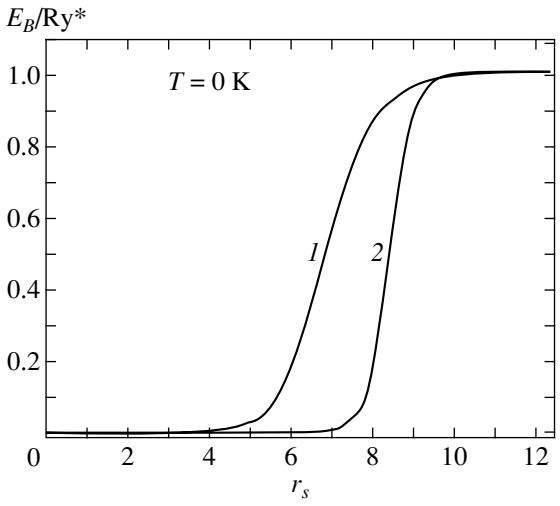

Fig. 4. Binding energy E_B of the exciton state in GaAs/AlGaAs quantum wells of width 200 Å as a function of the dimensionless parameter r_s in the local-field (*1*) and random phase (*2*) approximations.

Let us now determine the binding energy of an exciton directly in the local field approximation. We will use the procedure of variational self-consistent calculations described in Section 3, replacing $U_{\text{eff}}^{\text{scr}}(r)$ by $v_{\text{eff}}^{\text{scr}}(r)$ in it. Thus, we will calculate the ground-state energy of the exciton by varying functional (6) with the effective screened potential $v_{\text{eff}}^{\text{scr}}(r)$ in parameters r_0 and γ. Figure 4 presents the results of such a numerical calculation of the binding energy of an exciton for quantum wells of width 200 Å as a function of the dimensionless parameter r_s. It can be seen that, when the correlation corrections are taken into account, the exciton state collapses in the region of lower values of r_s (as compared to those observed in the random phase approximation), which corresponds to an increase in the threshold concentration of a quasi-two-dimensional electron gas and to a deteriorated screening. In order to explain this effect, we consider the compressibility as one of the parameters of the system.

5. COMPRESSIBILITY OF THE SYSTEM

The compressibility of the system is connected with the response function through the formula [16]

$$K = -\frac{1}{N_s^2} \lim_{q \to 0} \chi(q, 0). \tag{22}$$

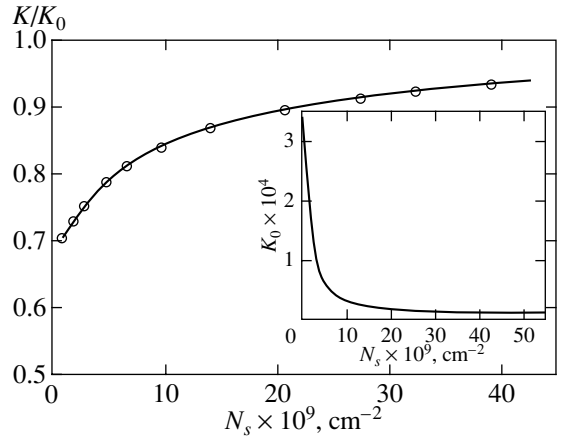

Fig. 5. The ratio of the compressibility K of the electron system in the local field approximation to the compressibility K_0 of the system of noninteracting electrons at $T = 0$ K as a function of the two-dimensional electron gas concentration N_s for a quantum well of width 200 Å. The inset shows the dependence of K_0 on N_s.

We denote by K_0 the compressibility of a system of free electrons. In the Lindhard approximation, we obviously have

$$N_s^2 K_0 = -\lim_{q \to 0} \Pi^0(q, 0).$$

Using the frequency sum rule [13, 16, 17] and formula (19) for the dielectric function, we obtain the compressibility in the local field approximation:

$$\begin{aligned} &-\lim_{q \to 0} \frac{2}{\pi U_{\text{eff}}(q)} \int_0^\infty \frac{d\omega'}{\omega'} \text{Im}[\epsilon^{\text{STLS}}(\omega')] \\ &= \lim_{q \to 0} \text{Re} \frac{\Pi^0(q, 0)}{1 - f_q(0)\Pi^0(q, 0)} \\ &= -\lim_{q \to 0} \frac{N_s^2 K_0}{1 + f_q(0) N_s^2 K_0} \equiv -N_s^2 K. \end{aligned} \tag{23}$$

It follows hence that

$$\frac{K_0}{K} = 1 + \lim_{q \to 0} f_q(0) N_s^2 K_0. \tag{24}$$

Using the above formulas, we can easily prove that (see Fig. 5) the compressibility of an electron system calculated in the local field approximation is smaller than the compressibility of a system in the random phase approximation for the same electron density in the systems. This result was obtained experimentally in [18]. This means that the system becomes more rigid[2]; i.e., the average electron spacing increases due to the short-range interaction. The change in the local concentration in this case turns out to be smaller than in a system of noninteracting electrons. This leads to a worse screening and, hence, to an increase in the threshold concentration for which the rearrangement of the exciton state takes place.

It should be noted that, strictly speaking, the local field approximation is applicable only for $r_s \leq 6$ (see [12]); consequently, the results obtained in the last two sections are rather qualitative by nature and are presented here to explain qualitatively the behavior of the system in the range of values of $r_s \sim 8$ we are interested in.

6. TEMPERATURE DEPENDENCE OF THE THRESHOLD CONCENTRATION OF A QUASI-TWO-DIMENSIONAL ELECTRON GAS

An analysis of the temperature dependence of the threshold concentration of a quasi-two-dimensional electron gas at which exciton rearrangement takes place is an important experimental problem. In this section, we will analyze this problem theoretically. As before, we will use the approximation of noninteracting electrons. In order to calculate the screening effect at a nonzero temperature, we will use a temperature-dependent dielectric function. In the high-temperature limit ($T \sim E_{\text{Fermi}}$), it has the following analytic expression [19]:

$$\epsilon^{\text{RPA}}(q) = 1 + U_{\text{eff}}(q) q_s(q), \tag{25}$$

$$q_s(q) = \frac{N_s}{k_B T} g_1(q\lambda). \tag{26}$$

Here, we have

$$\lambda = \sqrt{\frac{2\pi\hbar^2}{m_e k_B T}}, \quad g_1(x) = \frac{2\sqrt{\pi}}{x} \Phi\left(\frac{x}{4\sqrt{\pi}}\right),$$

$$\Phi(y) = \pi^{-1/2} \int_{-\infty}^{\infty} dz \frac{e^{-z^2}}{y - z}.$$

At low temperatures, the dielectric function can be obtained numerically.

It should be noted that the formula describing the change $\delta n(q)$ in the local electron concentration for a classic gas (see formula (8)) can be written in the form (see, for example, [20])

$$\begin{aligned} \delta n(q) &\approx N_s \exp\left(-\frac{U_{\text{eff}}(q)}{k_B T}\right) - N_s \\ &\approx -\frac{N_s}{k_B T} U_{\text{eff}}(q). \end{aligned} \tag{27}$$

[2] Zero compressibility corresponds to an absolutely incompressible liquid.

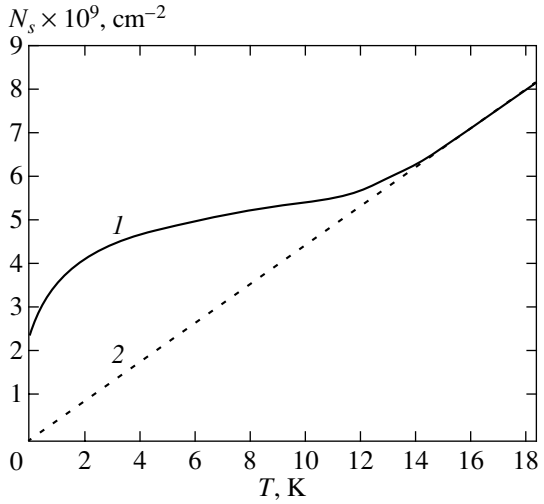

Fig. 6. Temperature dependence of the threshold concentration N_s of a quasi-two-dimensional electron gas for a GaAs/AlGaAs quantum well of width 300 Å in the random phase approximation (*1*) and in the Debye–Hückel approximation (classical limit) (*2*).

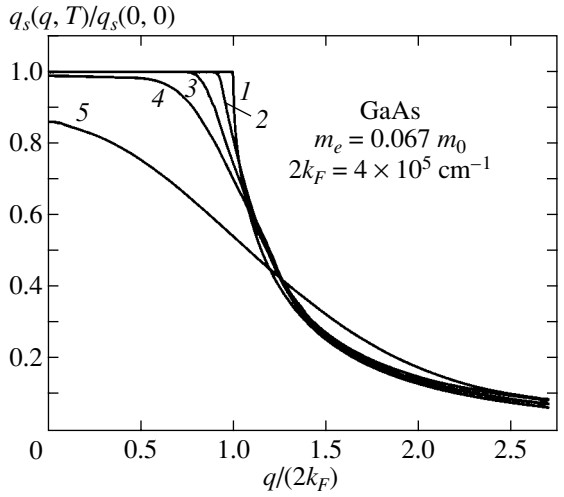

Fig. 7. The ratio of the effective screening parameter $q_s(q, T)$ to its value $q_s(0, 0)$ at $T = 0$ K on the wave vector q (in units of $2k_F$) at different temperatures, K: 0 (*1*), 0.1 (*2*), 0.5 (*3*), 1 (*4*), and 2.5 (*5*).

This leads to a formula of the type (25), but now we have

$$q_s^{\text{classic}}(q) = \frac{N_s}{k_B T}. \tag{28}$$

Thus, if the temperature of an electron system is of the order of the Fermi temperature, the quantum-mechanical formula (26) for the screening parameter $q_s(q)$ is transformed into the Debye–Hückel classical formula (28).

Using now the temperature-dependent dielectric function (25) in self-consistent variational calculations, we obtain the dependence of the critical concentration N_s^c of a quasi-two-dimensional electron gas as a function of temperature for a quantum well of width l_0 = 300 Å. These dependences are presented in Fig. 6. As the temperature increases, the Fermi step is "blurred" and, hence, the concentration of electrons with small values of q becomes lower. This leads to a considerable decrease in the screening parameter $q_s(q)$ for values of momenta $q \le 2k_F$ (see Fig. 7). Since the screening is mainly accomplished by electrons with $q \le 2k_F$, the effect of screening by the quasi-two-dimensional electron gas becomes weaker, and the value of the threshold concentration increases (see Fig. 6). On the other hand, at temperatures above 15 K, when the quasi-two-dimensional electron gas can be regarded as classic, we arrive at complete agreement between the results of calculations based on the quantum-mechanical Lindhard (solid curve in Fig. 6) and the classical Debye–Hückel (dashed line in Fig. 6) dielectric functions.

Another interesting problem emerging at a nonzero temperature is associated with the calculation of the phase diagram of dissociation of the exciton state. By

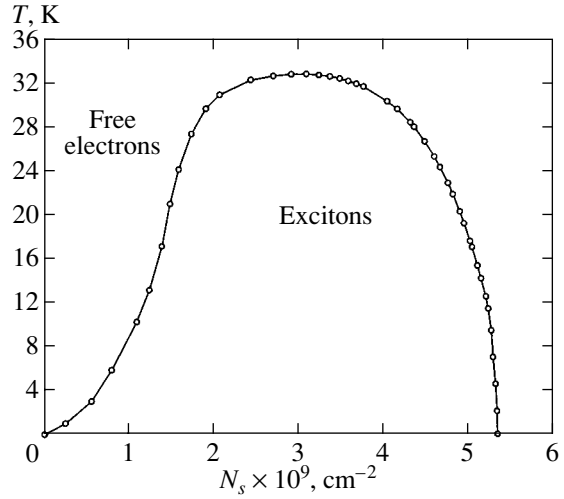

Fig. 8. Phase diagram of thermal dissociation of the exciton state in a GaAs/AlGaAs quantum well of width 300 Å.

definition (see [21]), the dissociation energy is the exciton binding energy. In analogy with [22], we obtain the following condition for the phase transition:

$$T = E_B(T, N_s)/\ln\left[\frac{m_e T}{\pi\hbar^2 N_s}\right], \tag{29}$$

where $\ln[m_e T/\pi\hbar^2 N_s]$ is a quantity describing the effect of ionization of the given exciton.[3] The result of calcu-

[3] For a high electron concentration, the ionization effect can be disregarded; consequently, $\ln[m_e T/\pi\hbar^2 N_s] \approx 1$. Conversely, for $N_s \longrightarrow 0$, we obtain $\ln[m_e T/\pi\hbar^2 N_s] \longrightarrow \infty$, and the ionization effect becomes predominant in the system.

lation of the phase diagram for a quantum well of width 300 Å is shown in Fig. 8. It can be seen that the screening effect reducing the binding energy and, hence, leading to dissociation of the exciton state can be observed for $N_s > 3 \times 10^9$ cm^{-2}. For lower concentrations, the ionization effect becomes predominant.

Thus, the temperature dependence of the exciton collapse determined above (see Fig. 6) is observed in the region of system parameters for which the phase transition is not yet observed. Consequently, it can be investigated experimentally.

7. CONCLUSIONS

In the present work, we have calculated the changes in the binding energy and oscillator strength for an exciton state which emerge as a result of screening by a quasi-two-dimensional electron gas in GaAs/AlGaAs quantum wells of width 50–300 Å. It is shown that the inclusion of the nonlinear response leads to a stronger screening of the Coulomb interaction as compared to that in the linear approximation and, as a result, shifts the threshold concentration for which the exciton states are rearranged towards lower densities of a quasi-two-dimensional electron gas and, accordingly, higher values of r_s (r_s = 8.3 for a well of width 300 Å). This concentration considerably exceeds the values calculated in the framework of dielectric screening (Mott transition) or in the framework of the linear screening by a quasi-two-dimensional electron gas. As the quantum well width decreases, the threshold electron concentration for which the transition takes place is displaced towards lower values of parameter r_s.

It is shown that the inclusion of correlation effects in the framework of the local field approximation makes the system more rigid and less capable of screening the introduced charged perturbation. This leads to an increase in the threshold concentration of the quasi-two-dimensional electron gas.

We have investigated the temperature dependence of the critical parameter r_s^c. As the temperature increases, the effectiveness of screening becomes lower and the critical parameter r_s^c decreases. This is a consequence of the decrease in the density of electrons with small values of q, which make the main contribution to the given effect. We have also demonstrated the transition from a purely quantum to the classical system upon an increase in temperature. We have constructed the phase diagram of the exciton state dissociation and indicated the region in which the temperature dependence of the exciton collapse can be observed experimentally.

ACKNOWLEDGMENTS

The authors are grateful to I.V. Kukushkin for his interest in this research and fruitful discussions of the results.

This work was financed by the Russian Foundation for Basic Research and the INTAS.

APPENDIX

In the local field approximation, the effective potential $v_{\text{eff}}(r)$ can be written in terms of the pair correlation function [15, 23]:

$$v_{\text{eff}}(r) = -\int_r^\infty dr g(r) \frac{dU_{\text{eff}}(r)}{dr}. \tag{A.1}$$

The expression for the pair correlation function $g(r) = g_{\uparrow\uparrow}(r) + g_{\uparrow\downarrow}(r)$ has the form

$$g(r) - 1 = \frac{1}{k_F^2}\int_0^\infty J_0(qr)[S(q) - 1] q dq, \tag{A.2}$$

where $S(q)$ is the static form factor, which can be expressed using the fluctuation-dissipative theorem [17] through the following formula:

$$S(q) = -\frac{2}{k_F^2}\int_0^\infty d\omega \text{Im}\chi(q, \omega). \tag{A.3}$$

In the initial approximation, the static form factor can be calculated proceeding from the Lindhard susceptibility defined by formulas (11) and (16). After evaluating the integrals, we obtain

$$S(q) = \begin{cases} 1 - \dfrac{\arccos(q/2k_F)}{\pi} + \dfrac{q\sqrt{1-(q/2k_F)^2}}{2\pi k_F}, & \text{if } q \le 2k_F, \\ 1, & \text{if } q > 2k_F. \end{cases} \tag{A.4}$$

REFERENCES

1. R. C. Miller, D. A. Kleinman, W. T. Tsang, and A. C. Gossard, Phys. Rev. B **24**, 1134 (1981).
2. G. Finkelstein, H. Strikman, and I. Bar-Josef, Phys. Rev. Lett. **74**, 976 (1995).
3. V. Huard *et al.*, Phys. Rev. Lett. **84**, 187 (2000).
4. S. I. Gubarev, I. V. Kukushkin, S. V. Tovstonog, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 469 (2000) [JETP Lett. **72**, 324 (2000)].
5. G. E. W. Bauer, Phys. Rev. B **45**, 9153 (1992).
6. D. A. Kleinman, Phys. Rev. B **32**, 3766 (1985).
7. R. J. Elliot, *Polarons and Excitons* (Oliver and Boyd, Edinburgh, 1963).
8. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory* (Nauka, Moscow, 1989, 4th ed.; Pergamon, New York, 1977, 3rd ed.).
9. H. Ehrenreich and M. H. Cohen, Phys. Rev. **115**, 786 (1959).
10. F. Stern, Phys. Rev. Lett. **18**, 546 (1967).
11. E. A. Andryushin and A. L. Silin, Fiz. Tverd. Tela (Leningrad) **21**, 219 (1979) [Sov. Phys. Solid State **21**, 129 (1979)].
12. K. S. Singwi, M. P. Tosi, R. H. Land, and A. Sjölander, Phys. Rev. **176**, 589 (1968).
13. K. Morawetz, cond-mat/0104229.
14. N. Iwamoto, E. Kroatscheck, and D. Pines, Phys. Rev. B **29**, 3936 (1984).
15. H. V. da Silvera, M. H. Degani, and K. S. Singwi, Phys. Rev. B **46**, 2995 (1992).
16. D. Pines and P. Nozières, *The Theory of Quantum Liquids* (Addison-Wesley, New York, 1968), Vol. 1.
17. R. Puff, Phys. Rev. **137**, A406 (1965).
18. J. P. Eisenstein, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **68**, 674 (1992).
19. A. L. Fetter, Phys. Rev. B **10**, 3739 (1974).
20. J. M. Ziman, *Principles of the Theory of Solids* (Cambridge Univ. Press, London, 1972; Mir, Moscow, 1974).
21. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1995; Pergamon, Oxford, 1980), Part 1.
22. L. V. Kulik, A. I. Tartakovskiĭ, A. V. Larionov, *et al.*, Zh. Éksp. Teor. Fiz. **112**, 353 (1997) [JETP **85**, 195 (1997)].
23. K. S. Singwi and M. P. Tosi, in *Solid State Physics*, Ed. by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1981), Vol. 36.

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 794–801.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 925–932.
Original Russian Text Copyright © 2002 by Elesin.

SOLIDS
Electronic Properties

High-Frequency Response of Two-Barrier Nanostructures

V. F. Elesin

Moscow Engineering Physics Institute (Technical University), Kashirskoe sh. 31, Moscow, 115409 Russia
e-mail: VEF@supercon.mephi.ru

Received December 13, 2001

Abstract—The linear rf response for a resonance-tunnel diode with asymmetric barriers is calculated analytically in the framework of a consistent quantum-mechanical model. It is shown that the response current is extremely sensitive to the asymmetry of the barriers. For example, if the "power" α_1 of the collector barrier becomes lower than the "power" α_2 of the emitter barrier (say, on account of bias voltage), the current reverses its sign at a certain frequency depending on the structure parameters. In the opposite case ($\alpha_1 \geq \alpha_2$), the sign of the current is preserved in the entire frequency range. This makes it possible to match, in principle, the experimental results obtained earlier with the theoretical results. At the same time, the quantum rf lasing mode of a resonance-tunnel diode, which was predicted earlier for $\alpha_1 = \alpha_2$, is realized for all values of α_1 and α_2. In this mode, high values of power can be attained at frequencies considerably higher than the resonance level width. It is also shown that the coherent amplification mechanism in resonance-tunnel diodes is closely connected with the quantum interference of resonantly tunneling electrons and differs significantly from the conventionally assumed mechanism. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The rf properties of two-barrier nanostructures including resonance-tunnel diodes still remain an unsolved theoretical problem. In spite of intense studies and an obvious practical interest in this field, no generally accepted theory of rf response and generation in resonance-tunnel diode has been proposed so far. Moreover, the publications contain contradictory results on the frequency dependence of the response even in an approximation linear in the field.

For example, it has been mentioned in the theoretical works [1–3] (numerical methods), [4] (analytical model), and [5, 6] (tunnel Hamiltonian method) that the polarization current (linear response) describing the amplification in a resonance-tunnel diode may reverse its sign at a certain frequency which is approximately equal to the resonance level width Γ. The sign reversal is also reported in some experimental works [7]. This leads to the conclusion about the existence of a limiting frequency of amplification and lasing for a resonance-tunnel diode. This is a widely accepted point of view (see, for example, [8]).

On the other hand, the experimentally attained lasing frequency 10^{12} s^{-1} [9] and the theoretical results [10–14] point to the contrary. Indeed, it is shown analytically and numerically [10–14] (see also [15]) that the current does not change its sign over a wide frequency range.

Leaving aside the publications [1–3] which are based on numerical methods (see the Conclusions below and [10]), it can be mentioned that either the above publications make use of the tunnel Hamiltonian method, or the Schrödinger equation is not solved explicitly in them. However, a consistent quantum-mechanical description and open boundary conditions are required for a system of electrons tunneling coherently in a resonance-tunnel diode and interacting with an electromagnetic field. This is due to quantum interference of electrons, which is quite sensitive to the energy of electrons supplied by the collector and to the boundary conditions. In our opinion, the approach developed in [4–6] does not satisfy the above requirements. On the contrary, the model used in [10–14] is quite rigorous.

However, the authors of [10–14] presumed complete symmetry of the barriers. It turned out that the asymmetry of the emitter and collector barriers (which is always observed in experiments and, apparently, in the numerical calculations [1–3] due to the bias field) may radically change the frequency dependence.

The present work aims at generalizing the results obtained in [10] to more general boundary conditions with a view to explain consistently the known theoretical and experimental data. The model [10] will be used to obtain an exact analytic solution and simple expressions for polarization currents. It will be shown that the response is extremely sensitive to the difference in the barrier parameters. For example, if the "intensity" α_1 of the collector barrier becomes smaller than the intensity α_2 of the emitter barrier (i.e., $\alpha_1 < \alpha_2$), the current reverses its sign at a certain frequency determined by barrier parameters. In the opposite case ($\alpha_1 \geq \alpha_2$), the current preserves its sign in the entire frequency range.

Thus, it becomes possible (at least in principle) to match the experimental results [7, 8] with numerical ([1–3, 11–13]) and analytical [10] theoretical results. Note that, according to [4–6], the response displays a weak dependence on the difference $\alpha_2 - \alpha_1$, thus demonstrating once again that the approach used for describing the coherent generation is not applicable (see the Conclusions for detail).

It is also shown that the mechanism of amplification in a resonance-tunnel diode is closely linked with quantum interference and differs significantly from the generally assumed mechanism.

2. WAVE FUNCTIONS OF A RESONANCE-TUNNEL DIODE IN THE FIELD-LINEAR APPROXIMATION

We will study a model of coherent tunneling in a resonance-tunnel diode, similar to the one used in [10]. Let us consider a 1D quantum well with δ-functional barriers at the points $x = 0$ and $x = a$ (see figure). A steady-state electron flux proportional to q^2 with an energy ε approximately equal to the energy ε_R of the resonance level is supplied to the quantum well from the left ($x \longrightarrow -\infty$). A varying electric field $E(t)$ with potential $U(x, t)$ acts in the region of the quantum well:

$$U(x,t) = U(x)\cos\omega t,$$
$$U(x) = \begin{cases} xU\theta(x), & x < a, \\ aU & x > a, \\ U = -eE/2. \end{cases} \tag{1}$$

The wave function $\Psi(x, t)$ satisfies the Schrödinger equation

$$i\frac{\partial\Psi}{\partial t} = -\frac{\partial^2\Psi}{\partial x^2} + [\alpha_2\delta(x) + \alpha_1\delta(x-a)]\Psi + U(x,t)\Psi. \tag{2}$$

Here, we have set $\hbar = 2m = 1$. The steady-state solution to Eq. (2) is sought in the form [2, 10, 16]

$$\Psi(x,t) = e^{-i\varepsilon t}[\psi_0(x) + e^{-i\omega t}\psi_{+1}(x) + e^{i\omega t}\psi_{-1}(x)]. \tag{3}$$

The partial wave functions ψ_0 and ψ_n ($n = \pm 1$) describe electrons with quasienergies ε and $\varepsilon + n\omega$, respectively.

The varying field induces polarization (response) currents

$$J^c(x,t) = J^c(x)\cos\omega t, \quad J^s(x,t) = J^2(x)\sin\omega t.$$

Here, J^c is the current synphase with the field and J^s is the reactive current. These currents can be expressed through the functions ψ_0 and ψ_n:

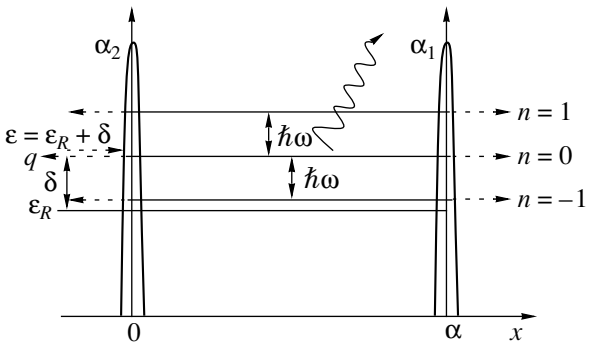

Figure.

$$J^c(x) = J^c_{+1}(x) + J^c_{-1}(x),$$
$$J^c_n(x) = -ie[(\psi_0^*\psi_n' + \psi_n^*\psi_0') - \text{c.c.}],$$
$$J^s(x) = J^s_{+1}(x) - J^s_{-1}(x), \tag{4}$$
$$J^s_n(x) = e[(\psi_0^*\psi_n' - \psi_n^*\psi_0') + \text{c.c.}], \quad \psi' \equiv \frac{d\psi}{dx}.$$

In the interval $0 < x < a$, the zeroth-order approximation function $\psi_0(x)$ satisfies the equation

$$\varepsilon\psi_0(x) + \psi_0''(x) = 0 \tag{5}$$

and the boundary conditions (see [8])

$$\psi_0(0)(1-\beta_2) + \psi_0'(0)/ip = q, \quad p^2 = \varepsilon,$$
$$\psi_0(a)(1-\beta_1) - \psi_0'(a)/ip = 0, \tag{6}$$
$$\beta_j = \alpha_j/ip, \quad i = 1, 2.$$

The corresponding equations and the boundary conditions for the functions $\psi_n(x)$ have the following form in the field-linear approximation:

$$p_n^2\psi_n(x) + \psi_n''(x) = U(x)\psi_0(x), \quad p_n^2 = p^2 + n\omega, \tag{7}$$
$$\psi_n(0)(1-\beta_{2n}) + \frac{\psi_n'(0)}{ip} = 0,$$
$$\psi_n(a)(1-\beta_{1n}) - \frac{\psi_n'(a)}{ip_n} = \frac{aU\psi_0(a)}{2p_n^2}, \tag{8}$$
$$\beta_{jn} = \frac{\alpha_j}{ip_n}, \quad j = 1, 2.$$

It is assumed in Eqs. (5)–(8) that $\psi_n \ll \psi_0$. The solution of this system of equations can be presented in the form

$$\psi_0(x) = A\exp(ipx) + B\exp(-ipx)$$
$$\equiv \gamma_0\cos px + i\delta_0\sin px, \tag{9}$$
$$A\Delta_0 = q(2-\beta_1)\exp(-2ipa), \quad B\Delta_0 = q\beta_1,$$

$$\Delta_0 = (2-\beta_1)(2-\beta_2)\exp(-2ipa) - \beta_1\beta_2 \approx \frac{2}{\sqrt{\Gamma_1\Gamma_2}}(i\delta - \Gamma), \quad (10)$$

$$\delta = \varepsilon - \varepsilon_R, \quad \Gamma = \Gamma_1 + \Gamma_2, \quad \Gamma_j = \frac{2p^3}{a\alpha_j^2},$$

$$\psi_n = A_n\exp(ip_nx) + B_n\exp(-ip_nx) - \frac{xU}{\omega_n}\psi_0 - \frac{2U}{\omega_n^2}\psi_0', \quad (11)$$

$$A_n\Delta_n = q_n(2-\beta_{1n})\exp(-2ip_na) + \beta_{2n}\tilde{q}_n,$$

$$B_n\Delta_n = q_n\beta_{1n} + (2-\beta_{2n})\tilde{q}_n, \quad \omega_n = -n\omega, \quad (12)$$

$$\Delta_n \approx \frac{2}{\sqrt{\Gamma_1\Gamma_2}}[i(\delta + n\omega) - \Gamma],$$

$$q_n = \frac{2Uip}{\omega_n^2}\left[A(2-\beta_{2n}) + \beta_{2n}B + \frac{\omega_n^2}{4p^4}(A+B)\right],$$

$$\tilde{q}_n = -\frac{2Uip}{\omega_n^2}\left[\beta_{1n}A + (2-\beta_{1n})Be^{-2ip_na} + \frac{\omega_n^2}{4p^4}(A + Be^{-2ipa})\right]e^{i(p-p_n)a}. \quad (13)$$

Formulas (9)–(13) give an exact solution to the problem, which is unfortunately cumbersome and hard to visualize. It was shown in [10], however, that the general formulas for ψ_n and currents J^c and J^s can be transformed to simple and physically visual expressions by using the natural small parameter ω/ε_R for a resonance-tunnel diode. Indeed, the smallness of frequency ω in comparison with the energy ε_R is inherent in a generator based on a resonance-tunnel diode.

In order to carry out this transformation, we present the quantities $\gamma_n = A_n + B_n$ and $\delta_n = A_n - B_n$ in the form of the sums of components

$$\gamma_n = \gamma_n^{(1)} + \gamma_n^{(2)} + \gamma_n^{(3)}, \quad \delta_n = \delta_n^{(1)} + \delta_n^{(2)} + \delta_n^{(3)}. \quad (14)$$

We set $p = p_n$ in the exponents of the components $\gamma_n^{(1)}$ and $\delta_n^{(1)}$ and form the difference between the exact and isolated expressions in $\gamma_n^{(2,3)}$ and $\delta_n^{(2,3)}$. Compensating a number of terms in $\gamma_n^{(1)}$ and $\delta_n^{(1)}$ and canceling out the determinant in the denominator, we obtain

$$\gamma_n^{(1)} = \frac{2Uip}{\omega_n^2}\delta_0, \quad \delta_n^{(1)} = \frac{2Uip}{\omega_n^2}\gamma_0. \quad (15)$$

Note that $\gamma_n^{(1)}$ and $\delta_n^{(1)}$ diverge for $\omega \longrightarrow 0$. The remaining terms are found to be finite in the low-frequency limit and are given by

$$\gamma_n^{(2)} = -\frac{4Uip}{\omega_n^2\Delta_n}[\beta_{1n}Az_n + (2-\beta_{1n})Be^{-2ip_na}z_n^*], \quad (16)$$

$$\delta_n^{(2)} = \gamma_n^{(2)}(\beta_{2n} - 1), \quad z_n = \exp(ipa - ip_na) - 1. \quad (17)$$

In these equations, we have omitted small terms of the order of ω/ε_R and Γ/ε_R. We will consider the most interesting case of a quantum well with "powerful" barriers in which $\Gamma/\varepsilon_R \ll 1$ and $p/\alpha_j \ll 1$. The remarkable properties of quantum wells are realized just in this limit. Taking into account the smallness of ω/ε_R and Γ/ε_R, Eq. (16) can be finally transformed as

$$\gamma_n^{(2)} = -\frac{4Ua^2\alpha_1 A}{p^2\Delta_n}. \quad (18)$$

It can be seen that the expression $\gamma_n^{(2)}$ and (in accordance with Eq. (17)) $\delta_n^{(2)}$ are finite for $\omega \longrightarrow 0$.

Partition of γ_n and δ_n (and, hence, of A_n and B_n) allows us to write the wave function $\psi_n(x)$ in a simpler form

$$\psi_n = \gamma_n^{(2)}\cos p_nx + i\delta_n^{(2)}\sin p_nx. \quad (19)$$

Indeed, it can be shown that the expressions in $\psi_n(x)$ diverging for $\omega \longrightarrow 0$ neutralize each other. Thus, the wave functions $\psi_n(x)$ are finite in the low-frequency limit and assume the following values at the boundaries of the well:

$$\psi_n(0) = \psi_n(a) = \gamma_n^{(2)}. \quad (20)$$

It should be noted that the functions $\psi_n(x)$ diverge for $\omega \longrightarrow 0$ in the publications [4–6] mentioned above. At the same time, it can be verified directly (by setting $\omega = 0$ in Eq. (1) from the very beginning) that the function ψ_n must not have singularities for $\omega = 0$. It is quite possible that this circumstance is responsible for the divergence of the frequency dependences of polarization currents.

Using Eqs. (9) and (20), we obtain the following linear-approximation criterion for $\omega \ll \Gamma$: $Ua/\Gamma \ll 1$. This criterion differs significantly from the corresponding criterion in [4–6]: $Ua/\omega \ll 1$. This difference is associated with the behavior of $\psi_n(x)$ for $\omega \longrightarrow 0$.

3. RF RESPONSE IN A RESONANCE-TUNNEL DIODE

To begin with, let us find the active component of the current J^c. Substituting ψ_0 from Eq. (9) and ψ_n from Eq. (19) into Eq. (4), we obtain

$$\begin{aligned} J_n^c(x) &= ep\{(K_n + \text{c.c.}) \\ &\times [\sin p_n x \sin px + \cos p_n x \cos px] \\ &- i(F_n - \text{c.c.})[\sin px \cos p_n x - \sin p_n x \cos px]\} \quad (21) \\ &\equiv ep\{(K_n + \text{c.c.})\cos(p - p_n)x \\ &- i(F_n - \text{c.c.})\sin(p - p_n)x\}, \end{aligned}$$

$$K_n = \delta_0^* \gamma_n^{(2)} + \gamma_0^* \delta_n^{(2)}, \quad F_n = \delta_0^* \delta_n^{(2)} + \gamma_0^* \gamma_n^{(2)}. \quad (22)$$

The contribution to the current $J_n^c(x)$ comes from four types of components: field-induced transition between states with wave functions $\sin px$ and $\sin p_n x$ with the weight $\delta_0^* \delta_n^{(2)}$; between $\cos px$ and $\cos p_n x$ with the weight $\gamma_0^* \gamma_n^{(2)}$; between $\sin px$ and $\cos p_n x$ with the weight $\delta_0^* \gamma_n^{(2)}$; and, finally, between $\cos px$ and $\sin p_n x$ with the weight $\gamma_0^* \delta_n^{(2)}$. It should be noted that the term in the current, which is proportional, for example, to $\sin px \cos p_n x$, appears as a result of a transition between the states $\sin px$ and $\sin p_n x$ since $J_n^c \propto (\psi^* \psi' - \text{c.c.})$. This term corresponds to a "laser"-type transition since the wave functions $\sin px$ and $\sin p_n x$ coincide with the eigenfunctions of an isolated quantum well. True, in the case under investigation, the momenta p and p_n differ by a small quantity $\omega_n/2p$. (For a laser, we have $p - p_n = \pm\pi/a$). Since the coefficients $\delta_0 \propto \beta\gamma_0$ and $\delta_n^{(2)} \sim \beta\gamma_n^{(2)}$, the contribution from this term is larger as compared to that from the second term $\gamma_0^* \gamma_n^{(2)}$ (between $\cos px$ and $\cos p_n x$) in the parameter $\alpha^2/p^2 \gg 1$.

The terms between the "mixed" states (typical of the current state in a resonance-tunnel diode and vanishing in an isolated well) $\sin p_n x$ and $\cos px$, $\sin px$ and $\cos p_n x$ appear with approximately equal weights $\delta_0^* \gamma_n^{(2)}$ and $\gamma_0^* \delta_n^{(2)}$, which permits the effective interference between these terms. It is precisely these transitions that lead to the peculiar frequency dependence $J_n^c(x)$ and to exceptional sensitivity of current to the difference $\alpha_2 - \alpha_1$ between the "powers" of the emitter (α_2) and collector (α_1) barriers of the resonance-tunnel diode.

Substituting Eq. (18) into Eq. (22), we arrive at the expression for K_n,

$$K_n = \frac{q}{\Delta_0^*} \gamma_n^{(2)} \varphi, \quad \frac{\varphi}{\Delta_0^*} = \beta_{2n}(A^* + B^*) - 2B^*, \quad (23)$$

$$\begin{aligned} \varphi &= \beta_2(2 + \beta_1)\exp(2ipa) - \beta_1\beta_2 + 2\beta_1 \\ &= \Delta_0^* - 2\Delta_{01}^* \exp(ipa), \end{aligned} \quad (24)$$

disregarding, as before, the terms of the order of ω/ε_R. Here, Δ_0 is defined by Eq. (10), and Δ_{01} is a "truncated determinant":

$$\Delta_{01} = (2 - \beta_1)\exp(-ipa) + \beta_1 \exp(ipa). \quad (25)$$

The function φ from (24) describes the superposition of the above-mentioned "nonlaser-type" transitions and strongly depends on the differences $\delta = \varepsilon - \varepsilon_R$ and $\alpha_2 - \alpha_1$. Here, ε_R is the energy of the resonance level, which is determined from the equation

$$\text{Re}\Delta_0(\varepsilon_R) = 0. \quad (26)$$

In the vicinity of the resonance and for $\Gamma/\varepsilon_R \ll 1$, the function φ can be presented in the form

$$\varphi = -\frac{i\alpha_1\alpha_2 a\delta}{p^3} + 2\left(\frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1}\right). \quad (27)$$

For $\alpha_2 = \alpha_1 = \alpha$, expression (23) is transformed to the corresponding expression for K_n, derived in [10].

Let us first consider the situation when $\alpha_2 = \alpha_1$. It can be seen from Eq. (27) that, at the resonance (when $\varepsilon = \varepsilon_R$), we have $\varphi = 0$. This means that the contributions to current J_n^c for $\cos(p - p_n)x$ (which are virtually independent of the coordinate and are the main contributions for $\omega \ll \Gamma$) are equal to zero separately (both J_{+1}^c and J_{-1}^c). If $\delta \neq 0$, both currents J_{+1}^c and J_{-1}^c differ from zero and have the same sign. It will be shown below that this corresponds to emission for $\delta > 0$ and absorption for $\delta < 0$. However, it is usually assumed that J_{+1}^c leads to absorption and J_{-1}^c to emission [4–6, 8], and the resultant sign of the response is determined by their difference. By the way, this made it possible to obtain finite expressions for current for $\omega \longrightarrow 0$ in [4–6], although the wave functions diverge (see Section 2).

On the other hand, in accordance with Eqs. (23), (27), and (18), the signs of the low-frequency contributions to J_n^c are identical and are determined by the interference of "nonlaser-type" transitions, which depends on the resonance conditions (i.e., on δ).

Let us now consider the effect of the boundaries. It follows from Eq. (27) that the difference between α_1 and α_2 leads to the emergence of an imaginary correc-

tion to δ, which is a function of the difference $\alpha_1/\alpha_2 - \alpha_2/\alpha_1$:

$$\varphi = -\frac{2i}{\sqrt{\Gamma_1\Gamma_2}}\left[\delta + i\sqrt{\Gamma_1\Gamma_2}\left(\frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1}\right)\right]. \quad (28)$$

Substituting the values of $\gamma_n^{(2)}$ from Eq. (18) and φ into Eq. (22), we arrive at the following expression for K_n + c.c.:

$$K_n + \text{c.c.} = -\frac{4U_a}{|\Delta_0|^2\Gamma_1^{3/2}\Gamma_2^{1/2}}\left[\delta\left(\frac{1}{\Delta_n} + \frac{1}{\Delta_n^*}\right) + i(\Gamma_1\Gamma_2)^{1/2}\left(\frac{\alpha_1}{\alpha_2} - \frac{\alpha_2}{\alpha_1}\right)\left(\frac{1}{\Delta_n} - \frac{1}{\Delta_n^*}\right)\right]. \quad (29)$$

Similarly, we can find the expression for "laser-type" transitions

$$F_n - \text{c.c.} = \frac{8iUp}{|\Delta_0|^2\Gamma_1^{3/2}\Gamma_2^{1/2}}\left(\frac{1}{\Delta_n} - \frac{1}{\Delta_n^*}\right), \quad (30)$$

which depends on the difference $\alpha_2 - \alpha_1$ only slightly. We can also prove that the contribution to J_n^c from $(F_n - \text{c.c.})$ gives absorption for $n = +1$ and emission for $n = -1$ in accordance with the conventional concepts.

Substituting Δ_n and Δ_0 into Eqs. (29) and (30) and summing up the results, we find the final expression for the current $J^c(x) = J_{+1}^c(x) + J_{-1}^c(x)$ and the reduced current J^c:

$$J^c(x) = -\frac{e^2EaQ\Gamma\Gamma_2\delta}{(\delta^2 + \Gamma^2)[(\delta+\omega)^2 + \Gamma^2][(\delta-\omega)^2 + \Gamma^2]} \times\left\{\left[(\delta^2 + \Gamma^2 + \omega^2) - \frac{\alpha_1^2 - \alpha_2^2}{\alpha_1^2 + \alpha_2^2}(\delta^2 + \Gamma^2 - \omega^2)\right] \times\cos\frac{\omega}{2p}x - \frac{4\omega p}{a}\sin\frac{\pi}{2p}x\right\}, \quad (31)$$

$$J^c = \frac{1}{a}\int_0^a J^c(x)dx = -\frac{e^2EaQ\Gamma\Gamma_2\delta\left[(\Gamma^2 + \delta^2) + \frac{\Gamma_1 - \Gamma_2}{\Gamma_1 + \Gamma_2}(\Gamma^2 + \delta^2 - \omega^2)\right]}{(\Gamma^2 + \delta^2)[(\delta+\omega)^2 + \Gamma^2][(\delta-\omega)^2 + \Gamma^2]}, \quad Q = pq^2. \quad (32)$$

The contribution from the "laser" term $(F_n - \text{c.c})$ to the reduced current J_n^c, which is proportional to ω^2, is completely compensated by the corresponding term $(K_n + \text{c.c})$. Consequently, the resultant expression (32) is of interference origin and stems from $(K_n + \text{c.c})$.

For identical barriers $\Gamma_1 = \Gamma_2 = \Gamma/2$, we arrive at the result which was obtained for the first time in [10]:

$$J^c = -\frac{e^2EaQ\Gamma^2\delta}{2[(\delta+\omega)^2 + \Gamma^2][(\delta-\omega)^2 + \Gamma^2]} \quad (33)$$

(after correcting the misprint: the denominator of expression (33) must contain 2 instead of 4). It can easily be seen that the current $J^c(\delta, \omega)$ does not change its sign in the entire frequency range. In the low-frequency limit, $\omega \ll \Gamma$, the quantity J^c can be expressed in terms of the static differential conductivity:

$$J^c(\delta, 0) = \frac{e^2Ea}{2}\frac{dJ_0(\delta)}{d\delta}, \quad (34)$$

$$J_0(\delta) = \frac{Q\Gamma^2}{2(\delta^2 + \Gamma^2)}, \quad (35)$$

where $J_0(\delta)$ is the static resonance current.

It was shown in [10] that, in addition to the conventional mode in which J^c has a peak for $\omega = 0$ (and for $\delta < \Gamma$), the so-called quantum mode also exists for $\delta > \Gamma$. It corresponds to the peak of J^c at a frequency ω_m:

$$\omega_m^2 = \delta^2 - \Gamma^2, \quad \delta > \Gamma. \quad (36)$$

Emission (absorption) occurs owing to quasi-resonant transitions between the states with energies ε and ε_R. It follows hence that for identical barriers $(\alpha_2 = \alpha_1)$, lasing is possible at frequencies considerably exceeding Γ if we choose the electron energy (an analogue of the constant bias voltage) $\varepsilon = \varepsilon_R + \omega$ outside the region of the maximum negative differential conductivity (where $\delta < \Gamma$).

Note that the results described above were confirmed to a high degree of accuracy by the numerical solution of the system of equations (5)–(8) and also directly by the solution to the time-dependent equation (2) in [12–14].

In the case of different barriers $(\Gamma_1 \neq \Gamma_2)$, the frequency dependence of current J^c may change radically. For example, for $\Gamma_1 > \Gamma_2$, the current changes its sign for a certain value of frequency ω_0:

$$\omega_0^2 = \frac{2\Gamma_1(\delta^2 + \Gamma^2)}{\Gamma_1 - \Gamma_2}. \quad (37)$$

At the same time, the "quantum" mode is realized for any values of Γ_1 and Γ_2. In particular, for $\Gamma_1 \gg \Gamma_2$, the

current peak is attained at the frequency $\omega_m^2 = 2(\delta^2 + \Gamma^2) - (\delta^4 + 10\delta^2\Gamma^2 + 9\Gamma^4)^{1/2}$. (If $\delta \gg \Gamma$, we have $\omega_m \approx \delta$.)

If the opposite inequality $\Gamma_1 < \Gamma_2$ is satisfied, the current is preserved for any frequency value. In the limiting case, when $\Gamma_2 \gg \Gamma_1$, the frequency dependence becomes unusual for a resonance-tunnel diode:

$$J^c = -\frac{e^2 EaQ\Gamma_2^2\delta\omega^2}{(\Gamma^2 + \delta^2)[(\delta + \omega)^2 + \Gamma_2^2][(\delta - \omega)^2 + \Gamma_2^2]}. \quad (38)$$

Indeed, for $\omega \longrightarrow 0$, the current vanishes and cannot be expressed in terms of differential conductivity. Only the quantum mode remains, the peak of $J^c(\delta, \omega)$ being attained for frequency

$$\omega_m^2 = \delta^2 + \Gamma^2 \quad (39)$$

for any $\delta > 0$.

Let us also determine the reactive current. After some calculations, we obtain

$$J_n^s(x) = ep[i(K_n - \text{c.c.})\cos(p - p_n)x + (F_n + \text{c.c.})\sin(p - p_n)x] \quad (40)$$

and the following expression for the reduced reactive current:

$$J^s = -\frac{e^2 EaQ\Gamma_2\delta\omega[\delta^2 - \omega^2 - 3\Gamma^2 + 4(\Gamma_2^2 - \Gamma_1^2)]}{2(\Gamma^2 + \delta^2)[(\delta + \omega)^2 + \Gamma^2][(\delta - \omega)^2 + \Gamma^2]}. \quad (41)$$

For $\Gamma_1 = \Gamma_2$, this expression is transformed into the corresponding expression derived in [10] after the appropriate sign reversal.

4. COMPARISON WITH THE RESULTS OBTAINED BY OTHER AUTHORS

It was proved by a detailed analysis in [10] and mentioned above in the Introduction that no unified approach to the limitation of the lasing frequency of a resonance-tunnel diode has been worked out. According to the conventionally used hypothesis (see, for example, [4, 8]), the lasing frequency (i.e., the frequency at which the current reverses its sign and the gain vanishes) is limited by a quantity equal to Γ.

This concept is based, among other things, on the results of theoretical works [1–4] in which the Schrödinger equation was solved as well as on the results of publications using the tunnel Hamiltonian method [5, 6].

Unfortunately, it is difficult to carry out a direct comparison with the results of numerical calculations, which are themselves contradictory (see, for example, [8, 17]). It was mentioned by us earlier in Section 3 that a possible reason for the sign reversal of the response in [1–3] is the asymmetry of the barriers due to the bias voltage.

It is important, in our opinion, to carry out a comparison with the theoretical results obtained analytically in the simplest formulation of the problem in order to eliminate the effect of nonessential complications. In doing so, we assume (as before) that the electron distribution function for the emitter is δ-shaped; i.e., electrons are assumed to be monoenergetic with energy ε. Then we carry out a comparison for other distribution functions.

In respect of the formulation of the problem, the work by Lju [4], who considered an analytical model of a resonance-tunnel diode on the basis of the Schrödinger equation, is the closest to our research (see also the references to previous publications cited in [4]). However, in contrast of the present work and [10], Lju [4] could not find an explicit solution of the Schrödinger equation in the quantum well region. He describes the wave function at the collector boundary ($x = a$) in the form

$$\Psi = [t_0 e^{-ipx} + t_{+1} e^{-ip_{+1}x - i\omega t} + t_{-1} e^{-ip_{-1}x - i\omega t}] \times \exp\left(i\varepsilon t - \frac{iV\sin\omega t}{\omega}\right), \quad (42)$$

where t_0 and $t_{\pm 1}$ are the amplitudes of electron transition through a well in zero field and in the first order in the field, respectively (analogues of our ψ_0 and $\psi_{\pm 1}$). The structure was assume to have identical values of $\Gamma_1 = \Gamma_2$.

The amplitudes were determined by summing the transmitted and reflected waves (the Fabry–Perrot resonator model) and were found to be

$$|t_{\pm 1}|^2 = \left(\frac{V}{2\omega}\right)^2 \frac{\Gamma^2[(\delta \pm \omega/2)^2 + \Gamma^2]}{(\delta^2 + \Gamma^2)[(\delta \pm \omega)^2 + \Gamma^2]}. \quad (43)$$

These amplitudes differ considerably from our amplitudes (see Eqs. (18) and (20))

$$|\psi_{\pm 1}(a)|^2 = |\gamma_\pm^{(2)}|^2 = \frac{(Va)^2\Gamma^2}{16(\delta^2 + \Gamma^2)[(\delta \pm \omega)^2 + \Gamma^2]}, \quad (44)$$

which were calculated according to the exact solution of the Schrödinger equation (2) with the boundary conditions (6) and (8). The basic difference lies in the divergence of $t_{\pm 1}$ for $\omega \longrightarrow 0$. The expression for current J^c from [4] has the form

$$J^c = -\frac{epV\Gamma^2\delta(\delta^2 + \Gamma^2 - \omega^2)}{(\Gamma^2 + \delta^2)[(\delta + \omega)^2 + \Gamma^2][(\delta - \omega)^2 + \Gamma^2]}. \quad (45)$$

This expression implies the sign reversal for $\omega_0^2 = \delta^2 + \Gamma^2$, the existence of the limiting frequency, and the absence of the quantum mode. The reasons for the difference between this relation and (33) are apparently associated with the approximations used in [4] (see for details the analysis in [10]). The remaining theoretical publications using the Schrödinger equation and known

to us do not give closed expressions for currents J^c and J^s.

In a large number of theoretical publications devoted to the calculation of the rf response in a resonance-tunnel diode, the method of tunnel Hamiltonian is used (see, for example, [5, 6] and the references cited therein).

The expression for current i_2 (analogue of J^c) derived in [5, 6] for a δ-shaped electron distribution function has the form

$$i_2 = -\frac{2e^2V\Gamma_1\Gamma_2\delta(\delta^2+\Gamma^2-\omega^2)}{(\Gamma^2+\delta^2)[(\delta+\omega)^2+\Gamma^2][(\delta-\omega)^2+\Gamma^2]}. \quad (46)$$

In deriving this expression, we assumed that a varying field is applied only to the emitter. It should be noted above all that the response i_2 weakly depends on the difference $\Gamma_1 - \Gamma_2$ and is similar to expression (45) from [4] for $\Gamma_1 = \Gamma_2$. It should be recalled that it was assumed in [4] that the field is applied to the entire well.

In order to carry out a comparison with [5, 6], we solved the Schrödinger equation (2) with the local potential

$$U(x, t) = \bar{U}\delta(x)\cos\omega t. \quad (47)$$

The expression for the reduced current has the form

$$J^c = -\frac{8\bar{U}(\Gamma_1\Gamma_2)^{3/2}\delta/p[\delta^2+\Gamma^2+\omega^2(3\Gamma/2\Gamma_1-1)]}{a(\Gamma^2+\delta^2)[(\delta+\omega)^2+\Gamma^2][(\delta-\omega)^2+\Gamma^2]}. \quad (48)$$

It can be seen that, in contrast to Eqs. (46) and (32), the current does not reverse its sign in the entire frequency range and weakly depends on the difference $\Gamma_1 - \Gamma_2$. In addition, current (48) is of the order of smallness Γ/ε_R as compared to expressions (46) and (32). The reason for this is obvious and lies in the following. An increase in function ψ_n by the factor ε_R/Γ due to the resonance in the well with the field applied to the entire well is absent if the field $U(x, t)$ in (47) is local. Thus, in the tunnel Hamiltonian method, the response weakly depends on $U(x, t)$ in contrast to the exact result. In addition, note that the wave functions diverge as $\omega \longrightarrow 0$ both in [5, 6] and in [4].

Let us also compare the responses for the electron energy distribution with the quasi-equilibrium function $f(\varepsilon/T)$ (T is the temperature). In the high-temperature limit, when $\omega, \Gamma \ll T$, we obtain the following expression instead of (46) for $\Gamma_1 = \Gamma_2 = \Gamma/2$:

$$\bar{i}_2 = e^2\frac{\partial f(\varepsilon_R)}{\partial\varepsilon}\frac{V\Gamma^3\pi}{4(\omega^2+\Gamma^2)}. \quad (49)$$

It can be seen that the amplification is limited by the frequency $\omega \approx \Gamma$. Carrying out a similar integration for Eq. (33), we obtain

$$J^c = \frac{e^2EaQ\Gamma\pi}{4}\frac{\partial f(\varepsilon_R)}{\partial\varepsilon}; \quad (50)$$

i.e., there is no limitation on the amplification in frequency. Apparently, the above contradictions are associated with the fact that the interference phenomenon and the open boundary conditions are taken into account incorrectly in the tunnel Hamiltonian method. Indeed, this method assumes the existence of a resonance level, and the jump between the well and the emitter (collector) is introduced phenomenologically. In fact, the tunnel Hamiltonian method described incoherent tunneling. The proof of this statement (at least, for $\omega < \Gamma$) is given in [18].

5. CONCLUSIONS

The expressions obtained for the polarization currents for asymmetric barriers make it possible in principle to explain consistently the experimentally obtained results as well as the results of numerical and analytical solution of the Schrödinger equation. Indeed, we can assume that the asymmetry in the barriers appeared in [1–3] and in [7] due to a constant electric bias field, which lowers the collector barrier (i.e., increases the value of Γ). A detailed comparison requires special computations and experiments with controllable Γ_1 and Γ_2, i.e., with real values of the emitter and collector barriers. It should also be interesting to verify the amplification in a resonance tunnel diode for $\Gamma_2 \gg \Gamma_1$, when the low-frequency amplification must tend to zero as ω^2. Obviously, this result will remain valid for any form of the electron distribution function $f(\varepsilon)$. It is also important to emphasize that the quantum lasing mode is preserved for all values of Γ_1 and Γ_2. The frequency for which the amplification satisfies the quasi-resonance conditions $\omega_m \approx \delta$ most closely is equal to $\omega \gg \Gamma$. Thus, according to our earlier results [12, 13], high lasing powers can be attained for ultrahigh frequencies.

A confirmation of the results predicted by the theory would serve as proof of a specific emission and absorption mechanism in structures with coherent resonance tunneling, which is connected with the purely quantum phenomenon of superposition of various types of radiative transitions.

As regards the tunnel Hamiltonian method, the correctness of its application for computing the rf response in resonance-tunnel-diode-type systems seems to be disputable. These systems are exceptionally sensitive to the boundary condition, the actual form of the variable field potential, and the correct description of the spatial quantization phenomenon. The tunnel Hamiltonian method is essentially phenomenological since the resonance level and the boundary conditions (replaced by the electron jump) are postulated.

ACKNOWLEDGMENTS

The author is obliged to Yu.V. Kopaev for fruitful discussions of the results.

This work was carried out in the framework of the special federal program "Integration" supported by the Ministry of Industry and Science of the Russian Federation under the program "Physics of Solid Nanostructures."

REFERENCES

1. W. R. Frensley, Appl. Phys. Lett. **51**, 448 (1987); Rev. Mod. Phys. **62**, 745 (1990).
2. R. K. Mains and G. I. Haddad, J. Appl. Phys. **64**, 3564 (1988); **64**, 504 (1988).
3. C. L. Fernando and W. R. Frensley, Phys. Rev. B **52**, 5092 (1995).
4. H. C. Liu, Phys. Rev. B **43**, 12538 (1991); Erratum **48**, 4977 (1993).
5. M. Büttiker, A. Pretre, and H. Thomas, Phys. Rev. Lett. **70**, 4114 (1993).
6. M. P. Anatram and S. Datta, Phys. Rev. B **51**, 7632 (1995).
7. J. P. Mattia, A. L. McWhorter, R. J. Aggarwal, *et al.*, J. Appl. Phys. **84**, 1140 (1998).
8. H. C. Lju and J. C. L. Sollner, Semicond. Semimet. **41**, 359 (1994).
9. E. R. Brown, J. R. Södestrom, C. D. Parker, *et al.*, Appl. Phys. Lett. **58**, 2291 (1991).
10. V. F. Elesin, Zh. Éksp. Teor. Fiz. **116**, 704 (1999) [JETP **89**, 377 (1999)].
11. V. F. Elesin and A. V. Krasheninnikov, Phys. Low-Dimens. Struct. **7/8**, 65 (1999).
12. V. F. Elesin, I. Yu. Kateev, and A. I. Podlivaev, Fiz. Tekh. Poluprovodn. (St. Petersburg) **34**, 1373 (2000) [Semiconductors **34**, 1321 (2000)].
13. V. F. Elesin, I. Yu. Kateev, and A. I. Podlivaev, Usp. Fiz. Nauk **170** (3), 333 (2000).
14. R. F. Kazarinov and R. A. Suris, Fiz. Tekh. Poluprovodn. (Leningrad) **6**, 148 (1972) [Sov. Phys. Semicond. **6**, 120 (1972)].
15. N. S. Wingreen, A. P. Janho, and Y. Meir, Phys. Rev. B **48**, 8487 (1993).
16. V. F. Elesin, Zh. Éksp. Teor. Fiz. **112**, 483 (1997) [JETP **85**, 264 (1997)].
17. F. A. Buot and A. R. Rajagopal, Phys. Rev. B **48**, 17217 (1993).
18. V. V. Afonin and A. M. Rudin, Phys. Rev. B **49**, 10466 (1994).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 802–820.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 933–954.
Original Russian Text Copyright © 2002 by Zavaritsky.

SOLIDS
Electronic Properties

Interlayer Magnetoresistance in the Normal and Mixed States of a $Bi_2Sr_2CaCu_2O_y$ Crystal with $T_c \geq 91$ K

V. N. Zavaritsky

Kapitza Institute for Physical Problems, Russian Academy of Sciences, Moscow, 117334 Russia
Loughborough University, Loughborough LE11 3TU, United Kingdom
e-mail: vlad@kapitza.ras.ru

Received October 25, 2001

Abstract—The effect of a strong magnetic field ($B \leq 55$ T), applied perpendicularly to the basal plane of a $Bi_2Sr_2CaCu_2O_y$ (BSCCO-2212) crystal with $T_c > 91$ K, on the out-of-plane electric conductivity of the sample was studied under the conditions of controlled ohmic crystal response and negligibly small induction overheating. The character of the field and temperature dependences of the sample resistance was studied in the region of the resistive state formation, in the flux flow regime, in the normal state, and under the conditions of superconductivity suppression by a strong magnetic field. Regular methods are proposed for estimating the interlayer resistance in the absence of superconductivity ($R_N(0, T)$) and the critical field for a superconducting nucleus formation ($H_{c2}(T)$). The results show a satisfactory agreement with published data. In particular, the $H_{c2}(T/T_c)$ estimates exhibit quantitative correlation in a broad range of dopant concentrations; the $H_{c2}(T)$ dependences determined by the proposed regular method using the plane and interlayer resistances of BSCCO-2201 show qualitative agreement. It is demonstrated that a self-consistent description of the whole body of experimental data can be provided within a restricted class of theoretical models. A new interpretation of the origin of the "quasiparticle" resistance of BSCCO-2212 is proposed, and it is established that the experimental $H_{c2}(T)$ curve can be satisfactorily described in terms of a one-parametric dependence predicted for the second critical field in the bosonic limit. An alternative explanation of the results, based on the theory of interlayer tunneling of the Cooper pairs and quasiparticles in a layered superconductor with *d* coupling, is analyzed in detail, and it is shown that the main conclusions of this model contradict the observed experimental facts. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The *H–T* diagram of cuprate high-temperature superconductors (HTSCs) combining high critical temperatures T_c of the superconducting transition and small coherence lengths ξ significantly differs from that of the "conventional" type II superconductors. The former phase diagram contains, in addition to the lines of critical fields $H_{c1}(T)$ and $H_{c2}(T)$, a line corresponding to the so-called irreversibility field $H_{irr}(T)$, which lies far away from the upper critical field $H_{c2}(T)$ and separates the region of existence of a vortex lattice and a nondamping supercurrent from a large domain of fields and temperatures featuring the state of a vortex fluid with ohmic response. The irreversibility field, like other features of the vortex system, was thoroughly studied both experimentally and theoretically for various HTSCs [1]. The experimental data indicate, in particular, that the lattice melting in perfect crystals of $Bi_2Sr_2CaCu_2O_y$ (BSCCO-2212) proceeds as a first-order phase transition [2, 3].

In contrast to the $H_{irr}(T)$ value, the magnitude of the upper critical field in HTSC and the character of its temperature variation are still under discussion. For example, some researchers believe that this field is not manifested experimentally because of the fluctuation effects [4, 5]. This notion is usually justified by the character of the magnetic field effect upon the superconducting transition smearing. As is known, a magnetic field applied to a conventional type II superconductor shifts the transition curve as a whole, while the initial part of the temperature dependence of magnetization is well approximated by a linear relationship, $M(T, B) \propto H_{c2}(T) - B$, thus admitting the direct measurement of H_{c2}. In contrast, the experiments with HTSC cuprates reveal a pronounced "rounding" of the curves, which increases with the field. Moreover, for the most anisotropic compounds such as BSCCO-2212, it was established that the $M(T)$ curves measured in the mixed state possess a common point of intersection (M^*, T^*) in a broad range of magnetic fields (10^{-3} T < B < 5 T) [6].

A satisfactory character of the approximation of a fragment of the experimental $M(T)$ curve provided, in a narrow temperature interval near T_c, by a three-parametric relationship predicted within the framework of the Bardeen–Cooper–Schrieffer (BCS) theory for a fluctuation contribution [7] to the magnetization of a layered quasi-two-dimensional superconductor in the normal state is frequently considered as direct evidence for adequacy of this model. Based on this opinion, most of the known estimates for $H_{c2}(0)$ were obtained from

an analysis of the fluctuation contribution to the magnetoresistance and magnetization in the vicinity of the critical temperature [6, 8, 9]. At the same time, it was noted [6, 8, 9] that the coefficients of such an approximation ($\xi_{ab}^2(0)$, $\xi_c^2(0)$, and $T_c(H)$) turn out to be field-dependent, so that each curve corresponds to an individual set of parameters; the values reported usually correspond to selecting a certain field, typically about 10–15 kOe. This choice is probably determined by the possibility of matching to the results of extrapolation [10, 11] of the $T_c(H)$ curve at T_c (obtained in the course of the above analysis) to zero temperature. Such a matching procedure is very sensitive to the value of T_{c0} (the transition temperature in a zero field) and requires this temperature to be overstated by 3–5 K as compared to the measured value (see, e.g., [8]).

Moreover, it was established for a large number of compounds representing various classes of HTSC cuprates that the results of a consistent analysis of the fluctuation corrections to the magnetization [12], thermal conductivity [13], and resistivity [14] contradict the BCS theory underlying all variants of the aforementioned analysis—both classical and those specially adapted to features of the HTSC materials studied [7, 15, 16].[1] For example, in a most thorough (to the author's knowledge) investigation of the magnetization of BSCCO-2212 crystals [6], the correlation lengths $\xi_{ab}(0)$ and $\xi_c(0)$ turned out to be on the order of the lattice parameter, the $H_{c2}(0)$ value proved to be close to or higher than the theoretical paramagnetic threshold, and the asymptotic dependences of $H_{c2}(T)|_{T \to T_c}$ exhibited a significantly nonlinear character in that very temperature interval where the best agreement with the theory (predicting the linear dependence) could be anticipated. Thus, the discrepancies between the results of experimental data treatment and theoretical premises apparently indicated inapplicability of the BCS theory to description of the properties of HTSC cuprates.

This conclusion is corroborated by the results of experimental determination of the upper critical field from variations of the magnetoresistance. By analogy with conventional superconductors, H_{c2} was estimated from a change in the superconducting transition temperature caused by an applied magnetic field. The temperature dependence of H_{c2} determined for a $Tl_2Ba_2Cu_1O_y$ crystal exhibited a negative curvature in a record-wide temperature interval, $10^{-3} \le T/T_c \le 1$ [17], which disagrees with predictions of both the BCS theory and the tight binding model. Later [18–20], the same anomalous character of the $H_{c2}(T)$ curves was reproduced in overdoped representatives of some other groups of HTSC cuprates. It should be noted that, in order to increase the experimentally accessible T/T_c range, these experiments were performed on the samples with suppressed critical temperature, although the employed method of doping still allowed the phases with different T_c to coexist. As was theoretically demonstrated using an idealized model [21], this inhomogeneity can, in principle, lead to analogous experimental dependences in the BCS theory as well. Such a scenario was *a priori* inapplicable to interpretation of the results obtained with optimum doped crystals [22]; however, a large number of assumptions and complexity of a multistep empirical extrapolation procedure employed in [22] required at least a more direct verification.

An attempt at such a verification is made in this study. Based on an analysis of the temperature-induced transformation of the magnetoresistance of a BSCCO-2212 crystal, the character of the sample magnetoresistance is determined in the vortex (fluxoid) flow regime, in the normal state, and under the conditions of superconductivity suppression by a strong magnetic field. From this analysis, the contribution of the superconducting state to the magnetoresistance at $T < T_{c0}$ was separated, the temperature dependence of the characteristic fields (identified with the field of formation of a superconducting nucleus and the irreversibility line) was determined, and the interlayer resistance in the absence of superconductivity was estimated. The obtained data show a satisfactory agreement with the results of independent investigations. A comparison to the theory allowed the class of models applicable to description of the properties of BSCCO-2212 crystals to be significantly reduced. The results presented below were partly reported in [23–25]. A systematic character of the present analysis, based on the data obtained for a large number of crystals, suggests that the results reflect the properties of the compound studied, rather than the individual features of particular samples. The field of a superconducting nucleus formation was evaluated using the point of intersection of the approximated curves representing the asymptotic behavior of the sample magnetoresistance in the regions of $H \ll H_{c2}$ and $H \gg H_{c2}$, which eliminates to a certain extent the uncertainty related to the fluctuational smearing of the transition.

The paper is organized as follows. The Introduction is followed by Section 2, describing the methods of measurement using pulsed magnetic fields. The main attention is devoted to the resistance measurements, the problem of the sample heating by the eddy currents, and the methods of controlling linearity of the system response. For a convenient presentation of the experimental results in Section 3, data on the properties of samples in the normal state, the features of the resistive response formation, and the flux flow region are described in separate subsections. Section 4 is devoted to a comparison of the obtained results to the published data, both experimental and theoretical; the final part of this section addresses the experimental verification of

[1] When applied to the interpretation of experimental data, the model proposed in [15] requires (besides leading to the aforementioned discrepancies) significantly (by approximately 30%) overstated values of the interlayer distance.

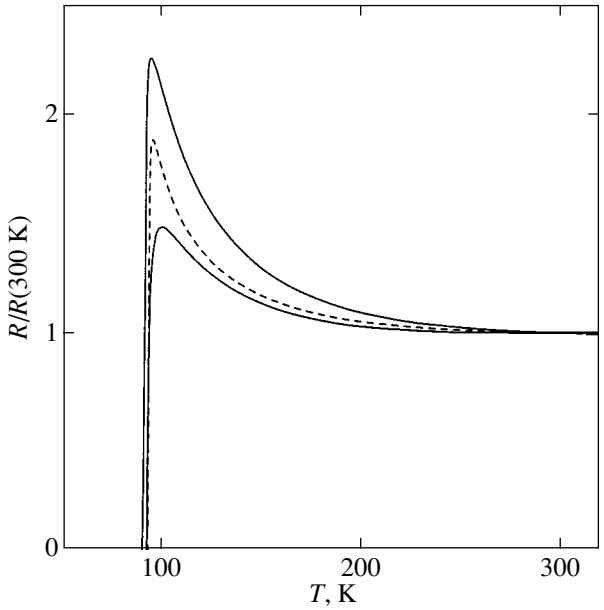

Fig. 1. Typical temperature dependences of the out-of-plane resistance for three BSCCO-2212 crystals measured in a zero magnetic field.

applicability of the theory of interlayer transport in a superconductor featuring nontrivial coupling.

2. EXPERIMENTAL METHODS

2.1. Samples and Contacts

We have studied the samples of crystals with a stoichiometric composition $Bi_2Sr_2CaCu_2O_y$ (BSCCO-2212) possessing $T_c > 91$ K. The high critical temperature, conventionally related to an optimum level of hole doping [26], was provided by partially substituting yttrium for calcium in the crystal lattice sites [27] immediately in the course of solid state synthesis.[2] In order to reduce the crystal overheating by the induction currents generated in a pulsed field, the experiments were performed on small samples cut from large crystals selected by the criteria of macroscopically homogeneous composition and the absence of blocks. The main results were obtained for a series of seven samples with thicknesses (along the *c* axis) from 0.8 to 5 μm. The samples were prepared by mechanically splitting large crystals along the (0 0 1) cleavage plane, followed by cutting almost rectangular fragments with dimensions in the basal plane varying from 26×28 to 32×60 μm^2.

The current and potential leads were glued to each freshly cleaved plane *ab* of the samples with the aid of a DuPont 6838 conducting composite. The geometric misfit of the contact areas on the opposite crystal faces did not exceed 5 μm. Upon the subsequent short-time (not exceeding 5 min) annealing of the conducting composite at 500–520°C in the oxygen flow, the contact resistance was 1–10 Ω at room temperature and exhibited a metal-type behavior with the temperature variation, with a (1.5–2.5)-fold drop on the sample cooling from room temperature to liquid nitrogen temperature. The crystal position relative to the quartz substrate was fixed with the aid of the current and potential leads made of a 5-μm-thick gold wire. This method ensured a quite satisfactory fixation of the samples during measurements in the pulsed magnetic field: each sample admitted 200 to 500 pulses (about quarter of which exceeded 50 T in magnitude) without significant changes in the contact resistance or any sign of mechanical displacement.

In the preliminary stage, all crystals were characterized by the temperature dependence of the out-of-plane resistance in a zero magnetic field. These potentiometric measurements were conducted using a standard ac bridge technique at 25–77 Hz; the heating/cooling rate was selected so as to ensure that the maximum temperature difference between the control gauge and the sample did not exceed 10 mK. The crystals selected for experiments in the pulsed fields exhibited coinciding $R(T)$ curves measured upon commutation of the contacts; the character of the temperature dependence and the values of the out-of-plane resistance ($\rho_c(100\ \text{K}) \approx 7$–$15\ \Omega$ cm) showed evidence of the absence of shunts. The temperature of the superconducting transition in a zero field (T_{c0}) was determined as the temperature corresponding to a resistance drop by four orders of magnitude as compared to the value at 100 K. The sample crystals possessed close T_{c0} values varying within 91–93.3 K. The temperature widths of the superconducting transitions were also slightly varying, not exceeding 1 K. At the same time, possessing qualitatively similar shapes, the $R(T)$ curves of various crystals exhibited a quantitative scatter within the limits indicated by solid curves in Fig. 1.

2.2. Method of the Crystal Resistance Measurement in a Pulsed Magnetic Field

The resistance measurements were performed for the most part on the setup for low-temperature investigations of the de Haas–van Alphen effect created at the Bristol University (United Kingdom), specially adapted for the potentiometric measurements in an extended temperature range (1.4 to 350 K). The field was generated by discharge of a high-voltage capacitor bank ($C \approx 12$ mF) via a liquid-nitrogen-cooled solenoid coil with $R_c(78\ \text{K}) \approx 20$–$25\ \text{m}\Omega$ and $L_c \approx 450$–500 μH. In order to increase the pulse duration (on reaching a maximum current in the solenoid), the coil outputs were shorted with an additional diode section. Therefore, it was expected a priori that the field on the leading (rise) and trailing (fall) fronts of the pulse would vary according to the sinusoidal $B = B_{\max}\sin(t/\sqrt{LC})$ and

[2] This method eliminated the problem of the sample surface degradation unavoidably encountered in the conventional method, whereby the doping level is adjusted to compensate for a change in the oxygen concentration as a result of prolonged high-temperature annealing of the crystal.

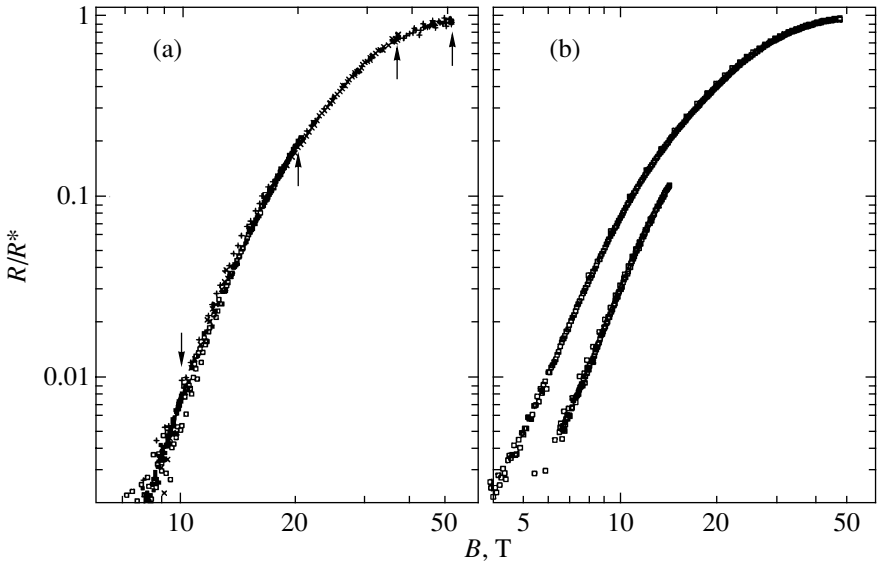

Fig. 2. Experimental verification of the effect of induction overheating at $T \approx 25$ K on the resistance of BSCCO-2212 crystals of different area: (a) $S \approx 53 \times 42$ μm^2, $B_{max} \approx 10$, 20, 36, and 50 T (indicated by arrows); (b) $S \approx 110 \times 85$ μm^2, $B_{max} \approx 14$ and 46 T (R^* is the sample resistance at maximum of the curves).

exponential $B = B_{max}\exp(-R_c t/L_c)$ laws, where B_{max} is the field amplitude. The real pulse shape somewhat differed from that anticipated: the field variation on the trailing front deviated from an exponential, and the total pulse duration (on an $0.1B_{max}$ level) depended on B_{max} in the zero approximation as $A\exp(-B_{max}/B_0)$ with $B_0 \approx 29$ T and a preexponential factor A of about 51 ms. At the same time, the B_{max} value produced a negligibly small effect on the shape and duration of both the leading front (rise time, 3–3.5 ms) and the top plateau, where the field was approximately constant (to within 0.02%) over about 80 μs. All these deviations were probably related to the Joule heating of the coil by current pulses, which slightly change the inductance on the background of a strong increase in the resistance. In order to provide for a correct solution of this problem, the shape of each $B(t)$ pulse was experimentally determined by integrating the signal taken from a probing coil; the absolute calibration was performed using the results of auxiliary measurements of the quantum oscillations of magnetization in a reference gold single crystal.

A high rate of the field variation during the pulse, together with unavoidably induced spurious signals and mechanical vibrations, poses additional requirements on the methods of measurements, setup design, and sample dimensions. In the experiments, a directly induced spurious signal in the measuring circuit was eliminated in two steps. First, the spurious emf induced by a field pulse was partly compensated by the signal from a probing coil placed near the sample; finally, the contributions were separated and the useful signal (with a working frequency) was determined by a special numerical processing of the data array. Obviously, the induction currents are generated not only in the measuring circuit, but in the sample as well, thus leading to an additional heat dissipation in the sampler volume and its overeating relative to the thermostat. In a rough approximation, the overheating is proportional to $(S\partial B/\partial t)^2/\rho$, where $\partial B/\partial t$ is the rate of the field variation, S is the sample cross section area perpendicular to the field, and ρ is the crystal resistivity in this plane. This circumstance shows the need in using small samples. At the same time, a quadratic dependence of the heating effect on $\partial B/\partial t$ can be used both for experimentally evaluating the degree of overheating in each particular experiment [28] and for estimating the admissible sample dimensions [24]. In the experiments reported here, the induction heating of a crystal was systematically monitored by comparing the $R(B)$ curves measured at each temperature using a series of pulses with various amplitudes differing by a factor of 3–20.

In a preliminary methodological stage, a series of BSCCO-2212 crystals of various dimensions were studied. These experiments revealed a strong increase in the effect with decreasing temperature of the sample and showed that the sample overheating induced by a 55-T pulse at $T \geq 10$–15 K was below the experimental detection threshold for a crystal area reduced to approximately 2×10^3 μm^2 (Fig. 2). These results determined the upper admissible area of the sample, thus posing restrictions on the dimensions of crystals selected for subsequent investigations. It should be noted that, by systematically using the results of measurements on both (leading and trailing) fronts of the magnetic field pulses, it was experimentally established that even the absence of hysteresis on the curves does not ensure that the induction overheating is negligibly small. This conclusion is illustrated in Fig. 2b showing virtually hys-

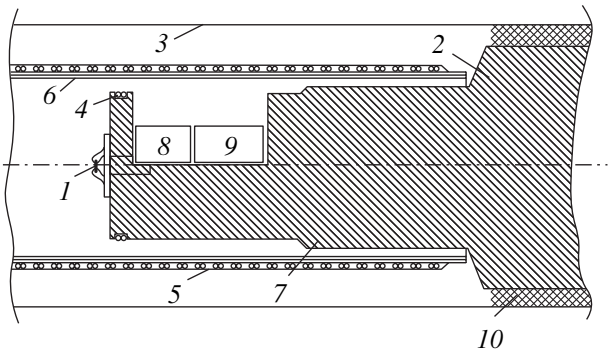

Fig. 3. A schematic diagram of the working part of the cryogenic insert: (*1*) crystal; (*2*) plastic insert; (*3*) inner surface of the Dewar vessel stem; (*4*) probing coil; (*5, 6*) heater winding and frame, respectively; (*7*) heater-fit part of inset; (*8, 9*) temperature sensors; (*10*) cotton tissue winding.

teresis-free plots with evident manifestations of the overheating.

The sample resistance was measured using a multistep procedure. Immediately during the magnetic field pulse, the kinetics of the magnetic field variation rate and the total voltage drop across the sample were monitored, and the resulting dependence of the sample resistance on the field was determined from these data by the numerical lock-in detection of the useful signal at a working frequency of the probing current.

The main measurements in the pulsed magnetic field were performed in the ac mode at a frequency selected according to the characteristic times of the field variation and the condition of minimum noise. The amplified signal from the potential contacts of a sample and a signal of the voltage drop in the probing coil were simultaneously processed by a high-speed (500 kHz) digital-to-analog converter with four independent channels. The useful signal $V(t)$ separated at a working frequency was determined at the stage of data processing upon numerically approximating elements of the experimental data array by a function of the type $V(t)\sin(\omega t + \phi) + A(t)$.[3] The resulting field dependence $V(B)$ was restored by combining these data with the magnetic field pulse profile $B(t)$. The method was developed using the results of a series of preliminary measurements employing calibrated RuO_2-based film microtransducers (Dale RSWP-575-40).

The experiments were conducted in an all-glass cryostat, the finger of which occurred on the axis of a solenoid coil through which the capacitor bank was discharged to generate the magnetic field. A sample crystal *1* (see Fig. 3) was mounted on the edge of a plastic (Melinex) cryogenic insert *2* placed into Dewar vessel stem *3*, so that the sample was located at a point of maximum field (to within 0.5 mm, which was ensured by measuring a signal from probing coil *4* used as the position sensor). The estimated accuracy of the sample orientation with the *ab* plane perpendicular to the field was not worse than 5°. The electric leads in the cryostat were made of bifilar manganin wires tightly set inside longitudinal channels sealed with an epoxy compound. In designing and assembling the cryogenic insert, special attention was paid to minimizing the area of wire loops. Heating coil *5* was also wound by a bifilar manganin wire onto a multilayer frame *6* made of insulated copper foil, tightly fit to flange *7* of the cryogenic insert. As a result, sample *1* and temperature sensors *8* and *9* were inside an additional shell providing for mechanical protection and temperature leveling. A tight fit of the plastic insert to the Dewar vessel stem *3* was ensured by a cotton-tissue winding *10*, which fixed the sample position relative to the solenoid, decreased the amplitude of vibrations, and provided for a thermal insulation making the measurements possible in a temperature range from 1.4 to 350 K in a cryostat filled with liquid helium.

The temperature was stabilized and monitored with the aid of two resistance detectors (copper and Cernox, with positive and negative temperature resistance coefficients, respectively). The temperature sensors were calibrated using a reference metrological thermometer (Lake Shore) mounted in place of a sample. The active temperature stabilization was provided by a standard LTC-21 controller (Neocera).[4] The temperature was stabilized prior to each field pulse application, at a level of not worse than 10^{-4}; the temperature variation during the pulse was not monitored, and the conclusion about a negligibly small instability is made taking into account a strong temperature dependence of the crystal resistance R, the absence of hysteresis on the $V(B)$ curves, and the results of checks for the induction heating.[5]

[3] We used a modification of the standard software intended for an analysis of the de Haas–van Alphen oscillations. As a rule, the size of the sinusoidal fitting interval was selected so that a change in the induction signal was approximated by a quadratic polynomial. It was empirically established that five periods of the measuring frequency ensure a satisfactory dynamic range at an acceptable rms deviation (about $10^{-3}R^*$, where R^* is the resistance at maximum of the curve, see Fig. 2). A relatively weak sensitivity of the method with respect to the fitting interval size was confirmed by perfect agreement between the $R(B)$ curves determined upon fitting over 1–12 periods. Coincidence of the results of analysis using significantly different windows was used in a regular manner for controlling the correctness of measurements.

[4] This temperature controller model was chosen because its processor is inherently suited to work with thermometers of any type, requiring no special interfaces. It was found *a posteriori* that the temperature stabilization in the experimental range studied requires a minimum adjustment of the control parameters.

[5] The statement concerning coincidence of the experimental curves is obviously valid to within the experimental uncertainty. In these experiments, the upper limit of the possible temperature variation during a field pulse was estimated from a maximum scatter of points on the $R_N(T)$ curve (see subsection 3.1). In fact, the scatter of experimental data obtained with various samples indicates that the temperature was constant with a relative accuracy of 5×10^{-3}–5×10^{-4}.

The experimental data presented below were obtained using ac resistance measurements at a frequency of 77.7 kHz; the working current densities used for various samples slightly varied within the interval $j = 1–5$ mA/cm^2. Note that the lower boundary of the current densities used in analogous measurements usually exceeds the values reported here by at least one order of magnitude [29, 30]. The amplitude of the probing alternating current was selected within the ohmic response of the crystals studied, which is confirmed by applicability of the sinusoidal approximation to the "raw" measured data. Linearity of the response was additionally verified by the $R(B)$ curves measured at various fixed temperatures using four or five current amplitudes differing by one to one and half orders (sometimes up to two orders) of magnitude. The current–voltage characteristics constructed by these data could be approximated by a power law $V \propto I^\gamma$ with $\gamma = 1.01 \pm 0.01$ for the fields within $0.1\ \mathrm{T} < B \le 55\ \mathrm{T}$. Finally, the conclusion concerning the ohmic response of the crystals in the resistive state was confirmed by direct measurements of the out-of-plane current–voltage characteristics in a constant magnetic field [31].

In order to compare the obtained data to the results of low-frequency measurements and to evaluate the frequency dependence of the resistance for each sample, a series of control measurements of the $R(B)$ curves were performed at several temperatures with the probing current frequency varied from 7 to 111 kHz. The observed coincidence of these curves (to within random scatter of the experimental points in one experiment) and a quantitative correlation of the results of dc and pulsed (low-field limit) measurements were indicative of insignificant frequency dependence of the resistance.

3. EXPERIMENTAL RESULTS

Figure 4 shows a typical dependence of the interlayer ohmic resistance of a BSCCO-2212 crystal on the external field. As can be seen from these data, the magnetoresistance below T_c contains contributions with different signs, reflecting the influence of various physical mechanisms. In a strong field, the system is characterized by a quasilinear negative magnetoresistance, while the region of weak fields shows an initial power increase in the resistance followed by a linear portion. The position of the maximum (B^*) of this curve, determined by the competition of two contributions, and the amplitude (R^*) of this maximum strongly increase with deviation from the critical temperature T_c and, on the contrary, decrease on approaching this point. A contribution related to the positive magnetoresistance component drops sharply on reaching the critical temperature of the crystal and completely disappears when the temperature increases 5–7 K above T_c. As the temperature grows further, the system exhibits a negative magnetoresistance in the entire range of magnetic fields.

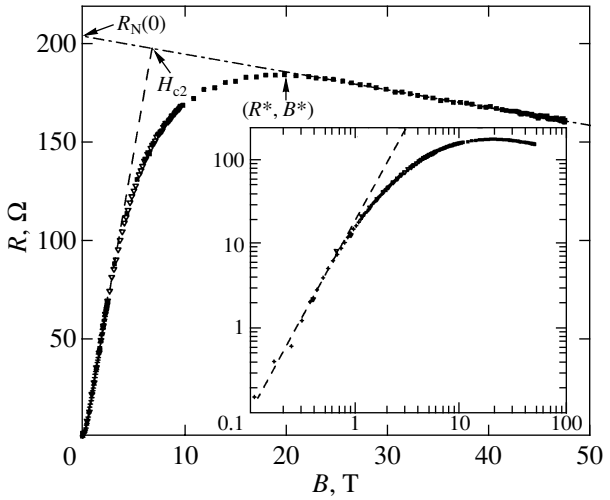

Fig. 4. Typical dependences of the out-of-plane resistance of a BSCCO-2212 crystals on the magnetic field at $T < T_{c0}$. The measurements were performed at $T = 56.6$ K using three pulses with the amplitudes $B_{max} \approx 2.7$, 9.7, and 47 T. The dash-dot and dash lines show approximations of the magnetoresistance under the conditions of field-suppressed superconductivity and flux flow, respectively; $R_N(0)$ and H_{c2} are estimates of the crystal resistance in the absence of superconductivity and the upper critical fields, respectively, obtained by extrapolation. The inset shows the same data plotted against a double logarithmic scale; here, the dashed line shows approximation of the initial (low-field) region by a power law $R(B) \propto B^\beta$.

Based on the results of these observations, it was natural to consider the positive magnetoresistance as a manifestation of the resistive state of the crystal, while the contribution of the opposite sign was attributed to properties of the normal state. It will be demonstrated in the subsequent subsections that (i) within the framework of these assumptions, one can separate strongly temperature dependent contributions to the resistance related to the normal and resistive state, and (ii) additionally assuming that the field dependence of the magnetoresistance in the normal state remains functionally unchanged, it is possible to evaluate the temperature variation of the upper critical field in BSCCO-2212 crystals and their resistance in the absence of superconductivity. In the following sections, a self-consistent physical mechanism is described which is responsible for the functionally different asymptotic behaviors observed in the experiment.

3.1. Longitudinal Magnetoresistance in the Normal State

The principal purpose of this study was to evaluate the upper critical field assuming that the formation of superconducting nuclei leads to deviation of the mag-

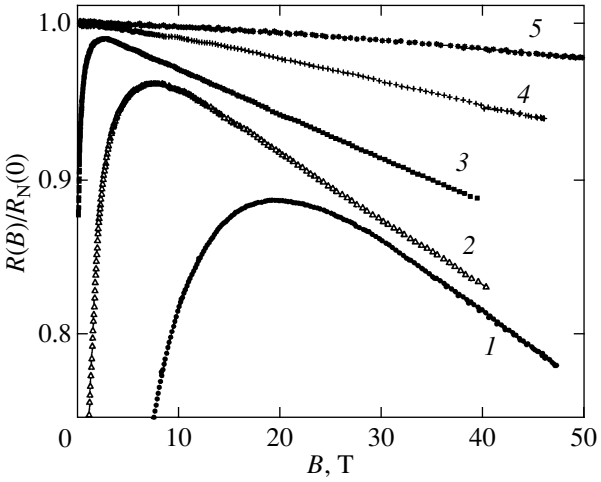

Fig. 5. Evolution of the normalized experimental plots of $R(B)/R_N(0)$ constructed by the data measured both above and below T_{c0}. Curves *1–5* refer to T = 56.6, 78.0, 90.05, 101.6, and 125 K, respectively.

netoresistance from a behavior typical of the normal state. It was suggested that the region of superconducting fluctuations is anomalously large; therefore, the characteristics of the normal state were studied in an extended temperature range. For each crystal, the field dependence of the out-of-plane resistance was measured in the temperature interval of 300–350 K and below; in all these experiments, the field was oriented perpendicularly to the basal plane of the crystal and parallel to the direction of the probing current (to within an accuracy estimated as 5°).

According to [23, 24], it was established that the longitudinal magnetoresistance of the crystal studied is negative and, to a first approximation, is satisfactorily approximated by a linear relationship. A special experiment showed that the effect is even with respect to reversal of the external field direction. The scatter of experimental data left a certain freedom in description of the asymptotic behavior of the magnetoresistance as $B \longrightarrow 0$, admitting both linear and a quadratic approximation [32].[6] As the temperature decreased, the magnetoresistance increased more rapidly than did the zero-field out-of-plane resistance $R_N(0)$. The normalized slope $S = -R_N^{-1}(0)\partial R/\partial B$ revealed an approximately 500-fold growth upon decreasing the temperature from 200 to 87 K, which can be satisfactorily described by the relation $S \propto \exp(T_0/T)$ with $T_0 \approx 820 \pm 70$ K.

[6] Previously [32], the effect was reported to exhibit a quadratic growth in a quasistationary field of up to 14 T, but the direct comparison or results is hindered both by a significant difference in the levels of crystal doping and by the possibility that the dependence observed in [32] was distorted by a systematic temperature drift in the course of the field sweep.

It should be emphasized that, albeit the character of temperature variation of the resistance was qualitatively the same for all samples, there were significant differences in the absolute values of characteristics for various crystals. Indeed, the normalized slope S exhibited an almost fivefold scatter, while the estimated relative zero-field resistance growth $\Delta R_N(0)/R_N(0, 300\ \text{K})$ showed a threefold change when the temperature decreased from 300 to 50 K. These variations probably reflect individual features of the microstructure of samples. Taking into account that all crystals had very close values of the critical temperature ($T_{c0} \approx 91$–93 K), the observed experimental scatter is additional evidence for the correctness of attributing the quasilinear negative magnetoresistance to properties of the normal state.

As was mentioned above, the character of the asymptotic behavior of the longitudinal magnetoresistance for the temperature approaching the critical temperature from above was retained in strong magnetic fields, while the low-field portion of the data showed deviations from a simple linear dependence. The deviations increased upon decreasing the temperature, and, at $T - T_{c0} \leq 2$–2.5 K, a region of the field appeared in which the initial portion of the field dependence exhibited a positive slope and the $R(B)$ curve acquired a maximum R^*. As the temperature decreased below the critical temperature, the region of a positive magnetoresistance expanded and the interval of fields appeared in which no resistive response at all was observed. A typical $R(B)$ plot for this temperature interval is presented in the inset to Fig. 4.

The character of transformation of the experimental curves observed on going through the critical temperature region is illustrated in Fig. 5 showing the $R(B)$ plots measured at several temperatures above and below T_c of the crystal studied. A comparison of these data unambiguously shows that the deviations from linearity observed in weak magnetic fields are related to the resistive state of the sample. The fact that the character of behavior in the high fields remains unchanged give grounds to believe that an approximate description of the field dependence of the magnetoresistance determined at high temperatures can also be used for estimating $R_N(0, T)$ from the results of $R(B)$ measurements below T_c. According to this approach, the experimental data were treated assuming that the region with a negative slope is related to the longitudinal magnetoresistance $R_N(B, T)$ under conditions of field-suppressed superconductivity, and $R_N(0, T)$ was determined by the linear extrapolation of this branch to $B = 0$. A typical result of such data treatment is presented in Fig. 6 together with the temperature dependence of the interlayer resistance of the same crystal measured on a direct current in a zero magnetic field. A satisfactory agreement between the data obtained by two methods at $T > T_c$ indicates, in particular, a weak effect of the

probing current frequency change on the results of the out-of-plane resistance measurements. According to these data, a strong increase in the interlayer resistance of BSCCO-2212 observed with decreasing temperature in the region above T_c continues at temperatures well below T_c, not showing pronounced tendency to weakening.

Unfortunately, the range of fields employed in this study was insufficient to trace the $R_N(0, T)$ dependence down to very low temperatures. Thus, the asymptotic behavior of the magnetoresistance as $T \longrightarrow 0$ remains undetermined, insistently requiring additional investigation. In fact, the lower temperature boundary of the region of validity of the extrapolated estimate obtained for $R_N(0, T)$ is determined by the character of evolution of the $R(B)$ curves. According to Fig. 5, as the temperature decreases, the (B^*, R^*) maximum exhibits a certain broadening and shifts toward higher fields, thus reducing the experimentally accessible region of a linear negative magnetoresistance and posing a natural restriction from below on the temperature interval in which a correct estimate of $R_N(0, T)$ can be obtained. As can be seen from Fig. 6, the (B^*, R^*) maximum as such can be reliably detected to much lower temperatures. Although the amplitude and position of this maximum are determined by the interplay of contributions due to resistive and normal states, the absence of pronounced features in the $R^*(T)$ behavior (as well as the lack of essential grounds for such features to exist) admits using these data for estimating the $R_N(0, T)$ values at low temperatures. The dashed line in Fig. 6 shows the extrapolation $R^*(T) \longrightarrow R_N(0, T)$ obtained assuming the validity of the empirical relationship $R_N(0, T) - R^*(T) \propto \exp(-\alpha T)$, which satisfactorily described the results in a broad range of temperatures where both $R_N(0, T)$ and $R^*(T)$ values are reliably determined. As seen in Fig. 6, the extrapolation converts the saturated curve of $R^*(T)$ into $R_N(0, T)$, retaining the tendency to rapid growth with decreasing temperature. The obtained results agree with a presumably different origin of the R_N and R^* quantities. Nevertheless, validity of such an extrapolation requires experimental verification.

Unfortunately, data on the behavior of $R_N(0, T)$ available in the literature are rather restricted. For this reason, the obtained relationship can be compared to the results of investigation of the magnetic field effect on $R(T)$ and with the experimental dependence of resistance of the BSCCO-2212 crystal in a zero field under the conditions of superconductivity broken by strong current pulses of microsecond duration [33]. It was found that the character of the $R_N(0, T)$ curve qualitatively agrees with the shape of envelope of the family of $R(T)$ curves measured in various fields [29]. The latter envelope was originally attributed to a quasiparticle tunneling via a Josephson medium [29], but the fact that this shape was retained at temperatures well above the critical temperature showed this interpretation to be inadequate, after which the notions were revised in favor of normal resistance [22]. In addition, the data presented in Fig. 7 show that, while the $R_N(0, T)$ curve satisfactory agrees with the results obtained by

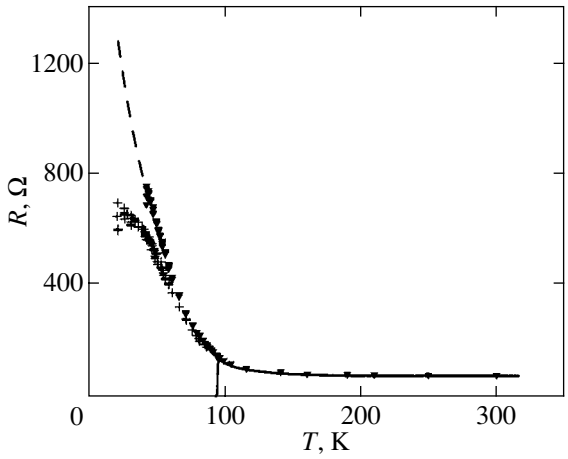

Fig. 6. Typical results of the experimental data processed as described in the text: triangles denote the values of resistance $R_N(0, T)$ measured in the absence of superconductivity; solid curve represents the $R(0, T)$ obtained by dc measurements; crosses show the values of resistance at maximum $R^*(T)$; and the dashed line shows the $R_N(0, T)$ curve empirically reconstructed from $R^*(T)$.

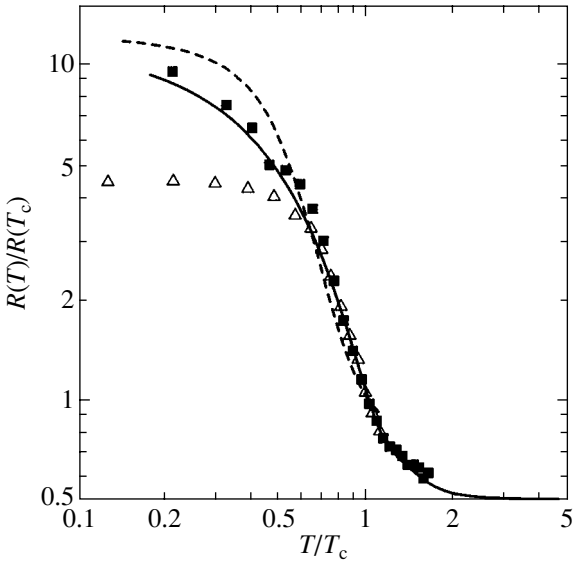

Fig. 7. A comparison of the experimentally determined $R_N(0, T)$ values (solid curve) to the published data. Dashed curve shows the resistance of a crystal sample measured in a field of 2 T at high current density (50 A/cm^2) [36]; triangles represent the $R_N(T)$ values measured using a series of microsecond current pulses in the absence of applied field [33]; squares represent the R_{qp} values determined for the return branch of the current–voltage characteristic in a zero field [35].

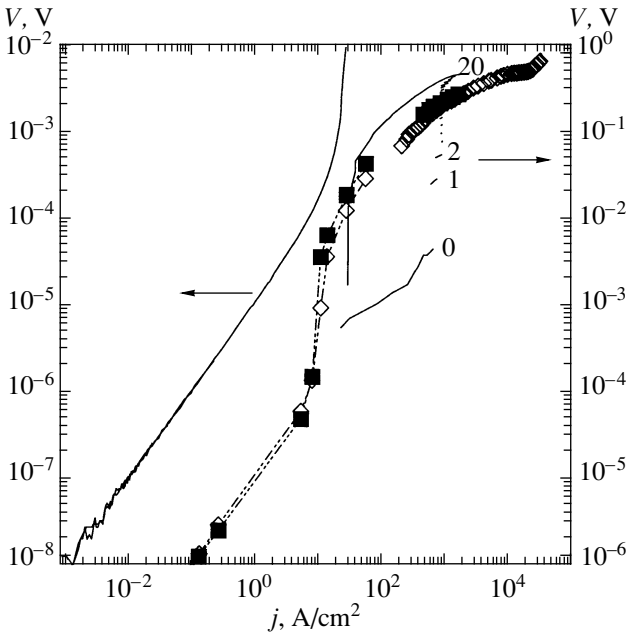

Fig. 8. The current–voltage characteristic (right-hand voltage scale) of a BSCCO-2212 crystal reconstructed from the data reported in [36]: symbols connected by dashed curve correspond to the values determined for 30 K (■) and 40 K (◇) from the $R(T)$ curves for $H = 2$ T. Solid curves with indexes at the first three direct (0, 1, 2) and return (20) branches represent a multibranch current–voltage characteristic measured in the regime of the internal Josephson effect in a zero field at 5.6 K. Nonconnected symbols represent a current–voltage characteristic measured using a series of current pulses for $H = 1$ T at 30 K (■) and 40 K (◇). Solid curve referred to the left-hand voltage scale shows the results of thorough measurements [31] of the current–voltage characteristic in the region of weak currents for $H = 3$ T at 35 K.

Yasuda *et al.* [33] for $T/T_c > 0.55$, a significantly different behavior is observed on further decrease in the temperature. Nevertheless, the discrepancy seems not to be as important, since the saturation observed in [33] was attributed by the authors to a possible overheating of the sample.

It should also be noted that the results obtained in this study confirm an interesting observation made previously [34] in the course of investigation of the internal Josephson effect in BSCCO-2212 crystals. In that study, a quantitative correlation was observed between the resistance R_N (determined under the conditions when superconductivity was suppressed by a strong magnetic field) and a "quasiparticle" resistance $R_{qp} \equiv (U/I)_{I \to 0}$ (estimated for the return branch of the current–voltage characteristic at the same temperature but in a zero field). As can be seen from Fig. 7, the curve of $R_N(0, T)$ is in fact correlated with R_{qp} from [35] (where the measurements analogous to those originally performed in [34] were reproduced in a wider temperature range). As will be demonstrated below, an analysis of the whole body of recent results gives grounds for a new interpretation of the original results reported in [34].

Until recently, a comparison between the results of resistance measurements and the aforementioned current–voltage characteristics was hindered by different scales of the current densities: the maximum current densities used in the former measurements rarely reached even 1 A/cm^2, whereas minimum current densities involved in the measurement of the internal Josephson effect were greater by one to two orders of magnitude. The gap was filled to a certain extent by Suzuki *et al.* in [36], where determination of the standard current–voltage characteristic was supplemented by measurements of the sample resistance at various current densities in the range from 0.1 to 50 A/cm^2. Figure 8 plots together the current–voltage characteristic constructed by using the $R(T)$ data from [36], a characteristic of the internal Josephson effect, and the results of pulsed measurements performed by the same researchers. Despite nonideal matching of these data, which is related to different experimental conditions, Fig. 8 provides for an insight into the true character of variation of the current–voltage characteristic in a wide range of currents.

An important fact, which can hardly be explained proceeding from the Josephson effect considerations, is the presence of a linear initial portion on the current–voltage characteristic. This fact was confirmed by direct measurements of the current–voltage characteristics [31] which are also presented in Fig. 8. As can be seen from this figure, a specific feature of the development of nonlinearity in the current–voltage characteristic is a strong growth in the voltage observed within a narrow region of currents, followed by a less pronounced increase and a tendency to still weaker variation with increasing distance from this region. This behavior can be explained taking into account a characteristic scale of currents, poor thermal conductivity, high contact resistance, and negative temperature coefficient of the normal resistance of BSCCO-2212. As is known, the main problem complicating the measurements of the current–voltage characteristics is the uncontrolled Joule heat evolving in the sample, the amount of which is determined by the thermal flux density W/S through the crystal surface. For a convenient analysis, Fig. 9 presents the same data from [36] plotted as resistance $R = V/I$ versus $\Delta T \propto W/S = VI/S$ (with neglect of the heat dissipated in the contacts). Similarity of the descending parts of this plot and the curve representing the results of direct resistance measurements is indicative of the "overheating" nature of the effect observed for the thermal flux densities exceeding 10 W/cm^2. Additional evidence for the adequacy of this interpretation is offered by experimental data [37] revealing quantitative coincidence of the current–voltage characteristics of the samples possessing different areas, represented in terms of the current density. Finally, a significant overheating in the near-contact

region of the sample at $W/S \approx 50$ W/cm^2 was revealed in the course of direct measurements [38] performed in the regime of the internal Josephson effect at 24 K. Within the framework of the "overheating" concept, it is possible to explain a bell-like shape of the return branch of the current–voltage characteristic (curve 20 in Fig. 9), the transition to which takes place at a thermal flux density exceeding 1000 W/cm^2. Since the overheating drops when the current decreases, the right-hand part of the characteristic reflects the temperature dependence of resistance in the normal state of the sample. A maximum in this dependence is apparently related to thermalization of the crystal and the electron subsystem, because the region of positive slope (usually omitted in published papers) evidently follows a nonlinear part of the current–voltage characteristic under weakly nonequilibrium conditions.

Based on the above considerations, the similarity of presumably dissimilar dependences presented in Fig. 7 is naturally explained by the fact that, in the zero-order approximation, the quasiparticle resistance R_{qp} corresponds to nonequilibrium conditions of the superconductivity breakage by the transport current and is implicitly related to the R_N value. Obviously, additional experimental and theoretical investigations are required, in particular, for refining the proposed relationship between R_{qp} and R_N. At the same time, the results of comparison of the current–voltage characteristics in Fig. 7 indicate that $R_{qp}(T)$ can be used as a rough estimate of $R_N(T)$ when direct measurements of the latter value are inaccessible.

3.2. Temperature-Activated Flux Flow

As was mentioned in the Introduction, it is a widespread conviction that a characteristic field determined from the resistance measurements reflects certain features in the behavior of the vortex system, but does not correspond to the upper critical field. The features of vortex dynamics in BSCCO-2212 depending on various possible factors were thoroughly studied both theoretically and experimentally [1]. It was established that, at sufficiently low temperatures, the compound occurs in a more or less ordered state, crystal or glass, depending on the degree of perfection. A significant difference of the HTSC cuprates from conventional superconductors consists in that there are wide regions of fields and temperatures featuring the state of flux flow with ohmic response. This region is separated from the region of ordered states by the so-called irreversibility line. Additional lines, occurring between $H_{irr}(T)$ and $H_{c1}(T)$, reflect fine features of the ordered state of the vortex system [1, 3]. On the contrary, the vortex fluid contains (as does any fluid) no structural features. For this reason, a comparison of the irreversibility field and the "resistive" critical field determined in the same experiment is important for elucidating that

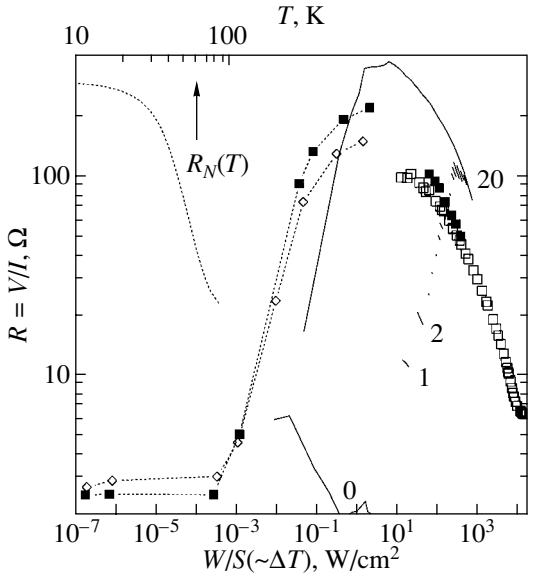

Fig. 9. The data from [36] (same as in Fig. 8) replotted in the coordinates of $R = V/I$ versus $\Delta T \propto W/S \approx VI/S$ and supplemented by the experimental curve of $R(T)$ for the same sample measured in a field of 2 T at a large current density of 50 A/cm^2 (top temperature axis). Other notation same as in Fig. 8.

there is much in common in the origin of these characteristics.

Thus, the problem to which this investigation was devoted required experimentally determining the irreversibility field of the crystal studied, which implied a thorough study of the character and origin of the dissipative response. An additional argument for such a study is the need for experimentally separating the contributions from various factors to the measured value. Indeed, although the character of transformation of the $R(B)$ value on going via the critical temperature region allows the region of positive magnetoresistance to be unambiguously related to the superconducting state, the experimental curves observed in this region exhibit a complicated behavior, apparently reflecting a change in the regime of charge carrier transfer. The results considered in this subsection, while providing an estimate for $H_{irr}(T)$, allow the initial part of the region under consideration to be related to the thermally activated vortex motion and significantly reduce the class of theoretical models applicable to description of the vortex state in BSCCO-2212.

The character of development of the resistive state in the crystal studied is illustrated in the inset to Fig. 4, where the curves from the main panel are replotted on a double logarithmic scale. As can be seen, the initial portion of this dependence (extending over a broad range of resistances up to $R/R^* \leq 10^{-1}$) admits approximation by a simple power law $R \propto B^\beta$ with a temperature-dependent exponent γ (see the main panel of Fig. 10). It was experimentally established that this

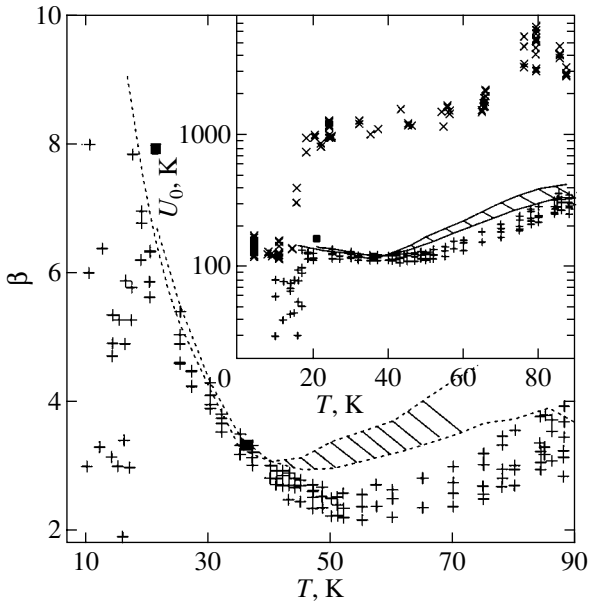

Fig. 10. Temperature variation of the approximation exponent in the relationship $R \propto B^{\beta}$: vertical crosses represent the results of determination in the pulsed field; cross-hatched region represents the dc measurements [42]; squares refer to the β values determined from dc current–voltage characteristics (Figs. 8 and 11). The inset shows the barrier $U_0 = \beta T$ versus temperature plot constructed by data from the main panel, supplemented by the values $U_0 = -T\partial \ln t/\partial M_{\text{rem}}$ determined from the measurements of logarithmic relaxation of the magnetic moment M_{rem} of the crystal (oblique crosses).

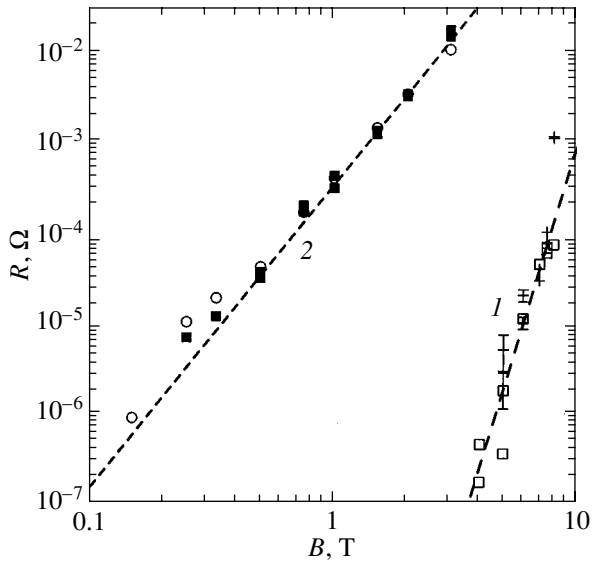

Fig. 11. Effect of the magnetic field on the linear resistance of BSCCO-2212 crystals determined from the ohmic portion of the current–voltage characteristic at (*1*) 21 K and (*2*) 36 K. Dashed lines show the results of approximation using the $R \propto H^{\beta}$ relation with $\gamma = 7$ (*1*) and 3.4 (*2*).

behavior is observed in the entire temperature range ($0.15 \le T/T_c \le 1$) of the resistive state of the crystal studied.[7] Taking into account that the temperature dependence of resistance in the same interval exhibits a temperature-activated character, $R(B, T) \propto \exp[-U(B)/T]$, the above effect is naturally identified with the temperature-activated flux flow regime. The development of this regime implies the appearance of an ohmic resistance. Therefore, in order to check for the correctness, it is necessary to study the character of the system response under these conditions.

The specific features of investigations in the pulsed field hinder direct measurement of the current–voltage characteristic. However, the absence of detectable deviations of the signal shape from sinusoidal and the coincidence of the results of resistance measurements using the current densities differing by one order of magnitude are indicative of linearity of the system response under these conditions. Additional evidence for this is provided by the results of direct measurement of the current–voltage characteristic in a steady field with a strength below 8 T, where the characteristic was linear in the range of current variation over several orders of magnitude (see Fig. 8). Note that the ohmic resistance determined from the initial part of these current–voltage characteristics (Fig. 11) is also satisfactorily approximated by the power law, in a quantitative agreement with the results of measurements in the pulsed field.

The experimentally confirmed linearity of the system response, together with the temperature-activated character of the $R(T)$ curve, is explicitly indicative of the fact that this part of the temperature dependence of the resistance reflects the regime of temperature-activated flux flow in the crystal studied. Moreover, the obtained results allow one to significantly reduce the class of theoretical models applicable to description of the vortex system in BSCCO-2212 crystals, since a power law describing the field dependence of the resistance is evidence for validity of the models developed in [39, 40], which predict a logarithmic dependence of the barrier height for this process: $U = U_0 \ln(H_0/B)$ and, accordingly, $R \propto \exp(-U/T) \propto B^{U_0/T}$. Additional evidence in favor of this interpretation is offered by Fig. 12 showing an experimental plot of $U(B)$—the barrier height versus applied field—which, according to the results of these measurements admits a logarithmic approximation in the region of strong fields with the parameters $U_0 \approx 100 \pm 20$ K and $H_0 = 200 \pm 30$ T, which are close to the values reported by Skvortsov and Geshkenbein [41]. At the same time, the obtained results demonstrate (in contrast to [41]) a significant increase

[7] It should be noted that the range of fields used in these experiments was insufficient to detect the resistive state at $T \le 10$–12 K. This fact provides for an estimate from below for $H_{\text{irr}}(4.2\text{ K})$ and presents additional evidence for an essentially "overheating" origin of the internal Josephson effect parameters measured for the most part in a zero field at 4 K.

in the barrier height in the region of weak fields, nevertheless also admitting a logarithmic approximation but with a different set of parameters: $U_0 \approx 1500 \pm 500$ K and $H_0 = 0.4$–0.7 T.

The U_0 values, determined within the framework of this approach using the $R(B) \propto B^{U_0/T}$ approximation of the field dependence of the out-of-plane resistance measured at various temperatures are shown in the inset to Fig. 10. As can be seen, there is a satisfactory quantitative agreement of the results of high-frequency measurements in a pulsed magnetic field and the data obtained from quasistationary measurements in a constant field [42]. To the first approximation, the "resistive" value of U_0 is virtually constant in a wide temperature interval from 20 to 60–70 K and agrees with the estimate obtained above from the logarithmic approximation of the temperature-activated process barrier. At the same time, the values outside this interval exhibit a systematic growth with the temperature increasing above 60–70 K and a tendency to sharp drop with the temperature decreasing below 16–18 K.

A comparison of the results presented in Fig. 10 (main panel) to the data on the crystal magnetic moment relaxation measured using a SQUID magnetometer [43] (presented in the inset to Fig. 10) reveals a qualitatively similar behavior. This is a remarkable fact, the more so that we compare the characteristics of the resistive and critical states occurring on different sides of the irreversibility line representing, in the opinion of many researchers, a real phase boundary [44]. A pronounced correlation between characteristics of the resistive and critical states is a clear indication of a common mechanism involved in these phenomena. A quantitative discrepancy between the results of the magnetic relaxation and resistance measurements presented in the inset to Fig. 10 is probably related to differences in the experimental conditions. Taking into account that (i) the scale of fields involved in the resistance measurements is higher by one to two orders of magnitude than that used in the magnetic relaxation experiments and (ii) a characteristic time t_1 of the resistance measurements is several orders of magnitude smaller than the corresponding time of the magnetometric experiments [45], the systematically greater values of the "magnetic" barrier can probably be related [45] to a logarithmic dependence of the barrier on B and t_1 values: $U \propto \ln(t_1/t_0)$, where t_0 is the characteristic time of the system studied.

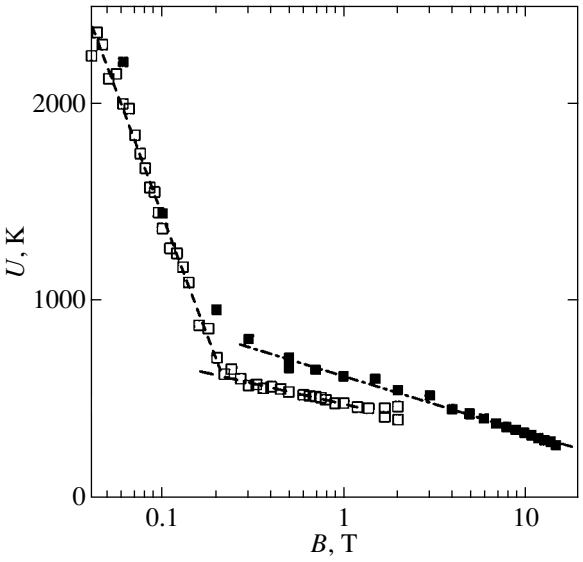

Fig. 12. The plot of barrier height U versus applied field for U values determined using the temperature activation approximation of the experimental $R(T)$ curves for two BSCCO-2212 crystals.

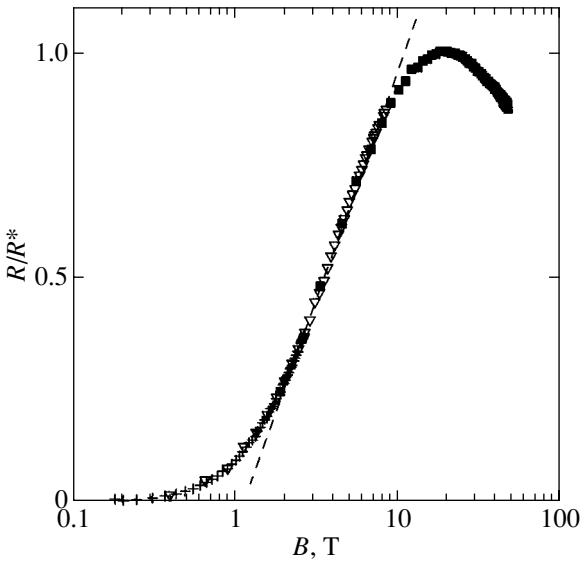

Fig. 13. A semilogarithmic plot of resistance versus field (data from Fig. 4) illustrating applicability of the $R \propto \ln B$ approximation (dashed line).

3.3. Flux Flow Region and the Upper Critical Field

Let us return to the field dependence of the magnetoresistance of a crystal in the resistive state (Fig. 4). Note that, as the field increases, the system exhibits deviation from the power asymptotic behavior at $R/R_N(0) \geq 0.05$–0.1. By virtue of the notions described in the previous subsection and by analogy with the case of "conventional" superconductors, this behavior is naturally related to a change in the flux flow regime in the crystal from temperature-activated to viscous. As can be seen from Figs. 4 and 13, the field dependence of the resistance in this region can be approximated both by a linear relationship $R \propto B$ and (in somewhat stronger fields) by a logarithmic law $R \propto \ln B$.

An analysis of the experimental data obtained in this region is complicated by the temperature dependence of the magnetoresistance under the conditions of field-

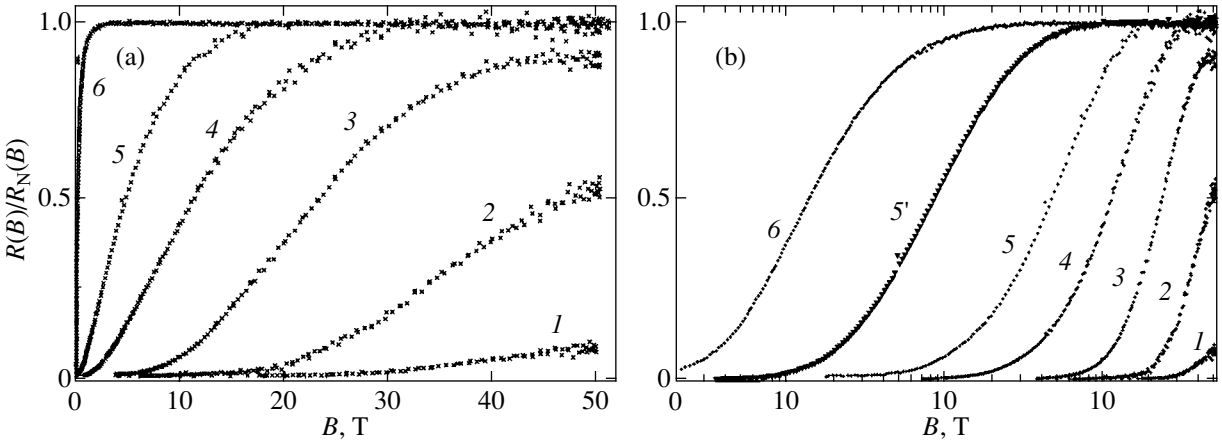

Fig. 14. Temperature evolution of the field dependence of the interlayer resistance of a BSCCO-2212 crystal in the mixed state: (a) normalized curves of $R(B)/R_N(B)$ measured at $T \approx 16$ (*1*), 20 (*2*), 30 (*3*), 45 (*4*), 57.5 (*5*), and 88.7 K (*6*); (b) the same data plotted on a semilogarithmic scale (curve *5*' refers to $T \approx 78$ K).

suppressed superconductivity. Taking into account the results of magnetoresistance measurements in the normal state, presented in Subsection 3.1, we can separate the temperature-dependent contributions from the normal and superconducting states. As a result, it was established that the effect of the temperature on the shape of the $R(B)$ curves describing the mixed state of BSCCO-2212 crystals significantly differs from that observed for typical low-temperature superconductors: instead of an almost parallel shift with decreasing temperature, the curves exhibit a decrease in the slope, while the initial point remains virtually undisplaced. This behavior is illustrated in Fig. 14a, where the results of $R(B)$ measurements at various temperatures are normalized to the corresponding field dependences of the magnetoresistance in the normal state. As is seen, the family of $R(B)$ curves exhibits a pattern different from that characteristic of a canonical superconductor in the magnetic field, although something like that is observed when the data are presented on a semilogarithmic scale (Fig. 14b).

The unusual character of evolution of the superconducting transition in Fig. 14 can be explained assuming that a linear positive magnetoresistance in the vortex fluid flow regime obeys the Bardeen–Stephen relationship $R_{FF} \propto R_N B/H_{c2}$. In this case, the character of evolution of the $R(B)$ curves measured at various temperatures is directly determined by the anomalous behavior of $H_{c2}(T)$. However, at first glance, applicability of this formula is not evident because the experiments were conducted under formally Lorentz force free configurations (the field parallel to the current). This formal discrepancy is removed by taking into account the layered structure of BSCCO-2212, according to which the vortex comprises a system of Abrikosov vortices in superconducting layers connected by the coreless Josephson segments. Since the positional correlation of vortices in the adjacent planes is readily violated (e.g., by thermal fluctuations), this leads to the appearance of a perpendicular component of the Josephson string perpendicular to the field and subject to the Lorentz force action. In the absence of theoretical calculations for these experimental conditions, these considerations give some grounds for using the Bardeen–Stephen formula, although the resistive response (in this experimental geometry) significantly depends on the character of interaction between the string and the core component of the vortex.

In any case, a strong temperature dependence of the slope of the linear part of the $R(B)$ curve is an experimental fact. A typical result is presented in Fig. 15, from which it is seen that the slope $\partial R_{FF}/\partial B$ exponentially increases with the temperature in the low-temperature region and then exhibits a much stronger dependence on approaching the critical temperature. The latter dependence obviously controls behavior of the "resistive" critical field, which is determined from the Bardeen–Stephen relationship as

$$H_{c2}(T) = \frac{R_N(0, T)}{\partial R_{FF}/\partial B}.$$

Indeed, as can be seen from Fig. 15, the slope varies over three orders of magnitude, while the normal resistance in the same temperature interval changes only by one order of magnitude (see Figs. 6 and 7).

In the course of a regular data processing, three approaches to determination of the characteristic field H_{c2}^* were employed (below, the notation H_{c2} is used irrespective of the method employed for determining the "resistive upper critical field). The first method is based on the Bardeen–Stephen formula in which the $R_N(0, T)$ was determined either by a linear extrapolation of the negative magnetoresistance to $B = 0$ under the conditions of field-suppressed superconductivity or by a resistance rescaling at the maximum of $R(B)$. Typical

results of the upper critical field evaluation for BSCCO-2212 by this method are presented in Fig. 16. As can be seen from these data, the resulting dependence differs from that predicted for the upper critical field within the framework of the BCS theory and shows a perfect quantitative agreement with the $H_{c2}(T)$ values obtained by the second method (the empirical determination of the upper critical field as a point of intersection of extrapolated linear approximations of the negative magnetoresistance in the normal state and the positive magnetoresistance in the flux flow region). Note that the second method differs only slightly from the conventional method, according to which the "resistive" critical field H_{c2} is taken equal to a value corresponding to a certain decrease in the sample resistance relative to that in the normal state. At the same time, the aforementioned quantitative coincidence of the results in the low-temperature region, where the extrapolation of the linear approximation of $R_{FF}(B)$ to $R = 0$ leads to a significantly nonzero field, is not a trivial result.

The third method used to evaluate the H_{c2} value in this study was based on the empirical approximation of the field dependence $R(B) \propto \ln B$, which also provides for a satisfactorily fit to experiment in the flux flow regime for the interval $0.2–0.3 \le R/R_N \le 0.7–0.8$. By analogy with the above construction, the H_{c2} value was determined by extrapolating this approximation to either (i) $R(B)/R^* = 1$ or (ii) $R(B)/R_N(0) = 1$. In the latter case, the $R_N(0)$ value at low temperatures was obtained by extrapolating the simplest empirical approximation $R_N(0) \propto \exp(-T/T_0)$. As is evident from the scheme of construction, the estimates of H_{c2} obtained in this way must quantitatively differ from those obtained by the first two methods. However, it was found that the differences can be eliminated by simple rescaling (using a temperature-independent coefficient close to two), after which the results obtained by the third method coincide (to within the experimental scatter of points) with the $H_{c2}(T)$ curve constructed by the first two methods.

This result can be considered as directly indicating that several functionally different empirical approximations of the experimental data obtained in the flux flow region are apparently related to the same physical mechanism. Unfortunately, to the author's knowledge, no theory has been developed so far that could be used for interpretation of the experimental data reported here. Therefore, selecting a reasonable scenario will require additional theoretical calculations. The available experimental data are insufficient for unambiguously identifying the characteristic field determined in this study with the thermodynamic value of H_{c2}. However, a qualitative agreement between the shapes of the temperature dependences $H_{c2}(T)$ determined by various methods, together with a quantitative coincidence of the results obtained by the same method for various crystals with close T_c values but significantly different characteristics of the normal state, indicates that the

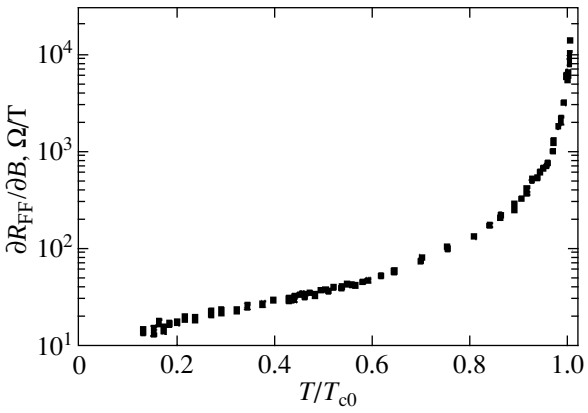

Fig. 15. A typical plot of $\partial R_{FF}/\partial B$ (the slope of a linear portion of the $R(B)$ curve) versus temperature in the flux flow regime.

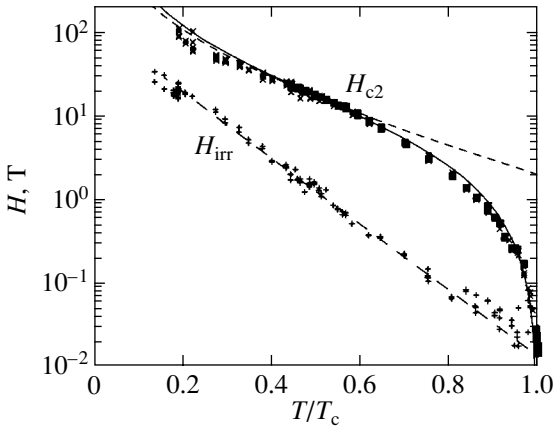

Fig. 16. Temperature variation of the characteristic fields of a BSCCO-2212 crystal determined by various methods: vertical crosses represent $H_{irr}(T)$ determined by the resistance drop to a fixed level of $R(B)/R_N(0) = 10^{-2}$; squares show the H_{c2} values determined by the method illustrated Fig. 4; oblique crosses correspond to the values determined from the Bardeen–Stephen relationship; thin solid curve shows the results of fitting to the one-parametric dependence [46]; dashed line represents "pseudo-H_{c2}" values [21]; long-dash line shows the H_{irr} approximation.

$H_{c2}(T)$ curves reflect the characteristics of the given condensed state, rather than individual features of the samples studied.

The latter statement requires additional justification, for example, by a comparative analysis of the $H_{c2}(T)$ values estimated using various components of the crystal resistance tensor. Such a comparison would be of special interest because essentially different character of the temperature dependences of the resistance measured in the basal plane (R_{ab}) and in the out-of-plane direction (R_c) implies significantly different mechanisms of the charge transfer in the normal state of the crystal. This situation poses a problem of correctly measuring the R_{ab} values in BSCCO-2212 crystals,

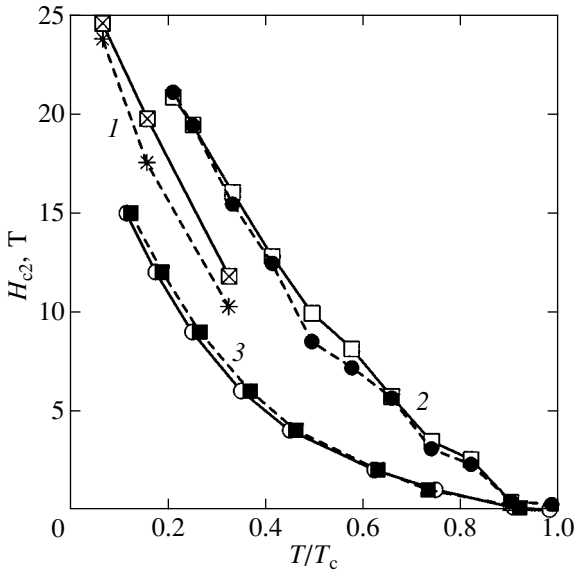

Fig. 17. A comparison of various estimates of the "resistive" upper critical field $H_{c2}(T)$, obtained from the temperature dependences of resistivities ρ_{ab} and ρ_c for the same crystal of BSCCO-2201 (solid and dashed lines connect points representing the H_{c2} values determined using $\rho_c(T)$ and $\rho_{ab}(T)$ curves, respectively): (*1*) an overdoped crystal with $T_c \approx 13$ K [28]; (*2*) an underdoped film with $T_c \approx 24.3$ K [47]; (*3*) estimates of the characteristic field for the latter film, determined on a level of $R/R_N = 0.3$.

which is an extremely difficult task because it requires using samples containing virtually no steps on the surface. Unfortunately, no such crystals have been prepared so far. Moreover, a thorough analysis of all papers reporting on the R_{ab} measurements in BSCCO-2212 revealed a dominating contribution of R_c, which makes these data absolutely useless for the purpose of this study. An almost equally unfavorable situation takes place for the study of BSCCO-2201, another representative of the class of strongly anisotropic layered cuprates characterized by significantly lower critical temperatures ($T_c \leq 27$ K). It was only quite recently that both components of the resistance tensor were studied for thin single crystal films of BSCCO-2201 grown on vicinal substrates [47]. At first glance, the results reported in [47] are free from the problem of intermixing of the resistance tensor components. The anisotropy of resistance in the films studied (about 6000) falls within the interval of values typical of the single crystals of this HTSC family, which confirms a high quality of the samples.

Figure 17 shows the estimates of the "resistive" upper critical field $H_{c2}(T)$ obtained in [48] (by analogy with [22]) using the $R_{ab}(T)$ and $R_c(T)$ data from [47]. As can be seen, there is more than satisfactory agreement between the values determined from various components of the resistance tensor. Moreover, there is no anomalous influence of the doping level on the resulting temperature dependence $H_{c2}(T)$. This conclusion follows from a comparison of the curves constructed using the data for slightly underdoped films [47] and an overdoped single crystal [28]. Finally, a good coincidence (in the scale of Fig. 17) is observed between the $H_{c2}(T)$ curves for the former films possessing close T_c values but significantly different parameters of the vortex system (as evidenced by a twofold difference in activation energies). In the absence of the results of direct measurements for BSCCO-2212, the above combination of data reported for a similar compound (Fig. 17) is indirect evidence in favor of the correctness of conclusions and estimates made in this paper based on an analysis of the interlayer resistance.

4. DISCUSSION OF RESULTS

The estimates of the upper critical field obtained in this study are well correlated with the results obtained previously [22]. This correlation confirms the adequacy of the assumptions and correctness of a complicated multistep extrapolation procedure proposed for evaluating $H_{c2}(T)$ [22]. In addition, the $H_{c2}(T)$ dependence presented in Fig. 16 was indirectly confirmed by the experiments of Suzuki *et al.* [36], which demonstrated that the current–voltage characteristic in this region is ohmic and showed dependences of the ohmic resistance on the magnetic field which qualitatively agree with Fig. 14 and allow the H_{c2} values to be estimated. As can be seen from the data in Fig. 18, the temperature dependence of the upper critical field determined from the curves reported in [36] quantitatively agrees with the results obtained in this study. It should be noted that, despite a rather unsatisfactory agreement between experiment and a theoretical approximation adopted in [36], the authors interpreted the $R(B)$ curves within the framework of a model relating the interlayer dissipation to fluctuations of the phase difference between layers.

As can be seen from Fig. 16, the obtained $H_{c2}(T)$ dependence exhibits a negative curvature in the entire range of temperatures far from T_c and in the range of the magnetic field variation over four orders of magnitude. This behavior contradicts the BCS theory but, to the first approximation, agrees qualitatively with the results of determination of the resistive upper critical field for compounds belonging to various families representing the so-called exotic superconductors (including HTSC cuprates) [20]. Although a limited range of the accessible magnetic fields did not allow the resistance of BSCCO-2212 to be studied at low temperatures, extrapolation of the empirical low-temperature asymptotics behavior as $H_{c2} = H_{c2}(0)\exp(-T/T^*)$ with $T^* = 18 \pm 2$ K provides for a rough estimate of $H_{c2}(0) \approx 220 \pm 30$ T (close to the paramagnetic threshold in the BCS model) and $\xi_{ab}(0) \approx 12$ Å. The latter value is on the order of a characteristic distance between charge carriers, which is also indicative of inapplicability of the canonical BCS theory. It is interesting to note that the

above estimates fall within the interval of values, H_{c2} = 22–400 T and $\xi_{ab}(0) \approx 9$–38 Å, obtained for BSCCO-2212 by means of the extrapolation to $T \longrightarrow 0$ proposed by Werthamer *et al.* [11] for the results of the fluctuation analysis of the magnetoresistance and magnetization in the vicinity of the critical temperature.

It should be noted that the anomalous dependence of the upper critical field obtained from the results of resistance measurements in overdoped HTSCs was attributed by some researchers (see, e.g., [4]) to the so-called irreversibility line separating the regions of existence of nondamping supercurrents and vortex fluids. In this study, the irreversibility field was determined on a fixed level of $R(H)/R_N(0) = 10^{-2}$, which provides for a very rough estimate from above for the true H_{irr} value [42]. With this circumstance borne in mind, evident quantitative and qualitative discrepancies between $H_{irr}(T)$ and $H_{c2}(T)$ for the same crystal (see Fig. 16) indicate that it would be incorrect to identify the "resistive" upper critical field $H_{c2}(T)$ with characteristics of the vortex ensemble.

As was noted above, the elegant construction [21] gave grounds for a more sophisticated interpretation of the anomalous behavior of $H_{c2}(T)$ in overdoped HTSC cuprates. Figure 16 shows the results of fitting our results to a two-parametric relationship for "pseudo-H_{c2}" values $H_{c2} \propto T^{-1}\exp(-T/T_0)$ derived in [21] within the framework of the aforementioned model description of a superconducting matrix containing a system of small inclusions with a higher T_c. As is seen from Fig. 16, the agreement between theory and experiment is far from satisfactory, especially at high temperatures. In addition, the experimental $R(T, B)$ curves obtained for the BSCCO-2212 crystals studied exhibited no significant features above T_c, as might be expected if the above model were applicable. Thus, the experimental results indicate that conclusions of this theory are inapplicable to the subject of investigation.

At the same time, the observed dependence can be satisfactorily approximated by a number of other relationships, in particular, by a two-parametric dependence predicted for the melting of a vortex lattice. This is apparently an accidental coincidence, since the existence of a vortex crystal under the condition of nonnegligible resistive response comparable with R_N is hardly probable. In addition, the initial portion of the observed $H_{c2}(T)$ curve satisfactorily agrees with a behavior predicted within the framework of a 3D-XY model stipulating a dominating contribution of the critical fluctuations near T_{c0} [50]. Note, however, that the latter conclusion was obtained in [50] using an unreasonably overstated value of T_{c0} which, in the author's opinion, decreases the reliability of this result. Finally, an $H_{c2}(T)$ curve with a negative curvature was recently reported in [51], where calculations using a 2D Hubbard model with strong repulsion were performed assuming that the van Hove singularity coincides with the Fermi level.

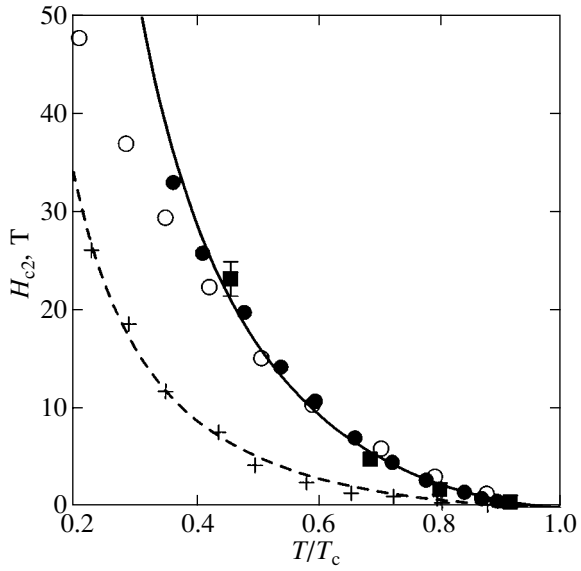

Fig. 18. Estimates of the $H_{c2}(T)$ values obtained in this work (solid curve) for a BSCCO-2212 crystal with $T_{c0} \approx$ 93.2 K compared to the published data for overdoped crystals with $T_c \approx 87$ K [36] (squares), 67 K [49] (open circles), and 78 K [49] (black circles). Crosses present the data for an underdoped sample [49] with $T_c \approx 68$ K; dashed curve shows an approximation of these data within the framework of the model [46].

Unfortunately, no indications concerning applicability of the derived formulas were provided in [51]. This circumstance somewhat decreases the significance of the generally satisfactory agreement between a three-parametric theoretical asymptotics and the experimental variation of dH_{c2}/dT observed in wide temperature interval ($T/T_c \le 0.9$).

A markedly better agreement between experiment and theory is observed, as can be seen from Fig. 16, when the experimental data are approximated by a one-parametric relationship $H_{c2} = H_0(\tau^{-1} - \tau^{1/2})^{3/2}$ predicted within the framework of the local pair theory [46] (here, $\tau = T/T_{c0}$, $H_0 \propto \Phi_0/\xi^2$, Φ_0 is the vortex quantum, and ξ is a correlation length). Taking into account that the parameter variation reduces to a trivial rescaling, the observed agreement can be considered as evidence for applicability of the theory [46].

In connection with this, it is expedient to mention the results of recent measurements of the resistance of BSCCO-2212 crystals with different levels of doping [49]. It was established that the curves of $H_{c2}(T)$, determined as described above for two overdoped crystals with T_c close to 78 and 67 K [49], not only qualitatively agree with our data (see Fig. 18), but virtually coincide with each other and with the results of Suzuki *et al.* [36] (being plotted against reduced temperature). Thus, the results reported in [49] can be considered as additional and independent evidence for adequacy of the model [46], which not only approximates the experimental data, but provides a qualitative description of the nontrivial result in Fig. 18. Indeed, a weak sensitivity to the

level of doping is inherent in the model [46] due to the smallness of exponent at the ratio l/n (l is the mean free path length, and n is the carrier concentration) in the parameter of the theory: $\xi = (l/n)^{0.25}$. In addition, the screening of impurity centers by carriers can, in principle, lead to a growth of l with an increase in the level of doping. In this hypothetical case, the l/n ratio will be virtually independent of n, which probably explains both the coincidence of results (Fig. 18) and a significant difference of data observed for an underdoped sample [49].

It should be noted that, besides a quite satisfactory description of the characteristics of superconductivity, the local pair approximation [46] (even specially refined for application to the system studied here) only qualitatively agrees with the experiment in description of the normal state [25]. At the same time, a quasilinear character of the longitudinal negative magnetoresistance in the normal state of the crystal, which is one of the most interesting results of this study, is worth a more detailed analysis. In this context, of special interest is the investigation performed by Morozov *et al.* [30] in Los Alamos, where the measurements and results of [23] were partly reproduced and an alternative explanation of the observed effect was proposed based on an analysis of the tunneling of Cooper pairs and quasiparticles in a layered superconductor with d coupling. According to [30], the maximum of $R(B)$ revealed in [23] results from a competition of these two contributions. A principal (and experimentally verifiable) conclusion from the description proposed in [30] was an essentially nonlinear character of the current–voltage characteristic in the regime of Josephson tunneling (in contrast to the linear character in the case the tunneling of quasiparticles and for the flux flow regime). It was believed [30] that direct proof of the response nonlinearity was provided by the observation of a strong distortion of the $R(B)$ curve observed upon a twofold increase in the probing current.

However, the above conclusion contradicts both the results of this study and the data reported by other researchers (see, e.g., [36]), which showed linearity of the system response in the region of small currents. The discrepancy between the results of apparently similar experiments could not be explained by features of the shape of the current–voltage characteristic (see Fig. 8), because the parameters for which deviations from the linearity were experimentally observed exceed the values employed in [30] by 1.5–2.5 orders of magnitude.[8] At the same time, it cannot be excluded that the discrepancies are related to the specific features of measurements in a pulsed field, complicated by the uncontrolled overheating of a sample by the induction currents (see Subsection 2.2). Indeed, although a maximum rate of the field buildup in [30] was almost 60 times lower than the $(\partial B/\partial t)_{max}$ value in this study, the sample crystal area in [30] was about 60 times as great as that of the sample exhibiting pronounced manifestations of the induction overheating in the experimental setup employed here (Fig. 2) and 500 times as great as the samples on which the main results were obtained in this study. Although the data reported in [30] do not allow a regular verification of the role of induction overheating (see Subsection 2.2), the coincidence of B^* values determined using the field pulses with $B_{max} = 30$ T at 50 K and 60 T at 42 K (with a fourfold difference in power dissipated in the crystal) is certainly indicative of a significant influence of overheating. An additional indication of the role of this phenomenon is a decrease in $R^*(T)$ and $B^*(T)$ below 40–50 K observed in [30].

It is obvious that, under the conditions of overheating, any additional dissipation would unavoidably increase this effect. This must be manifested, in particular, by a correlated decrease in the R^* and B^* values with increasing current, which was in fact observed in [30]. Within the framework of this scenario, an additional overheating of the crystal caused by a twofold growth of the current was quantitatively estimated [48] from a change in the position and amplitude of the $R(B)$ maximum [30]. Reduced to the measurements at the same current density, the crystal temperature was independently determined from both $B^*(T)$ and $R^*(T)$ experimental curves. Consistent estimates of the degree of overheating (3 ± 0.5 K and 2.1 ± 0.7 K) obtained in this way for the nominal temperatures of 35 and 55 K, respectively, offer convincing evidence for the "overheating" nature of the effect and provide for a reasonable explanation of the aforementioned discrepancy, thus making trivial the key experimental result obtained in [30].

Thus, it is highly probable that a discrepancy between the results obtained in this study and those reported in [30] are related to the determining influence of overheating in the latter experiment. This conclusion removes the experimental basis from interpretation of the negative magnetoresistance in terms of the tunneling of quasiparticles in a superconductor with nontrivial coupling. As a consequence, the approach to the separation of the Josephson and quasiparticle tunneling contributions proposed in [30] appears to be meaningless.

5. CONCLUSION

The interlayer magnetoresistance of a layered $Bi_2Sr_2CaCu_2O_y$ (BSCCO-2212) single crystal, representing a quasi-two-dimensional HTSC with $T_c > 91$ K,

[8] A crystal with $T_{c0} = 89$ K and $\rho_c(100\text{ K}) \approx 20\ \Omega$ cm [30] occupies an intermediate position between the samples studied in this work ($T_{c0} > 91$ K and $\rho_c(100\text{ K}) \approx 15\ \Omega$ cm) and the crystal used in [36] ($T_{c0} = 87$ K and $\rho_c(100\text{ K}) \approx 33\ \Omega$ cm. Taking into account that a characteristic level of the current densities at which the current–voltage characteristic deviates from ohmic in this study and [36] was 10–50 A/cm^2, there are no grounds to believe that the values lower by two orders of magnitude are sufficient in Los Alamos.

was studied in the normal and mixed states. A self-consistent description is provided for the experimental results obtained under the conditions of controlled ohmic crystal response and negligibly small induction overheating.

The character of the longitudinal magnetoresistance is determined for the crystal in the normal state and under the conditions of superconductivity suppression by a strong magnetic field. An estimate of $R_N(0, T)$, the interlayer resistance in the absence of superconductivity, obtained by extrapolation agrees well with the results of investigation of the superconductivity breakdown by a pulsed current. Based on these results and recently published data, an interpretation of the origin of the "quasi-particle" resistance of BSCCO-2212 is proposed. The character of variation of the ohmic resistance as a function of the magnetic field and the temperature in the region of the resistive state was experimentally determined, and it is demonstrated that a self-consistent description of the whole body of experimental data can be provided within a restricted class of theoretical models of the vortex state in HTSCs. A correlation is found between the effective barrier heights determined on the two sides of the irreversibility line $H_{irr}(T)$, and it suggested that the two barriers may be of a common origin. It is established that a change in the character of the resistance variation above the level of $(0.05–0.1)R_N$ reflects the transition from temperature-activated to viscous vortex (fluxoid) flow; the character of the field dependence of the sample resistance $R(B)$ in the latter state was determined. A method is proposed for separating the temperature-dependent contributions from the normal and resistive states to the total magnetoresistance of a crystal. An anomalous character of the temperature upon the shape of the superconducting transition in a magnetic field was established. A critical field $H_{c2}(T)$ for the formation of a superconducting nucleus in BSCCO-2212 was determined, and a quantitative correlation between various estimates of $H_{c2}(T/T_c)$ was found in a broad range of the doping level. A systematic analysis of the results of investigation of BSCCO-2201 crystals revealed a quantitative agreement between H_{c2} values determined using planar and interlayer resistances, which is evidence for adequacy of the method of $H_{c2}(T)$ determination used in this study.

The experimental results are compared to theoretical dependences for the upper critical field; inapplicability of the BCS theory to this case is demonstrated. The experimental $H_{c2}(T)$ is satisfactorily described by a one-parametric relationship predicted for $H_{c2}(T)$ in the bosonic approximation.

An alternative explanation of the results, based on the theory of interlayer tunneling of the Cooper pairs and quasiparticles in a layered superconductor with *d* coupling, is analyzed in detail, and it is shown that the main conclusions of this model contradict the known experimental facts.

ACKNOWLEDGMENTS

The author is especially grateful to M. Springford, A.S. Aleksandrov, and V.B. Geshkenbein for numerous fruitful discussions, and to S. Hayden and G.T. Janssen for valuable methodological suggestions and practical help in the development of experimental techniques. Many thanks to the staff of the Laboratory of Low Temperatures for hospitality, support, and technical assistance, and to several fellows of the Coffee-Club of the Kapitza Institute of Physical Problems for stimulating discussions and relaxed atmosphere.

This study was performed within the framework of projects supported by the Leverhulme Trust (Ref: F/00261/H), the Russian Foundation for Basic Research (project no. 01-02-16532), and partly by the Scientific Council on Superconductivity (project no. 96115).

REFERENCES

1. G. Blatter, M. V. Feigel'man, V. B. Geshkenbein, *et al.*, Rev. Mod. Phys. **66**, 1125 (1994).
2. M. Charalambous, J. Chaussy, P. Lejay, and V. Vinokur, Phys. Rev. Lett. **71**, 436 (1993).
3. E. Zeldov, D. Majer, M. Konczykowski, *et al.*, Nature **375**, 373 (1995).
4. J. R. Cooper, J. W. Loram, and J. M. Wade, Phys. Rev. B **51**, 6179 (1995).
5. A. A. Varlamov, Physica C (Amsterdam) **282-287**, 248 (1997); A. I. Larkin and A. A. Varlamov, in *Physics of Conventional and Unconventional Superconductors*, Ed. by K. H. Bennerman and J. B. Ketterson (Springer-Verlag, Berlin, 2001).
6. R. Jin, A. Shilling, and H. R. Ott, Phys. Rev. B **49**, 9218 (1994).
7. W. E. Lawrence and S. Doniach, in *Proceedings of the 12th International Conference on Low-Temperature Physics, Kyoto, 1971*, Ed. by E. Kanda (Keigaki Publ., Tokyo, 1971), p. 361.
8. Quiang Li, K. Shibutani, M. Suenaga, *et al.*, Phys. Rev. B **48**, 9877 (1993).
9. A. Pomar, M. V. Ramallo, J. Mosqueira, *et al.*, Phys. Rev. B **54**, 7470 (1996).
10. L. P. Gor'kov, Zh. Éksp. Teor. Fiz. **37**, 833 (1959) [Sov. Phys. JETP **10**, 593 (1960)].
11. N. R. Werthamer, E. Helfand, and P. C. Hohenberg, Phys. Rev. **147**, 295 (1966).
12. U. Welp, W. K. Kwok, K. G. Vandervoort, *et al.*, Phys. Rev. Lett. **62**, 1908 (1989).
13. T. T. M. Palstra, B. Batlogg, L. F. Schneemeyer, and J. V. Waszczak, Phys. Rev. Lett. **64**, 3090 (1990).
14. C. Escribe-Filippini *et al.*, Physica C (Amsterdam) **210**, 133 (1993).
15. Z. Tesanovic, L. Xing, L. Bulaevskii, *et al.*, Phys. Rev. Lett. **69**, 3563 (1992).
16. D. E. Prober, M. R. Beasley, and R. E. Schwall, Phys. Rev. B **15**, 5245 (1977).
17. A. P. Mackenzie, S. R. Julian, G. G. Lonzarich, *et al.*, Phys. Rev. Lett. **71**, 1238 (1993).

18. M. S. Osofsky, R. J. Soulen, S. A. Wolf, *et al.*, Phys. Rev. Lett. **71**, 2315 (1993).
19. D. J. C. Walker, O. Labrode, A. P. Makenzie, *et al.*, Phys. Rev. B **51**, 9375 (1995).
20. B. Brandow, Phys. Rep. **296**, 1 (1998).
21. V. B. Geshkenbein, L. B. Ioffe, and A. J. Millis, Phys. Rev. Lett. **80**, 5778 (1998).
22. A. S. Alexandrov, V. N. Zavaritsky, W. Y. Liang, *et al.*, Phys. Rev. Lett. **76**, 983 (1996).
23. V. N. Zavaritsky and M. Springford, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 420 (1998) [JETP Lett. **68**, 448 (1998)].
24. V. N. Zavaritsky, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 116 (2000) [JETP Lett. **71**, 80 (2000)].
25. V. N. Zavaritsky, M. Springford, and A. S. Alexandrov, Europhys. Lett. **51**, 334 (2000); Physica B (Amsterdam) **294-295**, 363 (2001).
26. S. L. Cooper and K. E. Gray, in *Physical Properties of High Temperature Superconductors*, Ed. by D. M. Ginsberg (World Sci., Singapore, 1994), Vol. IV, p. 61.
27. N. V. Zavaritsky, V. N. Zavaritsky, A. P. Mackenzie, and Yu. F. Orlov, Pis'ma Zh. Éksp. Teor. Fiz. **60**, 188 (1994) [JETP Lett. **60**, 193 (1994)].
28. Y. Ando, G. S. Boebinger, A. Passner, *et al.*, Phys. Rev. Lett. **77**, 2065 (1996); Physica C (Amsterdam) **282**, 240 (1997).
29. G. Briceno, M. F. Crommie, and A. Zettl, Phys. Rev. Lett. **66**, 2164 (1991).
30. N. Morozov, L. Krusin-Elbaum, T. Shibauchi, *et al.*, E-print archive, cond-mat/9912194; Physica C (Amsterdam) **341**, 1511 (2000).
31. V. N. Zavaritsky, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 543 (1999) [JETP Lett. **70**, 556 (1999)].
32. Y. F. Yan, P. Matl, J. M. Harris, and N. P. Ong, Phys. Rev. B **52**, R751 (1995).
33. T. Yasuda, M. Tonouchi, and S. Takano, Czech. J. Phys. **46**, 1265 (1996); Physica C (Amsterdam) **289**, 109 (1997).
34. A. Yurgens, D. Winkler, N. V. Zavaritsky, and T. Claeson, Physica C (Amsterdam) **282**, 2293 (1997).
35. Yu. L. Latyshev, T. Yamashita, L. N. Bulaevskii, *et al.*, E-print archive, cond-mat/9903256.
36. M. Suzuki, T. Watanabe, and A. Matsuda, Phys. Rev. Lett. **81**, 4248 (1998); **82**, 5361 (1999).
37. V. Krasnov, A. Yurgens, D. Winkler, *et al.*, E-print archive, cond-mat/0002094.
38. C. E. Gough, P. J. Thomas, J. C. Fenton, and G. Yang, E-print archive, cond-mat/0002246.
39. M. V. Feigel'man, V. B. Geshkenbein, and A. I. Larkin, Physica C (Amsterdam) **167**, 177 (1990); L. Burlachkov, V. B. Geshkenbein, A. I. Larkin, *et al.*, Phys. Rev. B **50**, 16770 (1994).
40. H. J. Jensen, P. Minnhagen, E. Sonin, and H. Weber, Europhys. Lett. **20**, 463 (1992).
41. M. A. Skvortsov and V. B. Geshkenbeĭn, Zh. Éksp. Teor. Fiz. **105**, 1379 (1994) [JETP **78**, 743 (1994)].
42. V. N. Zavaritsky, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 629 (1997) [JETP Lett. **65**, 663 (1997)].
43. V. N. Zavaritsky and N. V. Zavaritsky, Pis'ma Zh. Éksp. Teor. Fiz. **54**, 25 (1991) [JETP Lett. **54**, 22 (1991)]; V. N. Zavaritsky and N. V. Zavaritsky, Physica C (Amsterdam) **185–189**, 2142 (1991).
44. L. N. Bulaevskiĭ, private communication (2000).
45. Y. Yeshurun, A. P. Malozemoff, and A. Shaulov, Rev. Mod. Phys. **68**, 911 (1996).
46. A. S. Alexandrov, Phys. Rev. B **48**, 10571 (1993); A. S. Alexandrov, V. V. Kabanov, and N. F. Mott, Phys. Rev. Lett. **77**, 4796 (1996).
47. Y. Z. Zhang, R. Deltour, J.-F. de Marneffe, *et al.*, Phys. Rev. B **61**, 8675 (2000).
48. V. N. Zavaritsky, M. Springford, and A. S. Alexandrov, E-print archive, cond-mat/0011192.
49. T. Shibauchi, L. Krusin-Elbaum, Ming Li, *et al.*, E-print archive, cond-mat/0104261.
50. T. Schneider, E-print archive, cond-mat/0001258.
51. R. O. Zaĭtsev, Pis'ma Zh. Éksp. Teor. Fiz. **65**, 71 (1997) [JETP Lett. **65**, 74 (1997)].

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 821–833.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 955–970.
Original Russian Text Copyright © 2002 by Anishchenko, Luchinsky, McClintock, Khovanov, Khovanova.

NONLINEAR PHYSICS

Fluctuational Escape from a Quasi-Hyperbolic Attractor in the Lorenz System

V. S. Anishchenko[a], D. G. Luchinsky[b], P. V. E. McClintock[b], I. A. Khovanov[a, *], and N. A. Khovanova[a]

[a]*Saratov State University, ul. Universitetskaya 42, Saratov, 410026 Russia*
[b]*Lancaster University, Lancaster LA1 4YB, Great Britain*
**e-mail: igor@chaos.ssu.runnet.ru*
Received July 13, 2001

Abstract—Noise-induced escape from the basin of attraction of a quasi-hyperbolic chaotic attractor in the Lorenz system is considered. The investigation is carried out in terms of the theory of large fluctuations by experimentally analyzing the escape prehistory. The optimal escape trajectory is shown to be unique and determined by the saddle-point manifolds of the Lorenz system. We established that the escape process consists of three stages and that noise plays a fundamentally different role at each of these stages. The dynamics of fluctuational escape from a quasi-hyperbolic attractor is shown to differ fundamentally from the dynamics of escape from a nonhyperbolic attractor considered previously [1]. We discuss the possibility of analytically describing large noise-induced deviations from a quasi-hyperbolic chaotic attractor and outline the range of outstanding problems in this field. *© 2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Chaotic oscillations are observed in many fields of physics. At least in two of these fields, hydrodynamics and laser physics, chaotic dynamics is described by the same differential equations, the Lorenz system [2]. A minimum set of equations for Benard–Rayleigh convection (the Lorenz system) was derived by Salzman [3]. Subsequently, Haken [4] showed that the Lorenz system also describes the generation of a lasing transition in a single-mode laser with a uniformly broadened spectral line when the cavity eigenfrequency is exactly equal to the lasing transition frequency [5, 6]. The Lorenz system is a classical model of low-dimension chaos. It is used to study such fundamental problems of nonlinear dynamics [7] as synchronization, chaos control, hidden information transmission, invariant reconstruction from time series, and the like; whence the unremitting interest of scientists in this system. Increased interest in the Lorenz system also stems from the fact that it exhibits quasi-hyperbolic chaos under certain conditions. Since this chaos can be described mathematically [8, 9], the results of numerical experiments can be checked on the basis of a rigorous theory. Such a situation is not typical of systems with chaotic attractors; the principal and virtually the only tools for investigating the latter are the various numerical methods of solving differential equations and numerical bifurcation analysis.

In attempting to relate the Lorenz system to physical experiments, it is essential to make allowance for the influence of fluctuations on its dynamical regimes and transitions. Note that the influence of fluctuations on the fluctuational regimes of the Lorenz system was studied in considerable detail both theoretically [10, 11] and experimentally [12, 13]. However, despite a large number of studies, there are currently no answers to many questions, including the key ones: What is the dynamics of the Lorenz system during large noise-driven deviations of the trajectory from a chaotic attractor? Can the attractor quasi-potential be constructed? What is the mechanism of noise-induced escape from the basin of attraction of a quasi-hyperbolic attractor? The latter question, related to the problem of noise-induced transitions in nonequilibrium systems, is a fundamental problem in the fluctuation theory of nonequilibrium systems [14]. It is of broad interdisciplinary interest in terms of practical applications, for example, for Josephson junctions connected in series [15], biological transport [16], and laser systems [17], as well as in controlling transitions between attractors in multistable systems [18–20].

Recent progress in solving the problem of noise-induced transitions in nonequilibrium systems driven by colored noise [21–25] and a nongradient [26–29] or periodic external force [30–32], as well as in self-oscillating systems [33, 34], has been achieved through the use of Hamiltonian formalism [35] or its equivalent formalism in terms of path integrals. The idea of optimal trajectories along which the system fluctuates to states far removed from equilibrium [36] underlies the Hamiltonian approach. A mathematical variant of these physical concepts is embodied in the asymptotic formulas for solving the Fokker–Planck equation written in terms of rays (i.e., the solutions of the Hamilton equations) or fronts (i.e., the solutions of the

Hamilton–Jacobi equation). Since the Hamiltonian formalism is successfully used to solve many problems of the nonequilibrium fluctuation theory, the question of how it can be generalized to a broader class of dissipative systems, including chaotic ones, now acquires particular urgency. The difficulty of this generalization complicates the direct application of analytical methods. Recently, however, a new experimental approach to investigating optimal trajectories [37] has been proposed. It is based on the Hamiltonian formalism and on measurements of the so-called probability distribution of the fluctuational prehistory. This method basically involves continuously tracking the system dynamics and constructing the distribution of all realizations of the fluctuational trajectories that bring the system from equilibrium to a given remote state. The advantages of this approach were demonstrated [38–41] by studying some fundamental properties of the symmetry of optimal trajectories and singularities in their distribution. It has recently been shown [1, 20] that based on a statistical analysis of the dynamics of fluctuational trajectories, the method can be used to find the optimal trajectory for the noise-induced escape from the domain of attraction of a chaotic attractor in a nonlinear oscillator driven by a periodic external force. Beyond the boundaries of the chaotic attractor, this trajectory was shown to be unique and formed by the coalescence of several equally probable trajectories that emerge from the chaotic attractor. In [1, 20], we studied escape from a nonhyperbolic attractor. Here, we investigate noise-induced escape from a quasi-hyperbolic chaotic attractor in the Lorenz system and, thus, answer the question of how the mechanism of escape from a chaotic attractor depends on its structure.

The paper consists of three sections. In Section 2, we consider the Lorenz system and the structure of its phase space. The basics of the theoretical approach in terms of the Hamiltonian formalism and the experimental approach based on measuring the fluctuational prehistory are presented in Section 3. In this section, we also consider the influence of fluctuations on the probability measure of a chaotic attractor. Our results are presented and discussed in Section 4. In the Conclusion, we summarize our conclusions and outline the range of outstanding problems.

2. THE LORENZ SYSTEM

The Lorenz system was originally obtained as a simplified model to describe the convection of a fluid between two planes with different temperatures, with the temperature and gravitational gradients being directed oppositely [3]. This model can be written as

$$\begin{aligned}\dot{q}_1 &= \sigma(q_2 - q_1),\\ \dot{q}_2 &= rq_1 - q_2 - q_1q_3,\\ \dot{q}_3 &= q_1q_2 - bq_3 + \xi(t),\end{aligned} \qquad (1)$$

$$\langle \xi(t)\rangle = 0, \quad \langle \xi(t)\xi(0)\rangle = D\delta(t),$$

or, in vector form,

$$\begin{aligned}\dot{\mathbf{q}} &= \mathbf{K}(\mathbf{q}) + \mathbf{f}(t),\\ \mathbf{K} &= \{K_1, K_2K_3\}\\ &= \{\sigma(q_2 - q_1), rq_1 - q_2 - q_1q_3, q_1q_2 - bq_3\},\\ \mathbf{f}(t) &= \{0, 0, \xi(t)\}, \quad \langle \mathbf{f}\rangle = \mathbf{0},\\ \langle f_i(t) f_j(0)\rangle &= DQ_{ij}\delta(0), \quad Q_{ij} = \delta_{i3}\delta_{j3}.\end{aligned} \qquad (2)$$

Here, q_1, q_2, q_3 are the dynamical variables; and σ, r, b are the parameters that have the following physical meaning in the context of fluid convection: σ is the Prandtl number (the ratio of kinematic viscosity to thermal diffusivity), $r = R/R_c$ is the reduced Rayleigh number (R is the Rayleigh number and R_c is the critical Rayleigh number at which convection arises), $b = 4\pi^2/(\pi^2 + k^2)$ is a constant that characterizes the physical size of the system, k is a dimensionless number related to the spatial periodicity in the arrangement of fluid waves, q_1 is proportional to the intensity of the convective motion, q_2 is proportional to the temperature difference between the rising and sinking fluid flows, and q_3 is proportional to the deviation of the vertical temperature profile between the planes from the linear one. To simplify our analysis, we add a white-noise source $\xi(t)$ only to the third equation of system (1), which does not break down the system mirror symmetry; D is the noise intensity; $\delta(t)$ is the delta function; and δ_{i3} and δ_{j3} are the Kronecker symbols.

Equations (1) also describe traveling waves excited in a ring cavity filled with two-level atoms with a uniformly broadened spectral line [4]. With respect to a laser system, the variables and parameters of system (1) can be interpreted as follows: q_1 and q_2 are the normalized amplitude and polarization of the electric field, respectively; q_3 is the normalized population inversion in the atomic system; $\sigma = k/\gamma_1$, $r = \Lambda + 1$, and $b = \gamma_2/\gamma_1$, where k is the field relaxation rate in the cavity, γ_1 and γ_2 are the polarization and inversion constants, respectively, and Λ is the pumping parameter. Note that an infrared laser was suggested as a realization of the Lorenz system [12].

The Lorenz equations are simple in structure and contain only two nonlinear terms. Let us briefly consider the main bifurcations in system (1).[1] We fix $\sigma = 10$ and $b = 8/3$ and vary the parameter r. In this case, there are two global bifurcations. Their diagram is shown in Fig. 1. A supercritical pitchfork bifurcation (triple-equilibrium bifurcation) [7] is observed at $r = 1$;

[1] A more detailed and rigorous bifurcation analysis of the Lorenz system can be found in [8, 9].

the equilibrium position at the coordinate origin $O(0, 0, 0)$ becomes unstable and two stable states appear:

$$P_1(\sqrt{b(r-1)}, \sqrt{b(r-1)}, r-1)$$

and

$$P_2(-\sqrt{b(r-1)}, -\sqrt{b(r-1)}, r-1).$$

There are two stable equilibrium positions (stable foci) P_1 and P_2, a saddle equilibrium position O, and its one-dimensional unstable and two-dimensional stable manifolds in the phase space of the system. The two branches of the unstable manifold are the separatrices of the saddle point O.

The second bifurcation is observed at

$$r = \frac{\sigma(\sigma + b + 3)}{\sigma - b - 1} \approx 24.74, \qquad (3)$$

when the foci P_1 and P_2 lose their stability and the chaotic attractor becomes the only attractor in the system.

In addition, we note two local bifurcations.

A biasymptotic homoclinic contact of the separatrices of the saddle point O (not shown in Fig. 1) takes place at $r \approx 13.926$. When r passes through this bifurcation value, two saddle cycles, L_1 and L_2, are generated and a one-dimensional (in Poincaré section) hyperbolic chaotic set Ω_1 appears. Since this set is not attracting, the foci P_1 and P_2 that the separatrices of point O approach remain the system attractors. The second local bifurcation takes place at $r \approx 24.06$. The separatrices of the saddle point O are no longer closed at P_1 and P_2 but approach the unstable manifolds of the saddle cycles L_1 and L_2. As a result, a stable two-dimensional (in Poincaré section) chaotic set Ω_2, the Lorenz attractor, emerges at the location of the unstable set Ω_1 for $r > 24.06$. The rough saddle cycles with their stable manifolds transverse to the attractor are everywhere dense in the chaotic Lorenz attractor. Apart from the saddle cycles, the set Ω_2 contains the separatrices and the saddle point O. The latter is determined by the existence of two (because the Lorenz system is symmetric) trajectories, T_1 and T_2 (Fig. 2), which emerge from the chaotic attractor, fall on the stable manifold W^s of the saddle point at some point A, and then approach point O and, hence, separatrices Γ_1 and Γ_2. The existence of a saddle point in the chaotic attractor defines the nonhyperbolicity of the Lorenz attractor and the prefix "quasi" in the definition of a quasi-hyperbolic chaotic attractor [9]. The equilibrium positions P_1 and P_2 remain stable.

Thus, the two stable equilibrium positions and the chaotic attractor coexist in the phase space of the Lorenz system for $r \in [24.06; 24.74]$. Let us fix $r = 24.08$ in this range and consider noise-induced escape from the chaotic attractor to the basin of attraction of the equilibrium states. In other words, we consider

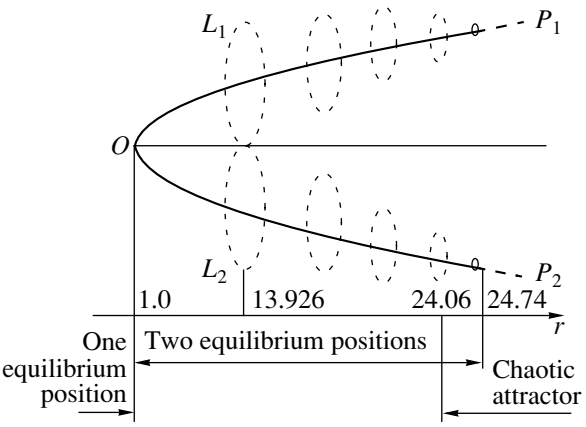

Fig. 1. A bifurcation diagram of the Lorenz system for $\sigma = 10$ and $b = 8/3$. The dashed and solid lines indicate unstable and stable states, respectively.

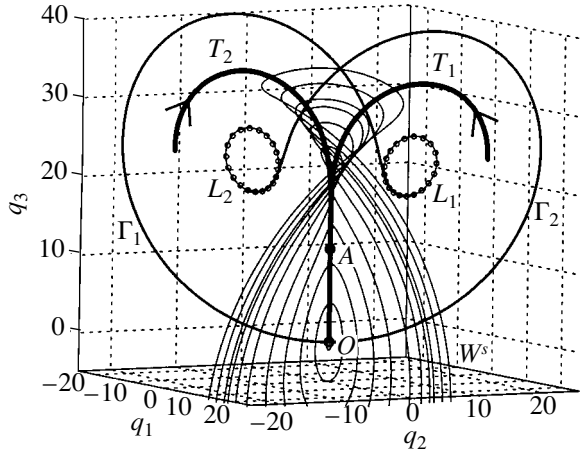

Fig. 2. Separatrices Γ_1 and Γ_2 and the two-dimensional stable manifold W^s of the saddle equilibrium position O. The saddle cycles L_1 and L_2 are indicated by circles. Trajectories T_1 and T_2 (see the text) are represented by the heavy lines.

large noise-induced deviations from the chaotic attractor.

Let us first investigate the structure of the phase space of system (1) for the coexisting stable equilibrium positions P_1 and P_2 and chaotic attractor at the following parameters: $\sigma = 10$, $b = 8/3$, and $r = 24.08$ (Fig. 3). The equilibrium positions P_1 and P_2 have three negative eigenvalues (–13.627, –0.020 + i9.504, and –0.020 – i9.504), with the last two being complex conjugate. This implies that the two eigendirections of points P_1 and P_2 form a two-dimensional surface: the rate of attraction on this surface is much smaller than the rate of attraction in the first eigendirection.

The saddle cycles L_1 and L_2 surround the states P_1 and P_2 and lie at the intersection of the two-dimensional unstable and stable manifolds [42]. The unstable manifold on one side goes to the equilibrium positions and

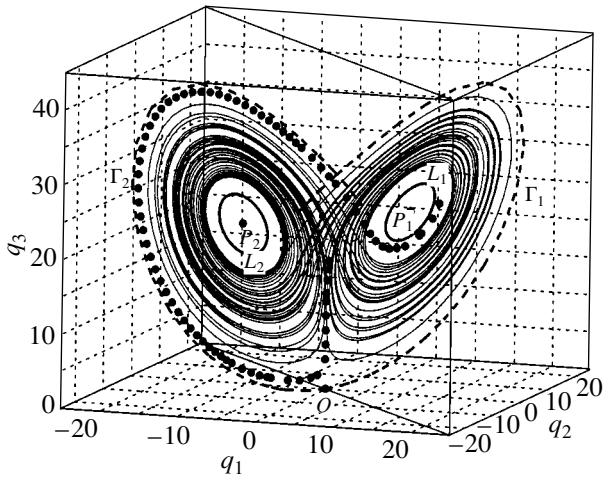

Fig. 3. Structure of the phase space for the Lorenz system. The thin solid line is the trajectory of the chaotic attractor, the dashed line indicates separatrices Γ_1 and Γ_2, and the initial segment of one of the escape trajectories is represented by filled circles.

is closed there, but on the other side it goes to the chaotic attractor. The stable manifold forms a tube near the equilibrium positions [42]. Since the saddle cycles have the multipliers (1.0000, 1.0280, 0.0001), the trajectories slowly go away from the cycles along the unstable manifold and rapidly approach along the stable manifold.

The boundaries of the chaotic attractor are specified by the initial segments of the separatrices that closely approach the saddle cycles L_1 and L_2 (see Fig. 2) and theoretically [9] belong to the attractor. However, numerical studies show that the probability of the trajectory falling in the neighborhood of the separatrices is exponentially small compared to the probability of its being in other segments of the chaotic attractor. Indeed, this probability is determined by the probability of motion in the neighborhood of trajectories T_1 and T_2 (Fig. 2). Computations indicate that, for the system trajectory to fall within $\epsilon = 0.1$ of the separatrix, it must pass in the close neighborhood of T_1 and T_2, $\epsilon \approx 10^{-7}$, which is unlikely.

Based on the calculated time during which a trajectory stays in a given neighborhood of the attractor, we computed the probability measure of the chaotic attractor for the Poincaré section $q_3 = r - 1$. The form of the Lorenz attractor in the section is known [8] to be similar to a one-dimensional curve in the q_1–q_2 plane. This allows us to pass from the two-parameter to the one-parameter probability measure [43] and to consider the probability measure $p(q_1)$ for one coordinate. The probability measure $p(q_1)$ in the Poincaré section $q_3 = r - 1$ (trajectories crossing the $q_3 = r - 1$ plane from the bottom upward were taken into account) has the form shown in Fig. 4a. We see that the probability of the chaotic-attractor trajectory falling in the neighborhood of separatrices Γ_1 and Γ_2 is exponentially small, and the typical trajectory rarely visits the neighborhood of the unstable manifolds of the saddle point O with a probability much larger than zero. The probability measure changes only slightly in the presence of noise, as previously shown theoretically in [10, 11]. In the presence of noise, however, the probability of the trajectory falling in the neighborhood of the separatrix increases (Fig. 4b). Thus, the probability measure of the chaotic attractor has a smooth structure without singularities, $p(q_1)$ is

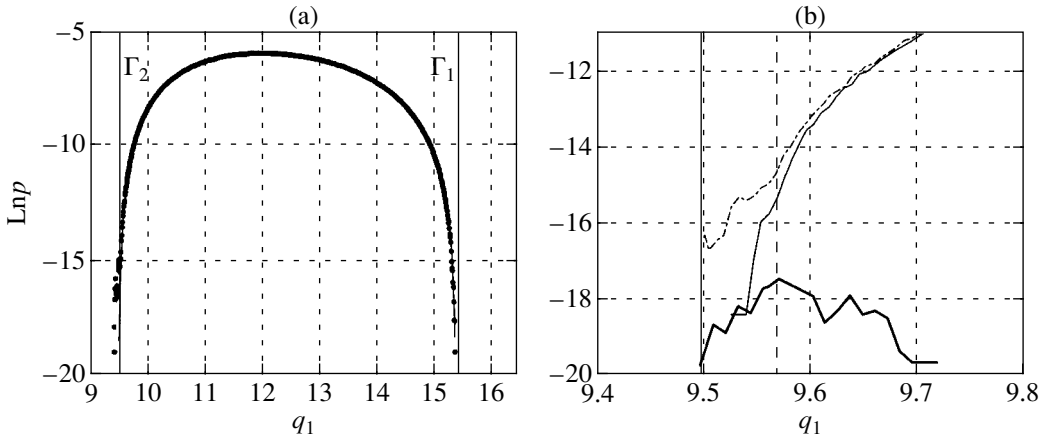

Fig. 4. (a) The logarithm of probability measure $p(q_1)$ for $q_1 > 0$ computed in the absence (solid curve) and presence (dots) of noise of intensity $D = 0.001$. The logarithm of probability measure $p(q_1)$ for $q_1 < 0$ has the same form, and it can be obtained by substituting $-q_1$ for q_1. The solid vertical lines mark the coordinates that correspond to the intersection of the $q_3 = r - 1$ plane with the separatrices of the saddle point O from the bottom upward. (b) The tail [an enlarged part of panel (a)] of probability measure $p(q_1)$ computed in the absence (thin solid curve) and presence (dash–dotted curve) of noise of intensity $D = 0.001$. The solid and dashed vertical lines mark the coordinates that correspond to the intersection of the $q_3 = r - 1$ plane with the separatrix and the optimal escape trajectory from the bottom upward. The heavy solid curve represents the distribution of escape trajectories in Poincaré section.

nonuniformly distributed over the attractor, there are regions with an exponentially small probability of the trajectory falling within them, and the presence of noise does not change qualitatively the distribution structure but increases the probability of the trajectory being in the neighborhood of the separatrices.

Our analysis of the basins of attraction of the stationary points P_1 and P_2 in three-dimensional space shows that the separatrices lie virtually at the boundaries of the basins of attraction in the interval $0 < q_3 < 5$. In addition, the separatrices come close to the basin boundaries of P_1 and P_2 in the neighborhood of the saddle cycles L_1 and L_2. Thus, within a large region of phase space, the separatrices and, hence, the chaotic attractor come close to the boundaries of the basins of attraction of the stationary points. Therefore, we may assume that escape from the chaotic attractor can be observed at any location where the separatrices come close to the boundaries.

Thus, we see that, in the presence of noise, there is a finite probability of fluctuational escape from the basin of attraction of the quasi-hyperbolic attractor in the Lorenz system. Before discussing the mechanism of fluctuational transitions and the possibility of analytically estimating the escape probability in the Lorenz system, we briefly describe the experimental methods for investigating the fluctuational dynamics in nonequilibrium systems.

3. THE FLUCTUATIONAL PREHISTORY PROBABILITY DISTRIBUTION AND ANALYSIS OF NONEQUILIBRIUM FLUCTUATIONAL TRANSITIONS

The probability characteristics of the Lorenz system in the presence of white Gaussian noise are described by the solution of the Fokker–Planck equation

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial q_i}K_i(\mathbf{q})\rho + \frac{D}{2}\frac{\partial^2}{\partial q_i \partial q_j}(Q_{ij}\rho), \quad Q_{ij} = \delta_{i3}\delta_{j3}, \tag{4}$$

which is written for the corresponding Langevin equations (2); here, $\rho = \rho(\mathbf{q}, t)$ is the probability density that the system will be in state $\mathbf{q}$ at time t. In principle, the solution of the Fokker–Planck equation provides the most complete information on the dynamics of the distribution function for system (2). However, for systems far from thermal equilibrium, for which the detailed balance condition is not satisfied [in particular, for our Lorenz system, the detailed balance condition [11] is not satisfied, because the vector field of system (2) is not potential], there are no general methods for solving (4). Therefore, various approximate approaches are used to describe the fluctuational dynamics of such systems. Thus, the idea of optimal trajectories and the concept of the fluctuational prehistory probability distribution prove to be very useful. By definition [37], the fluctuational prehistory probability distribution $p_h(\mathbf{q}_f, t_f; \mathbf{q}, t; \mathbf{q}_i, t_i)$ specifies the probability that the system is at point $\mathbf{q}$ at time t, provided that the system was at point $\mathbf{q}_i$ at initial time t_i and that its position at final time t_f is at point $\mathbf{q}_f$ ($t_i < t < t_f$). Thus, p_h describes the dynamics of the distribution function during the system transition from state $(\mathbf{q}_i, t_i)$ to state $(\mathbf{q}_f, t_f)$ normalized to the probability of the transition between these states. This quantity differs from the ordinary conditional probability, because both the initial and the final state of the system are fixed in the definition of p_h (see [37, 44] for details). Note also that the definition of p_h imposes no constraints either on the times t_i and t_f or on the noise intensity D.

It was shown in [37] that direct experimental measurements of p_h can reveal important information on the fluctuational dynamics of the system. Such measurements are of particular importance in investigating noise-induced transitions in nonequilibrium systems in the weak-noise limit. It turns out that the prehistory probability distribution p_h for $D \longrightarrow 0$ has a sharp peak along some deterministic trajectories in the system configuration space. The trajectories determined in this way are called optimal trajectories. In general, there is only one optimal trajectory that couples the specified initial and final states of the system.

To determine the optimal trajectories, we must turn to an asymptotic analysis of the solution to the Fokker–Planck equation for $D \longrightarrow 0$ (see, e.g., [21–41, 43–59]). In this approximation, D is a small parameter of the equation and, by analogy with the WKB approximation, the stationary (for $t \longrightarrow \infty$) probability density can be written as (see, e.g., [35])

$$\rho(\mathbf{q}) \approx C(\mathbf{q})\exp\left(\frac{-S[\mathbf{q}]}{D}\right) \text{ for } D \longrightarrow 0. \tag{5}$$

The quantity $S[\mathbf{q}]$ is the activation energy of the fluctuations near point $\mathbf{q}$ [11, 47], and $C(\mathbf{q})$ is the preexponential factor. Substituting (5) in (4) and collecting the terms proportional to D^{-1}, we obtain [35] the Hamilton–Jacobi equation for $S[\mathbf{q}]$ (see also [11, 29, 33, 45, 46])

$$\frac{\partial S}{\partial t} + H\left(\mathbf{q}, \frac{\partial S}{\partial q}\right) = 0, \quad H = \frac{1}{2}\mathbf{pQp} + \mathbf{pK}(\mathbf{q}),$$
$$\dot{\mathbf{q}} = \frac{\partial H}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = -\frac{\partial H}{\partial \mathbf{q}}, \quad \dot{S} = \frac{1}{2}\mathbf{p}^2, \quad \mathbf{p} = \frac{\partial S}{\partial \mathbf{q}}. \tag{6}$$

Here, H is the Wenzel–Freidlin Hamiltonian [35], and $S[\mathbf{q}]$ can be interpreted as the classical action of system (6) that determines the energy spent on the system switching from state $\mathbf{q}_i$ to state $\mathbf{q}_f$. In general, $S[\mathbf{q}]$ is a multivalued function of $\mathbf{q}_i$. However, the minimum activation energy $\min S[\mathbf{q}]$ has a physical meaning (i.e., is a physically observable quantity). The optimal force $\mathbf{p}_{\text{opt}}(t)$ and optimal trajectory $\mathbf{q}_{\text{opt}}(t)$ of system motion from $\mathbf{q}_i$ to $\mathbf{q}_f$ calculated from (6) corresponds to this energy. If the initial time corresponds to $-\infty$, then

$\min S[\mathbf{q}]$ defines the quasi-potential of the nonequilibrium system (2) [11, 43, 47]:

$$\rho(\mathbf{q}) \approx C(\mathbf{q})\exp\left(\frac{-S[\mathbf{q}_{\text{opt}}]}{D}\right) \text{ for } D \longrightarrow 0. \qquad (7)$$

Thus, in the weak-noise limit, the stochastic dynamics of the system is described in terms of motion along the optimal (in terms of energy expenditure) trajectories $\mathbf{q}_{\text{opt}}(t_i, \mathbf{q}_i; t; t_f, \mathbf{q}_f)$; here, $\mathbf{q}_i$ is the initial state of the system and $\mathbf{q}_f$ is the specified point in the space of system states located at distances much larger than $\sqrt{D}$ from the attractor.

It is worth noting that the Hamiltonian in (6) is similar to the Pontryagin Hamiltonian, which corresponds to the problem of switching system (2) from state $\mathbf{q}_i$ to state $\mathbf{q}_f$ while minimizing the energy of the additive control function [20] in the absence of constraints on the form of the function and on the switching time.

If we choose the initial state $\mathbf{q}_i$ on the attractor and the final state $\mathbf{q}_f$ at the attractor boundary, then the extreme trajectories described by Hamiltonian (6) correspond to the paths of escape from the attractor and the vector $\mathbf{p}$, which tends to zero when $t_i \longrightarrow -\infty$ and $t_f \longrightarrow \infty$, determines the fluctuational force that switches the system from one state to the other along a given escape path [40, 48].

This formalism corresponds to the following physical picture of noise-induced escape from the basin of attraction of an attractor: the system spends most of the time in fluctuating near the attractor but executes rare fluctuations that cross the attractor boundary. When such rare fluctuations take place, the system moves along almost a deterministic (optimal) trajectory. In general, the trajectory of escape from the domain of attraction is unique (or there are several such trajectories, because the system is symmetric) [30, 49–51]. For regular fixed-point and limit-cycle attractors, this physical picture of noise-induced escape was confirmed numerically [14, 15, 39, 46].

Note that the optimal trajectories in equilibrium systems are symmetric in time to the system relaxation trajectories [36] and, hence, these can be determined, at least in principle, as the time-reversed relaxation trajectories of a deterministic system. Since there is no such symmetry in nonequilibrium systems, investigating the fluctuational prehistory is of fundamental importance in understanding the mechanism of fluctuational escape from an attractor.

An experimental method proposed by Dykman [37], in which information on all trajectories in close neighborhood of the chosen state $\mathbf{q}_f$ outside the attractor is gathered, can be used to study the dynamics of large deviations from the attractor and to determine the optimal escape path. In this experimental method, the behavior of the dynamical variables $\mathbf{q}(t)$ and the random force $\mathbf{f}(t)$ is tracked continuously until the system makes a transition from the attractor to the neighborhood of state $\mathbf{q}_f$. Escape trajectories $\mathbf{q}^{\text{esc}}(t)$ of the required duration and noise realizations $\mathbf{f}^{\text{esc}}(t)$ of the same duration are then conserved; subsequently, the system forcefully returns to the basin of attraction of the attractor, with the initial conditions in the basin being chosen randomly. This is how an ensemble of trajectories is collected and how the fluctuational prehistory probability distribution $p_h(\mathbf{q}, t; \mathbf{q}_f, t_f)$ is constructed for the time interval during which the system is observed. This distribution contains all the information on the temporal evolution of the system immediately before the trajectory arrives at $\mathbf{q}_f$. The existence of an optimal path of escape from the attractor is diagnosed from the form of p_h: if there is no optimal escape trajectory, then p_h at a given time has a sharp peak at point $\mathbf{q}_{\text{opt}}(t; t_f, \mathbf{q}_f)$. Thus, when experimentally studying p_h, it is possible to find a region in the phase space of the system within which the optimal path is clearly seen, i.e., the region with a distinct, narrow (of the order of $\sqrt{D}$ [37, 52]) peak of the distribution. In this case, the optimal fluctuational force that moves the system trajectory outside the attractor along the optimal path can be estimated by averaging the noise realizations $\mathbf{f}^{\text{esc}}(t)$ over the ensemble. Note that investigating the fluctuation prehistory also allows us to verify the concept of optimal fluctuations and to determine the ranges of system parameters for which optimal paths exist. We used this experimental approach to study the escape from a quasi-hyperbolic Lorenz attractor.

By studying large noise-induced deviations, we can judge the system stability against noise perturbations and can develop ways of controlling the system dynamics in the absence and presence of noise [20, 48, 55]. Thus, for example, it was shown in [20, 55] that the problem of deterministic optimal (from an energetic point of view) control of the system transition from a chaotic state to a regular state can be solved by determining the optimal escape paths and the optimal fluctuational force. It was also shown in these papers that an experimental determination of the optimal trajectories based on the theory of large fluctuations is currently the only approach to solving the deterministic problem of optimal control of transitions between attractors in a chaotic system. The investigation of large fluctuations is therefore of importance in practical applications.

The Hamiltonian formalism presented above makes it possible to theoretically describe the fluctuational transitions in nonequilibrium systems. However, direct application of this approach to chaotic systems involves several fundamental difficulties. These difficulties are primarily related to the uniqueness of the solution, to the uncertainty of the boundary condition $\mathbf{q}_i$ on a chaotic attractor, and to the determination of the probability measure on the attractor itself. In general, when the chaotic attractor has fractal boundaries of its domains

of attraction, the problem of determining the second boundary condition $\mathbf{q}_f$ arises.

As yet no theoretical procedure has been developed to calculate the probability of fluctuational escape from the basin of attraction of a chaotic attractor. Previously, however, (see, e.g., [14, 45, 53, 54] it was shown that the Hamiltonian formalism can be generalized to chaotic systems. In this case, studying the uniqueness of the solution and boundary conditions becomes of current interest and importance. One possible approach to their solution is the experimental method described above. It allows us to find an approximate solution to the problem of fluctuational transitions in systems with chaotic attractors by statistically analyzing direct observations of the transition dynamics. We demonstrated the efficiency of statistical analysis of the experimentally measured fluctuation prehistories in [1, 20] in relation to the problem of escape from a chaotic attractor through a nonfractal boundary for a periodically driven nonlinear oscillator. Below, we present the results of our study of fluctuational transitions in the Lorenz system.

4. AN EXPERIMENTAL STUDY OF THE FLUCTUATION PREHISTORY

Let us consider fluctuational escape from the chaotic attractor of system (1) by using the experimental approach described above. Before analyzing the results of our studies, we note two fundamental points.

(1) The method described above allows the trajectory corresponding to a global minimum of the escape energy to be determined if the relaxation time to an equilibrium distribution on the attractor, t_{rel}, is much shorter than the time of the system fluctuational escape from the basin of attraction of a given attractor, t_{esc}: $t_{rel} \ll t_{esc}$. The fluctuational escape time exponentially increases with decreasing noise intensity D as $t_{esc} \propto \exp(S/D)$, where S is the escape "energy". Since D is always finite in practice (because the observing times are necessarily finite), the satisfaction of the condition $t_{rel} \ll t_{esc}$ is primarily related to t_{rel}, i.e., to the properties of the equilibrium distribution on the attractor.

It was shown in [1, 20, 55] that, for a chaotic attractor in a periodically driven nonlinear oscillator, the noise intensity can be chosen in such a way that the condition $t_{rel} \ll t_{esc}$ is satisfied. In this case, the optimal trajectory found does not depend on the initial conditions on the attractor or on the noise intensity.

For the Lorenz system, the situation is different. As was pointed out above, the probability of the system being in the neighborhood of the manifolds of the saddle point O is exponentially small. It may be said that the equilibrium distribution function on the attractor has tails similar to the tails of a Gaussian distribution and the trajectory falls on these tails rarely. Moreover, our reasoning that the trajectory must come arbitrarily close to the saddle-point manifolds when $t \longrightarrow \infty$ is based on a theoretical analysis of the Lorenz system, while numerical studies show that this probability is zero for sufficiently long but finite computational times (of the order of two weeks). This implies that the relaxation time to an equilibrium distribution on a quasi-hyperbolic chaotic attractor is much longer than the realistically feasible observing times. Consequently, for the Lorenz system, we have to investigate the dynamics of fluctuational escape within a finite time interval. In this case, the question of how the solution obtained depends on the initial conditions on the attractor and on the noise intensity for the Lorenz attractor is still an open question (cf. the studies of nonequilibrium trajectories [56, 57]). However, the experimentally found escape scenario (see below) suggests that a decrease in noise intensity cannot result in qualitative changes of the escape trajectory.

(2) The prehistory probability distribution was initially derived in [37] to investigate the motion of trajectories far from the attractor. The results presented in Section 2 show that the boundaries of the Lorenz attractor (separatrices Γ_1 and Γ_2) come very close to the boundaries of the basin of attraction; the attractor has a highly nonuniform distribution, and no motion along some parts of the attractor is observed in the absence of noise. Therefore, we extend the approach based on an experimental study of the fluctuational prehistory [37] to investigate the fluctuational deviations within the Lorenz attractor.

We investigated the fluctuational escape by numerically solving the stochastic differential equations of system (1) using the Heun method (equivalent to the Runge–Kutta method) [60] and a high-speed pseudorandom sequence generator [61]. The noise intensity was $D = 0.01$. For definiteness, we consider the transition from the chaotic attractor to state P_1 (see Fig. 3). Since the system is symmetric, the mechanism of the transition to state P_2 is the same and the quantitative characteristics of the transition can be obtained by changing the variables (q_1, q_2, q_3) to $(-q_1, -q_2, q_3)$. According to the experimental approach described above, we gathered the ensemble of escape trajectories $\{(q_1(t), q_2(t), q_3(t))_i\}$, $i = 1, N$ and the corresponding ensemble of noise realizations $\{(\xi(t))_i\}$, $i = 1, N$; here, i is the realization number and N is the number of realizations (in our studies, $N \approx 1000$).

Figure 5a shows several numerically simulated escape trajectories. We see that two narrow bundles of trajectories come close to the saddle point O from the neighborhood of points S and then slightly diverge and fall into the neighborhood of the saddle cycle L_1 while moving along separatrix Γ_2 (see also Fig. 3). The two bundles of trajectories from points S to point A exist, because the Lorenz system is symmetric, suggesting the only path along which the trajectories move in the neighborhood of point A. Note that the trajectories also approach the neighborhood of point S by narrow bun-

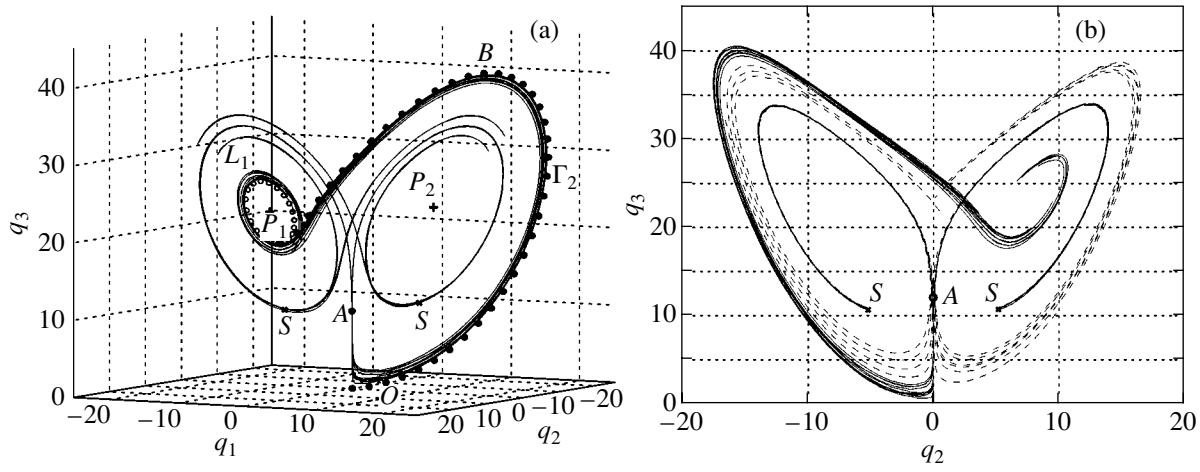

Fig. 5. (a) Ten trajectories of escape from a chaotic attractor in state P_1 obtained by numerically simulating system (1). The filled circles indicate separatrix *2*; the saddle cycle L_1 is indicated by open circles. (b) Ten trajectories of escape from a chaotic attractor (solid lines) and ten deterministic trajectories (dashed lines) with the initial conditions coincident with the coordinates of the escape trajectories near points *S*.

dles, suggesting that the behavior of the trajectories for the Lorenz system is predictable on long time scales.

The next step must be a statistical analysis of the ensemble of realizations: constructing the prehistory probability distribution $p_h(q_1, q_2, q_3, t)$ and the ensemble-averaged fluctuational force, which is an estimate of the momentum $p_3(t)$ and, accordingly, an estimate of the action $S[\mathbf{q}]$ for system (6) [20]. This requires matching the ensembles of realizations at some characteristic point. Strictly speaking, this characteristic point must be a point near the boundaries of the basins of attraction. However, because of the slow diffusion of the system near the attractor boundaries, an averaging over the initial conditions takes place (cf. [58, 59]), the fluctuational trajectories cross the basin boundary in a fairly wide region, and one point cannot be fixed in the neighborhood of the basin. In [1, 20], we chose a characteristic segment of the escape trajectory instead of a characteristic point. Based on the pattern of escape shown in Fig. 5a, we can choose a segment near point *A*, where the bundle of trajectories is narrow, as the characteristic segment. The statistical characteristics constructed for this characteristic segment are shown in Figs. 6a, 7a, and 7c. The prehistory probability distribution has a distinct and unique peak in the time interval $50 < t/h < 250$ (the region between the dashes in Fig. 6a). In this interval, the fluctuational force p_3 is zero (Fig. 7a) and the dispersion D_p of the distribution decreases to a minimum (Fig. 7c). Subsequently, the distribution peak spreads, the fluctuational force becomes nonzero, and the dispersion increases sharply. We see (Fig. 6a) that after passing point *A*, the escape trajectories diverge and there is no longer any distinct peak in the distribution $p_h(q_3, t)$. All trajectories after point *A* fall in the neighborhood of the saddle cycle L_1 but the times of system motion along different trajectories in this segment are different. This behavior of the escape trajectories is determined by the stable manifold of the saddle point *O*: the time of motion to point *O* along the stable manifold itself tends to infinity. If the trajectory runs alongside the manifold, then the time of motion along it is finite; the further the point is from the manifold, the larger is this time. For this reason, the trajectories near the saddle point *O* are therefore very sensitive to fluctuations; the escape trajectories traverse the segment from point *A* to cycle L_1 in different times, and, therefore, the peak of the fluctuation prehistory probability distribution spreads. This implies that choosing the neighborhood of point *A* as the characteristic segment allows us to analyze the behavior of the escape trajectories only up to point *A*. For the subsequent analysis, we must choose a different characteristic segment located after the saddle point *O*, for example, the neighborhood of point *B* (see Fig. 5a). Since the q_3 coordinate reaches its maximum in this segment, which can be easily diagnosed by the time series of the system, it will suffice to bring the escape trajectories into coincidence to construct the distribution $p_h(q_1, q_2, q_3, t)$. The statistical characteristics for point *B* are shown in Figs. 6b, 7b, and 7d. We see that the situation is opposite for point *B*: a distinct peak of the distribution $p_h(q_3, t)$ is observed at $t/h > 230$ (the region to the right of the dotted line in Fig. 6b), at which the fluctuational force p_3 is nonzero (Fig. 7b) and the dispersion D_p of the distribution reaches a minimum (Fig. 7d). Consequently, there is only one optimal escape trajectory along which the system moves from point *O* into the neighborhood of the saddle cycle L_1.

Our analysis of the prehistory probability and fluctuational force shown in Figs. 6 and 7 suggests the following scenario for the escape from a quasi-hyperbolic attractor. The escape trajectory can be broken down

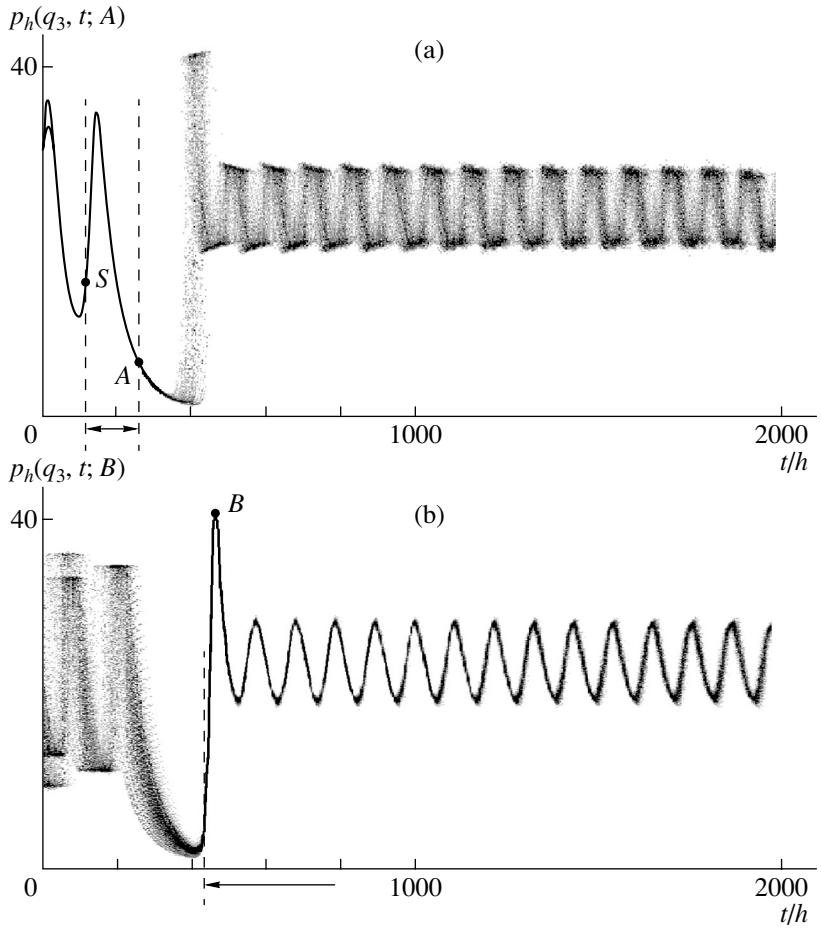

Fig. 6. Projections of the escape prehistory probability distributions onto the q_3 coordinate axis for the characteristic points (a) *A* and (b) *B*. The horizontal axes are normalized to the integration step $h = 2\pi/200$. The zero times for distributions (a) and (b) are different.

into three segments (Fig. 8). Segment 1 corresponds to the motion within the attractor, which requires no noise. The escape trajectory starts at some characteristic point of the chaotic attractor (point *S* in Fig. 5a). This point belongs to a typical attractor trajectory. The probability of the system trajectory being in a close neighborhood of point *S* in the absence of noise is much larger than zero. After point *S*, the escape trajectory falls into the neighborhood of the stable manifold (point *A* in Fig. 5a) of the saddle point *O*, where the dispersion of the prehistory probability distribution is at a minimum (Fig. 7c). At this time, the fluctuational force is close to zero (Fig. 7a). The motion in the segment from point *S* to point *A* is completely deterministic, and noise does not change qualitatively the trajectory behavior [10, 11]. Indeed, if the noise source is removed at the time when the trajectories pass point *S*, then the trajectories will fall into the neighborhood of point *A* in a deterministic way; the bundle will remain narrow (see Fig. 5b). However, after passing point *A*, the trajectories in the absence of noise are located far from the saddle point *O* and separatrices Γ_1 and Γ_2.

Subsequently, under the action of the fluctuational force, the escape trajectory approaches the saddle point *O* along the stable manifold of point *O* and then recedes from it along separatrix Γ_2 and falls into the neighborhood of the saddle cycle L_1. The trajectory motion from the neighborhood of the stable manifold of the saddle point *O* to the saddle cycle L_1 corresponds to segment 2 (Fig. 8; see also Fig. 1). In this part of the path, the fluctuational force follows the trajectory of the system q_3 coordinate; the force and the system trajectory may be said to be in phase. The escape trajectories in the second segment form a broad bundle (Figs. 5a and 6a) around some most probable escape path, which can be understood as follows. A narrow bundle of trajectories comes to point *A*, but this bundle has a finite width. As was already pointed out above, there is a large sensitivity to small deviations in the neighborhood of the manifolds of point *O* and the trajectories therefore diverge in a wide fan in the absence of a certain fluctuational force (see Fig. 5b). Under the action of fluctuations, however, the bundle width significantly decreases, but it remains much larger than the bundle size in the

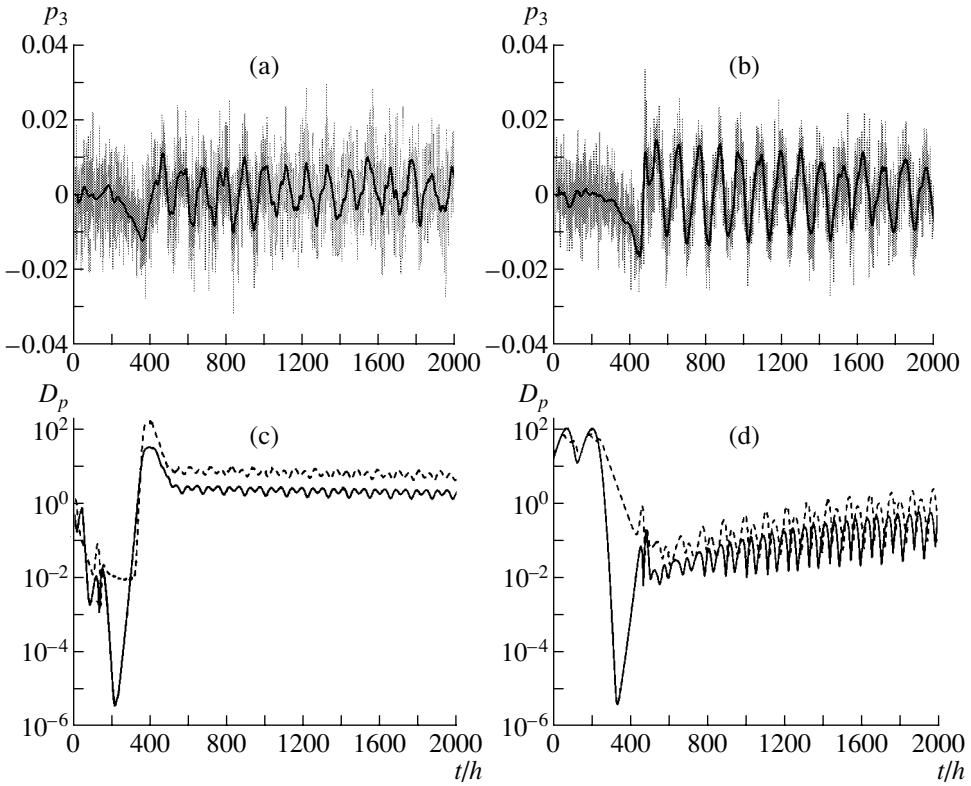

Fig. 7. Averaged fluctuational forces (a) $p_3(t; A)$ and (b) $p_3(t; B)$. The heavy line indicates the fluctuational forces passed through a low-noise filter with a zero-phase characteristic. (c, d) The dispersions of $p_h(q_1, t)$ (solid curve) and $p_h(q_3, t)$ (dashed curve) for the characteristic points A (c) and B (d). Since the dispersion of $p_h(q_2, t)$ behaves like the dispersion of $p_h(q_1, t)$, it is not shown in the figure. The horizontal axes in all figures are normalized to the integration step $h = 2\pi/200$. The zero times for panels 7a, 7c, 7b, and 7d are different.

neighborhood of point A. Thus, the fluctuational force begins to act after the trajectories passed point A (segment 1).

In segment 3, the escape trajectory moves along the unstable manifold of the saddle cycle L_1 and crosses cycle L_1. The escape trajectory and the fluctuational force are in antiphase, while the force itself shows a distinct oscillatory behavior, with its period being close to the system oscillation period (Fig. 7d). We also see that the fluctuational force does not vanish for a long time. The oscillatory behavior and long duration of the fluctuational force can be explained as follows. In the absence of fluctuations, the system trajectories go away from the saddle cycle along the unstable manifold very slowly, because the corresponding multiplicator is only slightly larger than unity. In this case, the fluctuational force can produce weak pushes during a long period toward the saddle cycle against the direction of trajectory motion along the unstable manifold. After the trajectory crosses the cycle, it also slowly relaxes to the equilibrium position P_1.

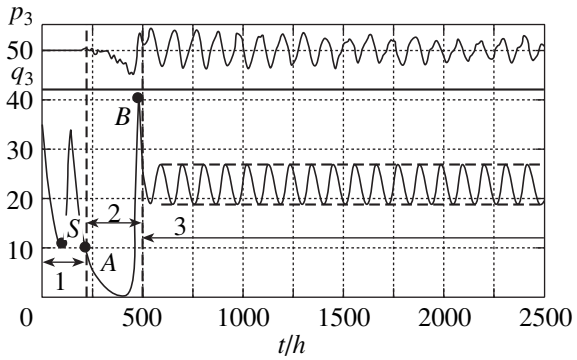

Fig. 8. The typical escape trajectory obtained by numerically simulating system (1). The horizontal dash–dotted lines indicate the boundaries of cycle L_1. The horizontal axis is normalized to the integration step $h = 2\pi/200$. The averaged fluctuational force is shown at the top.

The long duration of segment 3 in the escape trajectory leads us to conclude that there is no close relationship between this segment and the two preceding ones. Therefore, the escape process may be broken down into two independent stages: the trajectory expulsion into the neighborhood of the saddle cycle L_1 and its crossing.

5. CONCLUSION

Thus, the escape trajectory is unique and lies on the manifolds of system (1) that belong to the chaotic Lorenz attractor. The role of fluctuations is, first, to deliver the system trajectory to a rarely visited region of the attractor and, second, to induce the saddle-cycle crossing—the escape trajectory lies on the unstable manifold of the saddle cycle. Our results show that the most probable escape trajectory may be broken down into three segments (Fig. 8). The motion in the first segment is completely deterministic, while the trajectory path in the second and third segments is determined by a nonzero fluctuational force. Therefore, the probability of escape from a chaotic attractor can be calculated as a product of the probabilities of successive transitions, ρ_1, ρ_2, and ρ_3, where ρ_1 is the probability of the chaotic trajectory falling into the neighborhood of point S in the absence of noise, ρ_2 is the probability of the trajectory expulsion from the neighborhood of point A into the neighborhood of the saddle cycle L_1, and ρ_3 is the probability of crossing the saddle cycle L_1.

The probabilities ρ_1 and ρ_2 can be theoretically estimated by constructing the probability measure for a quasi-hyperbolic attractor using the technique from [11] but with allowance made for the finite noise intensity. Note that only the trajectories that are in a close neighborhood of cycle L_1 have a finite probability of escaping from the Lorenz attractor. Therefore, estimating the probability ρ_2 is related to solving the problem of a trajectory falling within a specified region when moving along the flux of a vector field in a finite time interval. The probability ρ_1 can be estimated with a high accuracy by using the attractor probability measure in the absence of noise, because noise in the first segment induces only diffusive motion around a deterministic trajectory (see Fig. 5a).

The probability ρ_3 is determined by the fluctuational dynamics on the two-dimensional unstable manifold of the saddle cycle L_1. Estimating ρ_3 is related to solving a problem similar to the problem of the fluctuational crossing of an unstable cycle into the domain of attraction of a stationary point in a Van der Pol oscillator with hard excitation. The trajectory motion near the saddle cycle (see Fig. 8) resembles periodic oscillations of slowly varying amplitude. In the absence of fluctuations, the oscillation amplitude increases and the role of fluctuations is to overcome the slow increase in amplitude and to cross the saddle cycle.

As we see, estimating each of the probabilities (ρ_1, ρ_2, ρ_3) is a separate serious problem that requires additional theoretical and experimental study.

The roles of noise in each of the escape segments are different. As was already mentioned above, in the absence of fluctuations, the probability of the trajectory falling into a close neighborhood of the separatrices and, hence, the saddle cycle is small but nonzero; i.e., the first and second segments can also be theoretically observed in a noiseless system: this is the system motion along trajectories T_1 and T_2 (see Fig. 2). In the presence of noise, the motion must take place in the tube formed around these trajectories. As our studies show, in the first segment (Fig. 8), the escape trajectories actually run in a close neighborhood of the trajectories T_1 and T_2. In the second segment, after point A (see Figs. 2 and 5), we observe the following: since the noisy escape trajectories differ from T_1 and T_2, the escape trajectories tend to go away from them, but, under the action of fluctuational force, they remain in the neighborhood of the stable manifold W^s and separatrix Γ_2. In other words, fluctuations hold the trajectory near the manifolds of the saddle point O. In the third segment, the escape from the basin of attraction of a chaotic attractor takes place precisely under the action of fluctuations.

As was pointed out above, an experimental approach based on investigation of the fluctuation prehistory was used in [1, 20] to study the fluctuational escape from a nonhyperbolic attractor in a periodically driven nonlinear oscillator. In these papers, the problem was solved in terms of the Hamiltonian formalism without reformulating the boundary conditions, which we did for a quasi-hyperbolic chaotic attractor by introducing a finite transition time. In [1, 20], we showed that the escape from a nonhyperbolic attractor is accomplished through successive transitions between saddle cycles and that no additional force is required to deliver the trajectory to the cycle from which the escape begins. For the quasi-hyperbolic attractor considered here, the situation is fundamentally different: the saddle cycles embedded in the attractor are not involved in the escape, and an additional fluctuational force is required to move the chaotic trajectory into a rarely visited region of the attractor. Thus, the dynamics of fluctuational escape depends significantly on the type of chaotic attractor.

Our analysis shows that the behavior of the trajectory of escape from the basin of attraction of a quasi-hyperbolic chaotic attractor is completely determined by the manifolds of the saddle point O at the coordinate origin and by the unstable manifold of the saddle cycle L_1. Therefore, the problem of escape from a chaotic attractor may be reformulated as the problem of trajectory motion along one- or two-dimensional manifolds in the presence of fluctuations. In other words, to get a more detailed picture of the escape from the attractor of a three-dimensional system, we must consider a similar problem in systems of lower dimension.

In Section 3, we noted that investigating large deviations from a chaotic attractor is important from a practical point of view and is related to the problem of attractor stability in the presence of fluctuations and to solving the problem of controlling the system dynamics. Our results, from the viewpoint of their applications to actual systems described by the Lorenz model, imply the following. First, the chaotic regime to which

a quasi-hyperbolic attractor corresponds is stable against fluctuations. Second, the noise-induced deviations from the chaotic regime of the system, i.e., large fluctuations, are extremely rare events, because these are related to the trajectory visit to exponentially unlikely regions of phase space. Third, the control of the system dynamics to switch it from the chaotic regime is implemented by a perturbation of a certain shape (see Fig. 8) in the time interval when the system is in a close neighborhood of the equilibrium state *O*. In other words, to realize or suppress transitions from the chaotic regime to a different regime requires keeping track of the system behavior in a close neighborhood of the system variables. The latter considerably simplifies the control procedure.

In conclusion, we note once again that the noise-induced escape from a quasi-hyperbolic attractor was investigated in terms of the theory of large fluctuations by experimentally analyzing the fluctuation prehistory [37]; here, we extended the experimental approach [37] to study the motion along the attractor itself.

ACKNOWLEDGMENTS

We are grateful to M.I. Dykman and V.N. Smelyansky for attention and helpful discussions. This study was supported by the INTAS Foundation (grant no. 01-867) and in part by a grant from the US Foundation for Support of Civil Research and Basic Development for the Commonwealth of Independent States (no. REC-006).

REFERENCES

1. D. G. Luchinskiĭ and I. A. Khovanov, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 782 (1999) [JETP Lett. **69**, 825 (1999)].
2. E. N. Lorenz, J. Atmos. Sci. **20**, 130 (1963).
3. B. Salzman, J. Atmos. Sci. **19**, 329 (1962).
4. H. Haken, Phys. Lett. A **53A**, 77 (1975).
5. A. N. Oraevskiĭ and V. Yu. Toronov, Kvantovaya Élektron. (Moscow) **16**, 2063 (1989).
6. C. O. Weiss, U. Hubner, N. B. Abraham, and D. Tang, Infrared Phys. Technol. **36**, 489 (1995).
7. V. S. Anishchenko, T. E.Vadivasova, and V. V. Astakhov, in *Nonlinear Dynamics of Chaotic and Stochastic Systems. Fundamental Principles and Selected Problems*, Ed. by V. S. Anishchenko (Saratovsk. Univ., Saratov, 1999), p. 1.
8. C. Sparrow, *The Lorenz Equations: Bifurcations, Chaos and Strange Attractors* (Springer-Verlag, New York, 1982), p. 1.
9. J. Marsden and M. McCracken, *The Hopf Bifurcation and Its Applications* (Springer-Verlag, Berlin, 1976; Mir, Moscow, 1980).
10. Yu. I. Kifer, Usp. Mat. Nauk **29**, 205 (1974).
11. R. Graham, in *Noise in Nonlinear Dynamical Systems*, Ed. by P. V. E. McClintock and F. Moss (Cambridge Univ. Press, Cambridge, 1989), Vol. 1, p. 225.
12. C. O. Weiss and W. Klische, Opt. Commun. **51**, 47 (1984); Phys. Rev. Lett. **57**, 2804 (1986).
13. V. S. Anishchenko and A. B. Neĭman, Zh. Tekh. Fiz. **60** (1), 3 (1990) [Sov. Phys. Tech. Phys. **35**, 1 (1990)].
14. R. Kautz, Phys. Lett. A **125**, 315 (1987).
15. R. L. Kautz, Rep. Prog. Phys. **59**, 935 (1996).
16. R. D. Astumian and I. Derenyi, Eur. Biophys. J. **27**, 474 (1998).
17. E. Arimondo, D. Hennequin, and P. Glorieux, in *Noise in Nonlinear Dynamical Systems*, Ed. by P. V. E. McClintock and F. Moss (Cambridge Univ. Press, Cambridge, 1989), Vol. 3, p. 119.
18. E. A. Jackson, Chaos **7**, 550 (1997).
19. B. Hubinger, R. Doerner, and W. Martienssen, Phys. Rev. E **50**, 932 (1994).
20. I. A. Khovanov, D. G. Luchinsky, R. Mannella, and P. V. E. McClintock, Phys. Rev. Lett. **85**, 2100 (2000).
21. A. J. Bray and A. J. McKane, Phys. Rev. Lett. **62**, 493 (1989).
22. T. J. Newman, A. J. Bray, and A. J. McKane, J. Stat. Phys. **59**, 357 (1990).
23. M. I. Dykman, Phys. Rev. A **42**, 2020 (1990).
24. M. I. Dykman and K. Linderberg, in *Contemporary Problems in Statistical Physics*, Ed. by G. H. Weiss (SIAM, Philadelphia, 1994), p. 41.
25. M. I. Dykman and V. N. Smelyanskiy, Superlattices Microstruct. **23**, 495 (1998).
26. R. S. Maier and D. L. Stein, Phys. Rev. Lett. **71**, 1783 (1993).
27. R. S. Maier and D. L. Stein, Phys. Rev. E **48**, 931 (1993).
28. M. I. Dykman *et al.*, Phys. Rev. E **49**, 1198 (1994).
29. R. S. Maier and D. L. Stein, J. Stat. Phys. **83**, 291 (1996).
30. M. I. Dykman and M. A. Krivoglaz, Zh. Éksp. Teor. Fiz. **77**, 60 (1979) [Sov. Phys. JETP **50**, 30 (1979)].
31. M. I. Dykman, H. Rabitz, V. N. Smelyanskiy, and B. E. Vugmeister, Phys. Rev. Lett. **79**, 1178 (1997).
32. V. N. Smelyanskiy, M. I. Dykman, H. Rabitz, and B. E. Vugmeister, Phys. Rev. Lett. **79**, 3113 (1997).
33. R. S. Maier and D. I. Stein, SIAM J. Appl. Math. **57**, 752 (1997).
34. V. N. Smelyanskiy, M. I. Dykman, and R. S. Maier, Phys. Rev. E **55**, 2369 (1997).
35. A. D. Ventsel' and M. I. Freĭdlin, *Fluctuations in Dynamic Systems under Random Perturbation Action* (Nauka, Moscow, 1979), p. 1.
36. L. Onsager and S. Machlup, Phys. Rev. **91**, 1505 (1953).
37. M. I. Dykman *et al.*, Phys. Rev. Lett. **68**, 2718 (1992).
38. M. I. Dykman *et al.*, Phys. Rev. Lett. **77**, 5229 (1996).
39. D. G. Luchinsky *et al.*, Phys. Rev. Lett. **79**, 3109 (1997).
40. D. G. Luchinsky, J. Phys. A **30**, L577 (1997).
41. D. G. Luchinsky and P. V. E. McClintock, Nature **389**, 463 (1997).
42. M. E. Johnson, M. S. Jolly, and I. G. Kevrekidis, Numer. Algorithms **14**, 125 (1997).
43. R. Graham and T. Tel, Phys. Rev. Lett. **52**, 9 (1984).
44. L. S. Schulman, Physica A (Amsterdam) **177**, 373 (1991).
45. R. Graham, A. Hamm, and T. Tel, Phys. Rev. Lett. **66**, 3089 (1991).

46. D. G. Luchinsky, P. V. E. McClintock, and M. I. Dykman, Rep. Prog. Phys. **61**, 889 (1998).
47. H. Haken, Rev. Mod. Phys. **47**, 67 (1975).
48. V. N. Smelyanskiy and M. I. Dykman, Phys. Rev. E **55**, 2516 (1997).
49. A. D. Ventsel' and M. I. Freĭdlin, Usp. Mat. Nauk **25**, 1 (1970).
50. D. Ludwig, SIAM Rev. **17**, 605 (1975).
51. V. A. Chinarov, M. I. Dykman, and V. N. Smelyanskiy, Phys. Rev. E **47**, 2448 (1993).
52. J. Hales, A. Zhukov, R. Roy, and M. I. Dykman, Phys. Rev. Lett. **85**, 78 (2000).
53. P. D. Beale, Phys. Rev. A **40**, 3998 (1989).
54. P. Grassberger, J. Phys. A **22**, 3283 (1989).
55. D. G. Luchinsky, I. A. Khovanov, S. Berri, *et al.*, Int. J. Bifurcation Chaos Appl. Sci. Eng. (2002) (in press).
56. S. M. Soskin, J. Stat. Phys. **97**, 609 (1999).
57. B. E. Vugmeister, J. Botina, and H. Rabitz, Phys. Rev. E **55**, 5338 (1997).
58. V. N. Smelyanskiy *et al.*, J. Chem. Phys. **110**, 11488 (1999).
59. V. N. Smelyanskiy, M. I. Dykman, and B. Golding, Phys. Rev. Lett. **82**, 3193 (1999).
60. R. Mannella, in *Supercomputation in Nonlinear and Disordered Systems*, Ed. by L. Vázquez, F. Tirando, and I. Martin (World Sci., Singapore, 1997), p. 100.
61. G. Marsaglia and W.-W. Tsang, SIAM J. Sci. Stat. Comput. **5**, 349 (1984).

Translated by V. Astakhov

Journal of Experimental and Theoretical Physics, Vol. 94, No. 4, 2002, pp. 834–851.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 121, No. 4, 2002, pp. 971–990.
Original Russian Text Copyright © 2002 by Mazur, Geogdzhaev, Gurevich, Zybin.

NONLINEAR PHYSICS

A Statistical Limit in the Solution of the Nonlinear Schrödinger Equation under Deterministic Initial Conditions

N. G. Mazur[a], V. V. Geogdzhaev[b], A. V. Gurevich[c, *], and K. P. Zybin[c, **]

[a]*Schmidt Joint Institute of Physics of the Earth, Russian Academy of Sciences, ul. Bol'shaya Gruzinskaya 10, Moscow, 123995 Russia*

[b]*Shirshov Institute of Oceanology, Russian Academy of Sciences, ul. Krasikova 23, Moscow, 117851 Russia*

[c]*Lebedev Physical Institute, Russian Academy of Sciences, Leninskii pr. 53, Moscow, 119991 Russia*

**e-mail: alex@lpi.ru*

***e-mail: zybin@lpi.ru*

Received July 25, 2001

Abstract—We investigate the semiclassical limit for the nonlinear Schrödinger equation in the case of a defocusing medium under oscillating nonperiodic initial conditions specified on the entire x axis. We formulate a system of integral conservation laws in terms of an infinite number of spatially averaged densities explicitly calculated from the initial conditions. We study the direct scattering problem and show that the scattering phase is a uniformly distributed random variable. The evolution of this system leads to the development of nonlinear oscillations, which become statistical in nature on long time scales. A modified inverse scattering method based on constructing a maximizer of the N-soliton solution in the continuum limit for $N \longrightarrow \infty$ is used to obtain an asymptotic solution. Using the maximizer, we found an infinite set of conserved averaged densities in the statistical state. This allowed us to couple the initial state with the limiting statistical steady (for $t \longrightarrow \infty$) state and, thus, to unambiguously determine the level spectrum in the statistical limit. © *2002 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

We consider the nonlinear Schrödinger equation (NSE) with defocusing,

$$iu_t + u_{xx} - 2|u|^2 u = 0. \quad (1)$$

This equation is known to be completely integrable. Its solution for the classical problem with a rapidly decreasing initial potential $u(x, 0)$,

$$\int_{-\infty}^{+\infty} |u(x, 0)|\,dx < \infty, \quad (2)$$

was first obtained by Zakharov and Shabat [1]. Condition (2) imposed on the initial function $u(x, 0)$ is of fundamental importance in the theory of integration by the inverse scattering method [2]. It ensures that the initial potential $u(x, 0)$ is analytic for $|x| \longrightarrow \infty$. This makes it possible to obtain a complete solution of the scattering problem, i.e., to accurately determine the system of initial phases. The phase shift in time is described by a simple ordinary differential equation, and the transformation to $u(x, t)$ is described by the linear inverse scattering problem (see [2], p. 36). This procedure forms the basis for the classical inverse scattering method. It allows one to solve the Cauchy problem for a nonlinear equation (e.g., the NSE), i.e., to unambiguously determine the strictly deterministic value of $u(x)$ at any instant in time t for a given $u(x, 0)$.

Our statement of the problem differs fundamentally from the classical one. We consider the initial function

$$u(x, 0) = u_0(x), \quad (3)$$

where $u_0(x)$ is an oscillating, nonperiodic, large-scale, bounded function specified on the entire x axis. The properties of this function are discussed in more detail below (in Section 2), while here an important fact is that it does not satisfy condition (2). As a result, the function $u_0(x)$ loses its analyticity when $|x| \longrightarrow \infty$. Moreover, it has an essential singularity here: the value of $u_0(x)$ proves to be indeterminate; it is only known to vary within specified limits [because $u_0(x)$ is bounded]. Thus, the solution of the scattering problem becomes fundamentally new in nature: it leads to random scattering phases. As a result, the purely dynamic evolution of the function $u(x, t)$ for uniquely determined initial potential (3) loses dynamical properties in the asymptotic limit $t \longrightarrow \infty$ and tends to the statistical limit. In other words, in the course of time, the solution takes the form of nonlinear oscillations with random phases. Such a system is statistical in nature and can be described in terms of a continuous random process. More specifically, the strictly deterministic value of

$u(x, t)$ loses its meaning in the asymptotic limit $t \longrightarrow \infty$, and we can only speak about the function $f(u, x, t)$, the probability density to find a given value of u at point (x, t), or the correlation functions. The problem is to determine the specific form of these functions and their dependence on the initial potential $u_0(x)$.

For the NSE with defocusing, this problem is stated for the first time. Previously [3, 4], we considered such a problem for the Korteweg–de Vries equation. In [3], we used the method of Whitham multiphase equations to show the development of a statistical structure in the solution with time, which becomes completely statistical in nature in the asymptotic limit $t \longrightarrow \infty$. In [4], we developed as asymptotic method of solving the above problem for the Korteweg–de Vries equation, which was called a modified inverse scattering method. The initial function $u_0(x)$ is assumed to be spatially homogeneous (ergodic). In this case, the probability density f depends on the velocity u alone, $f = f(u)$, and the spatial correlation function $K = K(x, x')$ depends on the difference $|x - x'|$ alone. In [3, 4], we determined the form of these functions for a specific initial potential $u_0(x)$.

In this paper, we develop the modified inverse scattering method for the NSE with defocusing. Note that the NSE describes the propagation of electromagnetic waves in nonlinear media, for example, in the Earth's ionosphere [5] and optical fibers [6]. The Gross–Pitaevskii equation, which describes the behavior of excitations in a gas in the state of a Bose–Einstein condensate [7], is similar in form.

Our paper has the following structure. In Section 2, we briefly describe the salient features of the NSE and formulate the initial-value problem in general form. In Section 3, we investigate the conservation laws. The NSE is known to have an infinite set of differential conservation laws. In the classical statement of the problem, these can be represented in integral form as time-independent integrals,

$$J_n = \int_{-\infty}^{\infty} R_n dx, \quad n = 1, 2, \ldots,$$

where R_n are the polynomials of the function $u(x)$ and its derivatives [2]. Under condition (3), however, the initial function $u_0(x)$ does not rapidly decrease with $|x| \longrightarrow \infty$ and the integrals J_n diverge. In Section 3, we introduce the averaged integral conservation laws or, in other words, the conservation laws represented as the invariant densities $\langle R_n \rangle$. The latter take on fixed, finite values determined by the ergodic initial function $u_0(x)$ for each n. In contrast to [4], we used the multitime technique and the Hamiltonian formalism to derive them. The values of $\langle R_n \rangle$ are shown to be conserved as the solution develops with time. We emphasize that there is no unique relationship between the invariant densities and the spectral function in the NSE, in contrast to the Korteweg–de Vries equation.

In Section 4, we solve the scattering problem. The scattering phases for the initial potential $u_0(x)$ under consideration are shown to be distributed randomly and uniformly. On the one hand, this rules out the direct application of the classical inverse scattering method but, on the other hand, serves as the basis for using the modified inverse scattering method to determine the asymptotic behavior of the solution for $t \longrightarrow \infty$. To this end, based on the N-soliton solution, we construct a maximizer for the system of interacting solitons specified on the spatial scale L with uniformly distributed random phases in Section 5. We show that, for $N \longrightarrow \infty$ and $L \longrightarrow \infty$, there is a continuum limit to the phase-averaged maximizer described by a linear integral equation. It is important to note that this equation includes the level spectrum in phase space $\rho_s(\lambda)$. Here, the fundamental difference between our problem and the classical problem again becomes significant. In the classical problem, the solitons go to infinity in the asymptotic limit $t \longrightarrow \infty$ and become free. For free solitons, the level spectrum in phase space is identical to the level spectrum in ordinary space $\rho(\lambda)$. Since the spectrum does not depend on time t, the following relation always holds in the classical statement of the problem:

$$\rho_s(\lambda) \equiv \rho(\lambda). \tag{4}$$

In our problem, the solitons never become free: they remain interacting ones at any t, even for $t \longrightarrow \infty$. Here, an essential singularity at infinity again shows up. In this case, $\rho_s(\lambda)$, the level spectrum in phase space, does not match the level spectrum in ordinary space; i.e., in our case, there is no relation (4). In Section 6, we use the modified inverse scattering method to determine $\rho_s(\lambda)$. It consists in the following. Using the maximizer calculated in the continuum limit, we determine the averaged invariant densities in the statistical state, $\langle R_n \rangle_s$. Based on the multitime hierarchy of the NSE, we prove that the maximizer determines the entire infinite set of invariant densities. Since the invariant densities do not vary with time, they match $\langle R_n \rangle$ specified by the initial condition

$$\langle R_n \rangle_s = \langle R_n \rangle. \tag{5}$$

Thus, the unique relationship between the statistical state and the deterministic initial state can be established by using an infinite set of conservation laws (5). The one-to-one relationship between the spectra $\rho_s(\lambda)$ and $\rho(\lambda)$ can be derived from relations (5). After the steady-state spectrum $\rho_s(\lambda)$ has been determined, we can calculate all the necessary parameters in the statistical state by using the maximizer.

In contrast to the solution of the Korteweg–de Vries equation in [4], where we initially considered only two conservation laws, here, we managed to prove that all conservation laws hold in the statistical state by using the multitime hierarchy.

2. STATEMENT OF THE PROBLEM

The NSE with defocusing has the form (1). With the change of variables

$$u = \rho^{1/2}\exp(i\phi), \quad v = \phi_x, \quad \tau = 2t,$$

Eq. (1) transforms to an equivalent system of dispersion hydrodynamic equations,

$$\rho_\tau + (\rho v)_x = 0,$$

$$v_\tau + vv_x = -\rho_x + \frac{1}{2}\left[\frac{(\sqrt{\rho})_{xx}}{\sqrt{\rho}}\right]_x.$$

We study the case where the hydrodynamic variables vary smoothly: $\rho = \rho(\varepsilon x, \varepsilon\tau)$, $v = v(\varepsilon x, \varepsilon\tau)$, $\varepsilon \ll 1$. Expanding the scale, $x \longrightarrow x/\varepsilon$ and $t \longrightarrow t/\varepsilon$, we can then explicitly separate out the small parameter of the dispersion term with higher derivatives in the hydrodynamic equations:

$$\begin{aligned}\rho_\tau + (\rho v)_x &= 0,\\ v_\tau + vv_x &= -\rho_x + \frac{1}{2}\varepsilon^2\left[\frac{(\sqrt{\rho})_{xx}}{\sqrt{\rho}}\right]_x.\end{aligned} \tag{6}$$

The system of equations (6) describes, in particular, the propagation of electromagnetic waves in a nonlinear defocusing medium. In contrast to the Korteweg–de Vries equation, it is a complete wave equation. The dispersion relation for system (6) linearized relative to a steady hydrodynamic flow with $\rho = \rho_0$ and $v = U_0$ is

$$\omega - kU_0 = \pm k\rho_0^{1/2}\left(1 + \frac{k^2}{4\rho_0}\right)^{1/2}.$$

It thus follows that the NSE describes two oppositely propagating waves. Accordingly, there are also two oppositely traveling systems of stationary waves, in particular, solitons

$$\begin{aligned}u &= \frac{\rho_0 + (v_s/2 - ik)^2\exp(2kX)}{\rho_0^{1/2}(1 + \exp(2kX))},\\ X &= \frac{x \pm v_s t}{\varepsilon}.\end{aligned} \tag{7}$$

We emphasize that, as follows from (7), the soliton of the defocusing NSE is a "well" in the background density ρ_0:

$$\rho = \rho_0 - \frac{k^2}{\cosh^2(kX)}, \quad v_s = 2(\rho_0 - k^2)^{1/2}.$$

When $|x| \longrightarrow \infty$, $\rho \longrightarrow \rho_0$. At $\rho_0 = 0$, Eq. (1) has no solitons. Note that the soliton velocity decreases with increasing depth of the well proportionally to $(\rho_0 - k^2)^{1/2}$. The soliton stops at $k^2 = \rho_0$; in this case, its velocity is equal to the steady-flow velocity U_0. Note also that the dispersion in the NSE is positive; i.e., both the wave phase and the group velocities in the medium increase with wave number k.

Let us now discuss the statement of the initial-value problem. We study the solution to the nonlinear equation (1) in the semiclassical limit. This implies that the solution to system (6) is considered for $\varepsilon \longrightarrow 0$; the initial hydrodynamic perturbations $\rho(x, 0)$ and $v(x, 0)$ do not contain the scale ε and are bounded on the entire x axis together with all derivatives. In addition, the initial functions $\rho(x, 0) \geq 0$ and $v(x, 0)$ are assumed to be nonperiodic functions oscillating about some means.

We restrict our analysis to initial conditions in the form

$$\begin{aligned}\rho(x, 0) &= \rho_0 - \sum_{m=-\infty}^{\infty} f_m\left(\frac{x - x_m}{\theta_m}\right),\\ v(x, 0) &= \sum_{m=-\infty}^{\infty} v_m\left(\frac{x - x_{mm}}{\theta_m}\right),\end{aligned} \tag{8}$$

where $\rho_0 > 0$, $f_m(\xi)$, and $v_m(\xi)$ are finite functions equal to zero outside the interval $\alpha_m < \xi < \beta_m$ ($\alpha_m < 0$, $\beta_m > 0$, $\beta_m - \alpha_m = 1$). Inside this interval, $f_m(\xi)$ is positive and has one maximum at $\xi = 0$ that does not exceed ρ_0. The latter condition ensures that $\rho(x, 0)$ is nonnegative. The intervals $(x_m + \alpha_m\theta_m, x_m + \beta_m\theta_m)$ on the x axis, in which different terms in (8) are nonzero, do not overlap. For brevity, the separate term f_m, v_m, which represents a localized perturbation of the homogeneous $\rho = \rho_0$, $U_0 = 0$ background, is called a well. Consider, for simplicity, specific initial conditions in the form of a sequence of wells with the same shape and depth,

$$\begin{aligned}\rho(x, 0) &= \rho_0 - \sum_{m=-\infty}^{\infty} f\left(\frac{x - x_m}{\theta_m}\right),\\ v(x, 0) &= \sum_{m=-\infty}^{\infty} g\left(\frac{x - x_m}{\theta_m}\right).\end{aligned} \tag{9}$$

The initial conditions are nonperiodic if, for example, the difference $x_{m+1} - x_m = l$ in (9) is constant and the well width varies as

$$\theta_m = \theta[1 + q\sin(\sigma m)], \tag{10}$$

where σ/π is an irrational number and $q < 1$.

Note that the initial conditions (10) we consider are ergodic; i.e., these are described by spatially homogeneous functions, for example, by quasi-periodic or almost periodic ones (see, e.g., [8] for an exact definition of ergodic functions).

Apart from the form of $f(x)$ and $g(x)$ and the background density ρ_0, the initial conditions (9) can be char-

acterized by the mean distance $\bar{l}$ between the wells and the mean well depth $\bar{\theta}$. The ratio

$$\gamma = \bar{\theta}/\bar{l} \tag{11}$$

defines the well packing density, i.e., the mean hydrodynamic perturbation of the homogeneous background.

3. INTEGRAL CONSERVATION LAWS

The NSE is integrable [2, 9]. In an exact integration of the nonlinear equation (1) and its equivalent hydrodynamic system (6), an auxiliary linear problem is the scattering problem for the system of the second order [9]

$$\begin{aligned} \varepsilon\psi_x^{(1)} &= -\frac{i}{2}\lambda\psi^{(1)} + \bar{u}\psi^{(2)}, \\ \varepsilon\psi_x^{(2)} &= \frac{i}{2}\lambda\psi^{(2)} + u\psi^{(1)}. \end{aligned} \tag{12}$$

Here, λ is the spectral parameter of the problem, and $\bar{u}$ is the function complex conjugate to u.

3.1. Differential Conservation Laws

Because of its complete integrability, the NSE is known to have an infinite set of independent differential conservation laws

$$\partial_t P_n + \partial_x W_n = 0. \tag{13}$$

For the initial-value problem under consideration, as in the solution of the Korteweg–de Vries equation [4], the conservation laws must be rewritten in an averaged form. This will be done in the present section. The Hamiltonian representation and the multitime technique $t \longrightarrow t_m$, $m = 1, 2, \ldots$ are used (in contrast to [4]) to derive the averaged integral conservation laws.

The densities of the conserved quantities P_n in (13) can be determined by expanding the logarithm of the transmission coefficient for the linear system (12) in powers of $1/\lambda$ for $\lambda \longrightarrow \infty$. This allows a recurrence formula to be derived for P_n. Indeed, after eliminating $\psi^{(2)}$ and substituting

$$\psi^{(1)} = \exp\left[i\varepsilon^{-1}\left(-\frac{\lambda}{2}x + \int P dx\right)\right],$$

system (12) transforms to the Riccati equation

$$\lambda P = |u|^2 + P^2 - i\varepsilon\bar{u}\partial_x\left(\frac{P}{\bar{u}}\right).$$

For the expansion coefficients of the solution to this equation with $\lambda \longrightarrow \infty$

$$P = \sum_{n=1}^{\infty} P_n \lambda^{-n}, \tag{14}$$

we derive the recurrence system [9, pp. 37, 39]

$$\begin{aligned} P_{n+1} &= -i\varepsilon\bar{u}\partial_x\left(\frac{P_n}{\bar{u}}\right) + \sum_{m=1}^{n-1} P_m P_{n-m}, \\ n &= 1, 2, \ldots, \end{aligned} \tag{15}$$

with the initial recurrence condition being $P_1 = |u|^2$. The densities in the differential conservation laws (13) can be successively determined from system (15):

$$P_2 = -i\varepsilon\bar{u}u_x, \quad P_3 = -\varepsilon^2\bar{u}u_{xx} + |u|^4,$$

$$P_4 = i[\varepsilon^3\bar{u}u_{xxx} - \varepsilon|u|^2(4\bar{u}u_x + u\bar{u}_x)], \ldots.$$

These are the polynomials of $u(x)$ and $\bar{u}(x)$ and their derivatives.

Recurrence (15) gives a definite set of polynomial densities P_n, but, in general, these densities are determined ambiguously. Clearly, we can make any linear combination of these or add an arbitrary constant. Moreover, adding the total derivative of the polynomial of $u(x)$ and $\bar{u}(x)$ and their derivatives to P_n again leads to the density of the conserved quantity.

3.2. Integral Conservation Laws and the Hamiltonian Form

If the function $u(x)$ and all its derivatives rapidly converge to zero when $|x| \longrightarrow \infty$ so that the integrals

$$I_n = \int_{-\infty}^{\infty} P_n(x)dx \tag{16}$$

exist, then the time independence of these integrals follows from relations (13).

Equation (1) can be represented in the Hamiltonian form [9] by using (16). If the functionals of $u(x)$ defined by integrals (16) with $n = 3, 4, \ldots$ are taken as the Hamiltonians, then we derive the hierarchy of equations

$$\frac{\partial u}{\partial t} = -i\frac{\delta I_n}{\delta\bar{u}}. \tag{17}$$

We have Eq. (1) at $n = 3$ and the so-called higher NSEs for $n > 3$.

The initial condition (8) or (9) that we consider is a nonperiodic sequence of isolated localized perturbations (wells). Therefore, the total spectrum of problem (12) can be obtained in the semiclassical limit by a simple superposition of the spectra for individual wells. Accordingly, when studying the contribution of each well, we used the theory developed in [9] for Eq. (1) under the condition of a finite density: when $|x| \longrightarrow \pm\infty$,

$$|u(x, t)|^2 \longrightarrow \rho_0 \neq 0. \tag{18}$$

Subsequently, these contributions were combined (for details, see the next section). Under condition (18), it is

convenient to transform Eq. (1) [by adding the phase factor $u \longrightarrow u\exp(-2i\rho_0 t)$] to the form [9]

$$iu_t + u_{xx} - 2(|u|^2 - \rho_0)u = 0. \tag{19}$$

In this case, the hydrodynamic form (6) remains unchanged.

It is important to note that the Hamiltonian form (17) holds only when $u(x)$ rapidly decreases together with its derivatives, because only in this case are the functionals I_n determinate and do their variational derivatives rapidly decrease as $|x| \longrightarrow \infty$. For a finite density, to ensure the convergence of the conserved integrals, it will suffice to subtract the following constants [9, p. 72] from the densities $P_n(x)$:

$$P^{\infty}_{2k-1} = \frac{1}{2}(-1)^{k+1}\omega^{2k}b^+_k, \quad P^{\infty}_{2k} = 0, \quad k = 1, 2, \ldots,$$

where $\omega = 2\sqrt{\rho_0}$ and b^+_k are the coefficients defined by the expansion

$$(1 + x)^{1/2} = \sum_{k=0}^{\infty} b^+_k x^k.$$

However, apart from the existence of functionals, the integrals with infinite limits of

$$Q_n(x) = P_n(x) - P^{\infty}_n,$$

it is also necessary to ensure that their variational derivatives rapidly decrease with $|x| \longrightarrow \infty$ in order to represent Eq. (19) and the corresponding higher equations in the Hamiltonian form. As was shown in [9, pp. 73, 193], certain linear combinations $R_n(x)$ of the densities $Q_n(x)$ related to the asymptotic expansion of the logarithm of the transmission coefficient in negative powers of the variable $\sqrt{\lambda^2 - \omega^2}$ may be used instead of expansion (14).

Denoting

$$\lambda^{-1} = \xi, \quad (\lambda^2 - \omega^2)^{-1/2} = \eta,$$

we have

$$\xi(\eta) = \eta(1 + \omega^2\eta^2)^{-1/2}.$$

The generating function of (14), i.e.,

$$P(\xi) = \sum_{n=1}^{\infty} P_n\xi^n,$$

corresponds to an infinite series of the densities P_n; similarly, we define

$$Q(\xi) = \sum_{n=1}^{\infty} Q_n\xi^n.$$

According to [9], the densities R_n are specified by the generating function

$$\sum_{n=1}^{\infty} R_n\eta^n = R(\eta) = Q(\xi(\eta)). \tag{20}$$

The following expansion is used to establish the relationship between R_n and Q_n that follows from (20):

$$\xi^n = \eta^n \sum_{p=0}^{\infty} \beta_{np}(\omega)\eta^p, \tag{21}$$

where the coefficients β_{np} are known; these can be expressed in terms of binomial coefficients and even powers of ω, some of which are zero. As a result, we obtain the relation

$$R_n = \sum_{p=0}^{n-1} \beta_{n-p,\,p}(\omega)Q_{n-p}. \tag{22}$$

Substituting

$$Q_j = P_j - P^{\infty}_j$$

into this relation yields a formula that expresses R_n in terms of P_n. Given the explicit form of $\beta_{np}(\omega)$ and the constants P_n, it is more convenient to write this formula separately for odd and even numbers:

$$R_{2k+1}(x) = (2k-1)!!\left[\sum_{m=0}^{k} \frac{(-1)^m\omega^{2m}}{2^m m!(2k-1-2m)!!}\right.$$

$$\left. \times P_{2k+1-2m}(x) + \frac{(-1)^{k+1}\omega^{2k+2}}{2^{k+2}(k+1)!}\right], \tag{23}$$

$$R_{2k+2}(x) = k!\sum_{m=0}^{k} \frac{(-1)^m\omega^{2m}}{m!(k-m)!}P_{2k+2-2m}(x).$$

Clearly, the differential conservation laws

$$\partial_t R_n + \partial_x W^{(R)}_n = 0, \tag{24}$$

which follow from (13), hold for new combinations R_n.

If we now take the integrals

$$J_n = \int_{-\infty}^{\infty} R_n(x)dx \tag{25}$$

as the Hamiltonians, then we derive the hierarchy of NSEs for a finite density:

$$\frac{\partial u}{\partial t} = -i\frac{\delta J_n}{\delta \bar{u}}, \quad n = 3, 4, \ldots. \tag{26}$$

It begins with Eq. (19).

Below, we give the first several densities R_n calculated with the recurrence formula (15) and relations (23):

$$R_1 = |u|^2 - \rho_0, \quad R_2 = -i\varepsilon\bar{u}u_x,$$
$$R_3 = -\varepsilon^2\bar{u}u_{xx} + (|u|^2 - \rho_0)^2, \tag{27}$$
$$R_4 = i[\varepsilon^3\bar{u}u_{xxx} - 4\varepsilon(|u|^2 - \rho_0)\bar{u}u_x - \varepsilon|u|^2u\bar{u}_x].$$

The expressions of these quantities in terms of the hydrodynamic variables ρ and v are

$$R_1 = \rho - \rho_0, \quad R_2 = \rho v - \frac{1}{2}i\varepsilon\rho_x,$$
$$R_3 = \rho v^2 + (\rho_0 - \rho)^2 - i\varepsilon(\rho v)_x - \varepsilon^2\rho^{1/2}(\rho^{1/2})_{xx},$$
$$R_4 = \rho v^3 + (3\rho - 4\rho_0)\rho v \tag{28}$$
$$- i\varepsilon\left(\frac{3}{2}\rho v^2 + \frac{3}{2}\rho^2 - 2\rho_0\rho\right)_x + O(\varepsilon^2).$$

The densities of the conserved quantities

$$p = \frac{\varepsilon(u_x\bar{u} - u\bar{u}_x)}{2i} = \rho v = \mathrm{Re}R_2,$$
$$h = \varepsilon^2|u_x|^2 + (\rho_0 - |u|^2)^2 \tag{29}$$
$$= \rho v^2 + (\rho_0 - \rho)^2 + \frac{1}{4}\varepsilon^2\rho_x^2\rho^{-1} = \mathrm{Re}R_3 + \frac{1}{2}\varepsilon^2\rho_{xx}$$

are physically interpreted as the momentum and energy densities, respectively. As we see from (29) and (27), the differences $p - R_2$ and $h - R_3$ are the total derivatives.

3.3. Multitime Formalism

Note an important property of the hierarchy of NSEs: any finite number of Eqs. (26) may be considered as the overdetermined system

$$\frac{\partial u}{\partial t_m} = -i\frac{\delta J_{m+2}}{\delta\bar{u}}, \quad m = 1, 2, \ldots \tag{30}$$

for the function of several variables

$$u = u(x, t_1, t_2, \ldots). \tag{31}$$

The nontrivial property of the NSE attributable to its complete integrability is that fluxes (30) commute:

$$\frac{\partial^2 u}{\partial t_m\partial t_n} = \frac{\partial^2 u}{\partial t_n\partial t_m},$$

i.e., system (30) is compatible and actually defines the function of many "times" (31). Each of the functionals J_n with $n \geq 3$ is the Hamiltonian for the corresponding time t_{n-2}.

The conservation laws (24) also apply to all higher times:

$$\frac{\partial R_n}{\partial t_m} + \frac{\partial W_{nm}^{(R)}}{\partial x} = 0. \tag{32}$$

Relations (32) at $n = 1$ are of particular importance, because these are used in Section 6. In this case, the fluxes can be represented as

$$W_{1m}^{(R)} = \sum_{k=1}^{m+1} k\beta_{k+1,\,m+1-k}Q_k + \partial_x(\ldots), \tag{33}$$

where β_{np} are defined in (21) and $\partial_x(\ldots)$ denotes a total derivative. To derive formula (33), let us determine $\partial R_1/\partial t_m$ from the general rule of the Hamiltonian dynamics by using the Poisson brackets [9]:

$$\frac{\partial R_1}{\partial t_m} = \{J_{m+2}, R_1\} = i\left(u\frac{\delta J_{m+2}}{\delta u} - \bar{u}\frac{\delta J_{m+2}}{\delta\bar{u}}\right).$$

Hence, relation (22) yields

$$\frac{\partial R_1}{\partial t_m} = i\sum_{p=0}^{m+1}\beta_{m+2-p,\,p}\left(u\frac{\delta\tilde{I}_{m+2-p}}{\delta u} - \bar{u}\frac{\delta\tilde{I}_{m+2-p}}{\delta\bar{u}}\right), \tag{34}$$

where $\tilde{I}_n = \int_{-\infty}^{\infty} Q_n dx$ are the functionals defined for the functions $u(x)$ that satisfy the finite-density condition. Since the density Q_n differs from the density P_n only by the constant P_n^{∞}, we have

$$\frac{\delta\tilde{I}_n}{\delta u} = \frac{\delta I_n}{\delta u}, \quad \frac{\delta\tilde{I}_n}{\delta\bar{u}} = \frac{\delta I_n}{\delta\bar{u}},$$

which allows the combination of variational derivatives in (34) to be expressed in terms of P_n:

$$u\frac{\delta\tilde{I}_n}{\delta u} - \bar{u}\frac{\delta\tilde{I}_n}{\delta\bar{u}} = i\partial_x[(n-1)P_{n-1} + \partial_x(\ldots)].$$

Substituting Q_{n-1} for P_{n-1} under the derivative sign and using the derived relation in (34) yields expression (33).

3.4. Averaged Integral Conservation Laws

For our problem with the restricted ergodic initial conditions (9), it would be natural, as in [4], to pass to an integral form of the conservation laws in terms of the averaged densities of conserved quantities. Consider the mean quantities

$$\langle R_n\rangle = \lim_{L\to\infty}\frac{1}{L}\int_{-L/2}^{L/2} R_n(x)dx. \tag{35}$$

Since the functions $R_n(x)$ are bounded, limits (35) exist. In view of the local conservation laws (24),

$$\frac{d\langle R_n\rangle}{dt} = \lim_{L\to\infty}\frac{1}{L}\int_{-L/2}^{L/2}\partial_t R_n dx = -\lim_{L\to\infty}\frac{1}{L}W_n^{(R)}(x)\Big|_{-L/2}^{L/2},$$

and the fluxes $W_n^{(R)}(\pm L/2)$ through the segment ends are bounded; therefore, $d\langle R_n\rangle/dt = 0$. Consequently, the means (35) do not depend on time and are conserved quantities. This is the complete system of averaged integral conservation laws for our initial-value problem.

The invariant densities $\langle R_n\rangle$ can be easily determined at initial time $t = 0$ from the specified $f(x)$ and $g(x)$ and $\gamma = \bar{\theta}/\bar{l}$ by using the fact that the initial conditions (9) are semiclassical. Using expressions (28) and (29), we obtain

$$\langle R_1\rangle = -\gamma\int_{-\infty}^{\infty} f(x)dx,$$

$$\langle R_2\rangle = \langle p\rangle = \gamma\int_{-\infty}^{\infty}(\rho_0 - f(x))g(x)dx,$$

$$\langle R_3\rangle = \langle h\rangle = \gamma\int_{-\infty}^{\infty}[(\rho_0 - f(x))g^2(x) + f^2(x)]dx + O(\varepsilon^2).$$

Similarly, by expressing (23) in terms of the hydrodynamic variables ρ and v, we can calculate $\langle R_n\rangle$ in the initial state for any n. Below, we will see that the invariants $\langle R_n\rangle$, as in the case of the Korteweg–de Vries equation [4], are uniquely related to the level density of the scattering problem. These play an important role in determining the spectrum in a statistical state.

4. THE DIRECT SCATTERING PROBLEM IN THE SEMICLASSICAL LIMIT

Let us study the spectral problem for the linear system (12) in the semiclassical limit, when $\varepsilon \longrightarrow 0$. Let us first consider one localized initial perturbation of $\rho(x)$ and $v(x)$ to which

$$u = \sqrt{\rho}\exp\left(i\varepsilon^{-1}\int^x v\,dx\right)$$

corresponds and then investigate the more general initial condition (9) composed of a large (infinite, in the limit) number of such localized perturbations with nonoverlapping carriers.

For our "hydrodynamic" initial-value problem (9), it is convenient to pass from system (12) to an equation of the second order, which allows the standard formulas of the Wentzel–Kramers–Brillouin (WKB) approximation to be used. After eliminating $\psi^{(1)}$ and substituting $\psi^{(2)} = \psi\sqrt{u}$ to remove the first derivative, we derive the equation

$$\varepsilon^2\psi_{xx} + \tilde{r}(x, \lambda; \varepsilon)\psi = 0,$$

where

$$\tilde{r} = \frac{1}{4}\lambda^2 - |u|^2 + \frac{1}{2}i\varepsilon\lambda(\ln u)_x + \varepsilon^2\left[\frac{1}{2}(\ln u)_{xx} - \frac{1}{4}(\ln u)_x^2\right].$$

Expressing $\tilde{r}$ in terms of the hydrodynamic variables ρ and v helps to ascertain the actual orders of the terms of expansion in powers of ε, given that

$$(\ln u)_x = \frac{1}{2}(\ln\rho)_x + \frac{i}{\varepsilon}v.$$

Disregarding the small terms $O(\varepsilon)$ and $O(\varepsilon^2)$ in the coefficient $\tilde{r}(x, \lambda; \varepsilon)$ in the principal order of the WKB approximation, we obtain a simpler equation,

$$\varepsilon^2\psi_{xx} + r(x, \lambda)\psi = 0, \tag{36}$$

where

$$r(x, \lambda) = \frac{1}{4}[\lambda - v(x)]^2 - \rho(x). \tag{37}$$

It is convenient to explicitly separate out the trivial dependence of (9) on well width θ in (37) (considering one well, we omit the subscript m). Taking into account (9), we have

$$r(x, \lambda; \theta) = r_1\left(\frac{x}{\theta}, \lambda\right),$$
$$r_1(x, \lambda) = \frac{1}{4}[\lambda - g(x)]^2 - \rho_0 + f(x). \tag{38}$$

Let us describe the spectrum of problem (36) in the semiclassical limit $\varepsilon \longrightarrow 0$. Denote $\omega = 2\sqrt{\rho_0}$. Let the velocity perturbation be small enough for the overall pattern of variation in the function $r(x, \lambda)$ to be the same as that for $v(x) \equiv 0$. More specifically, $r(x, \lambda)$ is positive for $|\lambda| > \omega$ and either changes sign with increasing x according to the scheme –, +, – or remains always negative for $|\lambda| < \omega$ (Fig. 1a). In that case, problem (36) has a continuous spectrum with a zero reflection coefficient (in the semiclassical approximation under consideration) for $|\lambda| > \omega$ and a discrete spectrum with closely spaced eigenvalues $|\lambda_{n+1} - \lambda_n| \sim \varepsilon$ for $|\lambda| < \omega$. There are

no eigenvalues in the interval between $m_+ = \min\lambda_+(x)$ and $m_- = \max\lambda_-(x)$, where

$$\lambda_\pm(x) = v(x) \pm 2\sqrt{\rho(x)}$$

are the roots of the equation $r(x, \lambda) = 0$. The restriction

$$-2(\sqrt{\rho_0} - \rho(x)) < v(x) < 2(\sqrt{\rho_0} - \sqrt{\rho(x)}) \qquad (39)$$

is the condition of sufficient smallness for $v(x)$. If it is satisfied, then $\lambda_+(x) \le \omega$ and $\lambda_-(x) \ge -\omega$ for all x.

4.1. The Discrete Spectrum

The discrete spectrum of problem (36) is described (in our approximation) by an equation similar to the Bohr–Sommerfeld quantization condition:

$$S(\lambda_n) = \frac{\varepsilon\pi}{\theta}\left(\frac{1}{2} + n\right), \quad n = 0, 1, \ldots, \qquad (40)$$

where

$$S(\lambda) = \int_{x_-(\lambda)}^{x_+(\lambda)} \sqrt{r_1(x, \lambda)}\,dx, \qquad (41)$$

and $x_-(\lambda) < x_+(\lambda)$ are the roots of the equation $r_1(x, \lambda) = 0$. Since $S(\lambda)$ is a nonmonotonic function (Fig. 1b), Eq. (40) describes two sequences of spectral levels: λ_n^-, $n = 0, 1, \ldots, N^- - 1$ in the interval $(-\omega, m_-)$ and λ_n^+, $n = 0, 1, \ldots, N^+ - 1$ in the interval (m_+, ω).

In the limit $\varepsilon \longrightarrow 0$, the discrete spectrum is quasicontinuous. It can be described by the density $d(\lambda)$ with the normalization condition

$$\int_{-\omega}^{\omega} d(\lambda)\,d\lambda = 1. \qquad (42)$$

Thus, the number of eigenvalues in the interval $(\lambda, \lambda + d\lambda)$ is $Nd(\lambda)d\lambda$, where $N \sim \varepsilon^{-1}$ is the total number of discrete levels, which is calculated below.

Clearly, the spectral density $d(\lambda)$ is inversely proportional to the interval $|\lambda_{n+1} - \lambda_n|$ between adjacent levels. Therefore, taking into account (40), we have

$$d(\lambda) = A\left|\frac{dS}{d\lambda}\right|, \qquad (43)$$

where the proportionality coefficient

$$A = [S(-\omega) + S(\omega)]^{-1} \qquad (44)$$

is specified by the normalization condition (42).

The total number of levels is

$$N = N_- + N_+ = \frac{\theta}{\varepsilon\pi A}, \qquad (45)$$

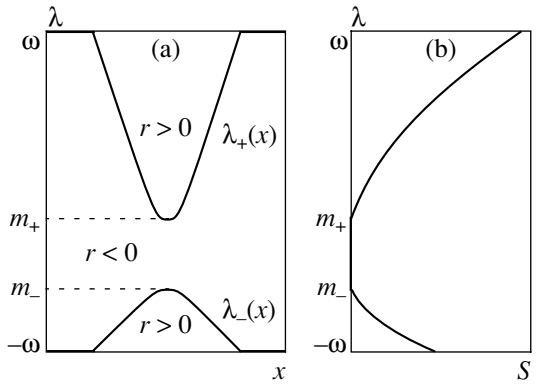

Fig. 1. (a) The domains of constant sign of the function $r(x, \lambda)$; (b) the Bohr–Sommerfeld integral $S(\lambda)$.

because it follows from (40) that

$$N_- = \theta(\varepsilon\pi)^{-1}S(-\omega) + O(1)$$

levels lie in the range $(-\omega, m_-)$ and

$$N_+ = \theta(\varepsilon\pi)^{-1}S(\omega) + O(1)$$

levels lie in the range (m_+, ω).

The above semiclassical description of the scattering problem (12) can now be used to characterize the spectral data for the potential $u(x, 0)$ that corresponds to the infinite sequence of wells (8).

We emphasize that an accurate determination of the spectrum for the Schrödinger equation with a nonperiodically oscillating potential specified on the x axis is a complex mathematical problem [10]. In our case, the problem is simplified, because we are interested not in the detailed structure of the spectrum but only in the density of discrete levels in the semiclassical limit $\varepsilon \ll 1$. The density $d(\lambda)$ can be determined from the part of the spectrum for the potential $u(x, 0)$ that refers to a finite segment $L \gg L_0$, where L_0 is the characteristic homogeneity scale of $u(x, 0)$. Since the initial function is ergodic, we may assume that the normalized density $d(\lambda)$ does not depend on L and tends to a definite function as $L \longrightarrow \infty$.

Each eigenvalue in the (everywhere) dense point spectrum for the infinite sequence of wells (8) refers to a certain well in accordance with the localization of the eigenfunction $\psi_\lambda(x)$ that belongs to this eigenvalue. Therefore, if we take a sufficiently long interval with many wells, then we can determine the total number $N(L)$ of spectral levels λ, each referring to a particular well in this interval. Because of the ergodicity of $u(x, 0)$, the total linear level density $N(L)/L$ has a limit for $L \longrightarrow \infty$. Since the number of levels for each individual well is inversely proportional to the dispersion parameter ε, it is convenient to normalize this limit:

$$C = \varepsilon \lim_{L\to\infty} \frac{N(L)}{L}. \qquad (46)$$

The dimensionless density C of discrete levels defined in this way no longer depends on ε.

The fraction in the density C of spectral levels in the interval $(\lambda, \lambda + d\lambda)$,

$$\frac{dC}{C} = d(\lambda)d\lambda, \tag{47}$$

determines the spectral density $d(\lambda)$, which is clearly normalized according to (42).

Let us calculate the linear level density C for our initial conditions (9) composed of wells with the same shape and depth. According to (45), $N_m = \theta_m/\varepsilon\pi A$ discrete levels correspond to each individual well. The total number of levels divided by the length of the interval L containing a large number ($M \gg 1$) of wells is

$$\frac{N(L)}{L} = \frac{\sum\theta_m}{L\varepsilon\pi A} = \frac{M}{L}\frac{1}{\varepsilon\pi A}\frac{\sum\theta_m}{M}.$$

Here, the summation is performed over those values of m from (8) and (9) for which the point x_m lies in the above interval L. Substituting this expression into (46) and passing to the limit, we obtain using (11)

$$C = \frac{\gamma}{\pi A}, \quad \gamma = \frac{\bar{\theta}}{\bar{l}}. \tag{48}$$

In our case, the spectral density $d(\lambda)$ can also be easily calculated. Since there are no resonances between the levels in different wells for a nonperiodic variation in their width (10), the total spectrum can be obtained by a simple superposition of the spectra for individual wells with the same $d(\lambda)$, because formulas (43) and (44) do not contain the well width θ. Thus, for an infinite sequence of wells, the spectral density $d(\lambda)$ is given by the same expressions (43) and (44).

Let us establish an important relationship between the conserved densities $\langle R_n\rangle$ from (35) and the spectral density $d(\lambda)$. For the initial perturbation (9), the integral in definition (35) can be represented as

$$\frac{1}{L}\int_{-L/2}^{L/2} R_n(x)dx = \sum_m J_{nm}, \tag{49}$$

where the summation is performed over the well numbers within the interval $(-L/2, L/2)$ and

$$J_{nm} = \frac{1}{2L}\int_{-L}^{L} R_{nm}(x)dx$$

is the conserved integral calculated for the individual mth well.

The invariant J_{nm} for a localized perturbation can be represented not only as an integral over the x coordinate but also as a spectral decomposition. This is the so-called trace identity [9]. In the semiclassical case, where the integral over the continuous spectrum may be disregarded, it reduces to

$$J_{nm} = \sum_{i=0}^{N_m-1} \varepsilon j_n(\lambda_i^{(m)}). \tag{50}$$

Here, $\lambda_i^{(m)}$, $i = 0, \ldots, N_m - 1$ are the discrete spectral levels in the mth well, and $\varepsilon j_n(\lambda)$ are the conserved integrals for an isolated soliton (7) propagating at velocity $v_s = \lambda$. The factor ε appears, because the soliton linear size is of the order of ε on the scale under consideration.

The trace identity means that the calculated invariant J_n for a localized perturbation is represented as the sum of its values for individual solitons into which this perturbation breaks up in the limit $t \longrightarrow \infty$ and its part that corresponds to the continuous spectrum and that scatters over the entire x axis for $t \longrightarrow \infty$. In the semiclassical case under study, this part is negligible compared to the soliton component.

The functions $j_n(\lambda)$ are given by the formulas [9, p. 75]

$$j_{2k+1}(\lambda) = (-1)^{k+1}\frac{\nu^{2k+1}}{2k+1},$$
$$j_{2k+2}(\lambda) = (-1)^{k+1}\frac{\lambda\nu}{2(k+1)}\sum_{p=0}^{k}(-1)^p b_p^- \omega^{2p}\nu^{2(k-p)}, \tag{51}$$

where $k = 0, 1, \ldots$; $\nu = \sqrt{\omega^2 - \lambda^2}$; and b_p^- are the coefficients from the expansion

$$(1+x)^{-1/2} = \sum_{p=0}^{\infty} b_p^- x^p.$$

The first three functions in (51) are

$$j_1(\lambda) = -\nu, \quad j_2(\lambda) = -\frac{\lambda\nu}{2}, \quad j_3(\lambda) = \frac{\nu^3}{3}.$$

Substituting the integral for the sum in (50) in the semiclassical limit, when $\lambda_{i+1} - \lambda_i \sim \varepsilon \ll 1$, yields

$$J_{nm} = \varepsilon N_m \int_{-\omega}^{\omega} j_n(\lambda)d(\lambda)d\lambda.$$

Since the spectral density $d(\lambda)$ for the sequence of wells that differ only in width does not depend on m, we obtain integral (49) in the form

$$\int_{-L/2}^{L/2} R_n(x)dx = \varepsilon N(L)\int_{-\omega}^{\omega} j_n(\lambda)d(\lambda)d\lambda,$$

where $N(L) = \sum_m N_m$ is the total number of spectral levels in the interval L. Using the latter expression in the definition of invariant densities (35) and taking into

account (46), we find the sought-for relationship between $\langle R_n \rangle$ and the spectral parameters in the initial state:

$$\langle R_n \rangle = C \int_{-\omega}^{\omega} j_n(\lambda) d(\lambda) d\lambda. \tag{52}$$

The simple relation (52) between the mean invariant densities and the spectrum of the scattering problem is a remarkable general property of our "hydrodynamic" initial-value problem (6) and (9) under the ergodicity conditions (cf. [4]).

4.2. The Scattering Phase

Let us now describe the set of scattering phases. Since the spectral equation (36) in the semiclassical limit is analogous to the spectral equation in the case of the Korteweg–de Vries equation, the result of our scattering-phase study matches that from [4]. Recall briefly how the parameters of the shift

$$s_n = \frac{\varepsilon}{\lambda_n} \ln |b_n|, \tag{53}$$

where $b_n = C_{n+}/C_{n-}$ is the ratio of the coefficients in the asymptotic expressions of eigenfunctions ψ_n for $x \longrightarrow \pm\infty$, are calculated. Let us determine the phase shift for the general case of M wells. To this end, lets us ascertain how the phase of the wave function localized in the well under consideration changes when passing the jth well to which the turning points x_r and x_{r+1} $(r = 2j - 1)$ refer. We separated three domains in the passage of this well (see Fig. 1 in [4]). In domain I, we have an exponentially decreasing solution of the initial wave function with an amplitude A_{r-1}:

$$\psi_{\mathrm{I}} = \frac{A_{r-1}}{\sqrt{p}} \exp\left(-\frac{1}{\varepsilon} \int_{X_1}^{x} |p| dx \right). \tag{54}$$

Here, $p = \sqrt{-r(x, \lambda)}$.

In domain II, the wave function is

$$\psi_{\mathrm{II}} = \frac{A_r}{\sqrt{p}} \exp\left(\frac{i}{\varepsilon} \int_{X_1}^{x} |p| dx \right) + \frac{B_r}{\sqrt{p}} \exp\left(-\frac{i}{\varepsilon} \int_{X_1}^{x} |p| dx \right). \tag{55}$$

Going around the point X_2 in the complex domain from above, we derive from (54) and (55) the relationship between the coefficients

$$A_r = A_{r-1} \exp\left(i\frac{\pi}{4} - \frac{1}{\varepsilon} \int_{X_1}^{x} |p| dx \right).$$

Hence, up to the point X_3, we have the solution

$$\psi_{\mathrm{II}}^{+} = A_{r-1} \exp\left(i\frac{\pi}{4} - \frac{1}{\varepsilon} \int_{X_1}^{X_2} |p| dx + \frac{i}{\varepsilon} \int_{X_2}^{x} |p| dx \right).$$

Similarly, going around the point X_3, we obtain

$$A_{r+1} \exp\left(-\frac{1}{\varepsilon} \int_{X_3}^{x} |p| dx \right) = A_{r-1} \exp\left(i\frac{\pi}{4} - \frac{1}{\varepsilon} \int_{X_1}^{X_2} |p| dx + \frac{i}{\varepsilon} \int_{X_2}^{X_3} |p| dx - \frac{1}{\varepsilon} \int_{X_3}^{x} |p| dx \right). \tag{56}$$

Only the absolute value of the amplitude ratio is of interest in determining the shifts. From (56), we derive the following relation between the amplitudes when passing the jth potential well, the levels of which do not coincide with the levels in the first well because of the nonperiodic variation in well width (10):

$$\left| \frac{A_{r+1}}{A_{r-1}} \right| = \exp\left(-\frac{1}{\varepsilon} \int_{X_1}^{X_2} |p| dx \right).$$

Repeating this procedure M times and taking into account the integrals over the segments $(-\infty, x_1)$ and $(x_M, +\infty)$, we have

$$\left| \frac{C_+}{C_-} \right| = \exp\left\{ \frac{\lambda}{2\varepsilon}(x_1 + x_M) - \frac{1}{\varepsilon} \sum_{j=1}^{M-1} \int_{x_{2j}}^{x_{2j+1}} |p| dx + \frac{1}{\varepsilon} \left(\int_{x_1}^{-\infty} + \int_{x_M}^{+\infty} \right) \left(\frac{\lambda}{2} - |p| \right) dx \right\}. \tag{57}$$

The integration in (57) is performed over the ranges where the wave function exponentially decays. Using (57) and (53), we finally obtain

$$s_n(k, M) = (x_1 + x_M) - \frac{1}{\lambda} \sum_{j=1}^{M-1} \int_{x_{2j}}^{x_{2j+1}} |p| dx + \frac{1}{\lambda} \left(\int_{x_1}^{-\infty} + \int_{x_M}^{+\infty} \right) \left(\frac{\lambda}{2} - |p| \right) dx. \tag{58}$$

It is important to emphasize that the minus sign in front of the sum in expressions (57) and (58) implies that we found the corrections to the shifts from the wells to the right from the well under consideration. If we perform a similar procedure for the wells to the left, then [here, the amplitudes B_r (55) are significant) formulas (57) and (58) retain their form; only the sign in front of the

sum changes to plus. Formula (58) is similar to expression (40) from [4], where it was shown that the sum has no finite limit for $M \longrightarrow \infty$ and, consequently, the phase uniformly fills the interval of length l between the minima of the potential wells at points x_0 and x_1. The filling of this interval is quite similar to the generation of chaos in dynamical systems [11].

For the "nonsymmetric" passage to $M \longrightarrow \infty$, i.e., in a situation where the numbers of wells to the left and to the right differ by a value proportional to M, all of the above properties are preserved but the interval in which the phase is smeared shifts proportionally to an arbitrary parameter M.

Thus, under the initial conditions (8) and (9), i.e., for an infinite number of potential wells, the scattering phase of the wave function localized in any initial well randomly and uniformly fills any of the segments $[x_m, x_{m+1}]$. It was shown in [3, 4] that allowance for the mutual influence of an infinite number of initial wells corresponds to the asymptotic limit $t \longrightarrow \infty$. Consequently, in the asymptotic limit $t \longrightarrow \infty$, our initial-value problem is described by an infinite system of solitons with random phases.

5. THE STATISTICAL LIMIT OF THE N-SOLITON SOLUTION WITH RANDOM PHASES

Let us now describe an infinite system of solitons with random phases. According to statistical physics [12], to do this requires separating out an arbitrary subsystem of N solitons located in some segment and then averaging over an ensemble of states with random phases s by assuming the segment length to indefinitely increase as the number N increases proportionally (the statistical limit). In this case, it would be natural to use the standard exact solution for N interacting solitons [9], which is convenient to transform to

$$u_N(x, t) = \sqrt{\rho_0}\frac{D_1(x, t)}{D(x, t)}. \tag{59}$$

Here, D and D_1 are the sums over the binary multi-indices $\mu = \mu_1, \ldots, \mu_N$ ($\mu_n = 0$ or 1):

$$D = \sum_{\mu} \exp(\Phi_\mu), \quad D_1 = \sum_{\mu} \exp(i\theta_\mu)\exp(\Phi_\mu), \tag{60}$$

where

$$\Phi_\mu = \varepsilon^{-1}\sum_{n=1}^{N} \mu_n \nu_n (x - \lambda_n t - s_n) + \sum_{m=1}^{N}\sum_{n=1}^{N} \mu_m \mu_n l_{mn}, \quad \theta_\mu = \varepsilon^{-1}\sum_{n=1}^{N} \mu_n \theta_n. \tag{61}$$

The quantities ν_n, θ_n, and l_{mn} are determined by the spectrum λ_n:

$$\nu_n = \sqrt{\omega^2 - \lambda_n^2} > 0, \quad |\lambda_n| < \omega, \quad \omega^2 = 4\rho_0,$$
$$\exp(i\theta_n) = \bar{z}_n/z_n, \quad z_n = \lambda_n + i\nu_n, \tag{62}$$
$$0 < \theta_n < 2\pi,$$

$$l_{mn} = \ln\left|\frac{z_m - z_n}{z_m - \bar{z}_n}\right| = \frac{1}{2}\ln\frac{\omega^2 - \lambda_m\lambda_n - \nu_m\nu_n}{\omega^2 - \lambda_m\lambda_n + \nu_m\nu_n}, \tag{63}$$

$s_1, \ldots, s_N$ is the set of shift parameters. Note an analogy between expressions (59)–(63) and the standard Kay–Moses–Hirota formulas [13] for the N-soliton solution of the Korteweg–de Vries equation that we used in [4].

The absolute value of the N-soliton solution (59) is given by the relation

$$R_1 = |u|^2 - \rho_0 = -\varepsilon^2\partial_{xx}\ln D, \tag{64}$$

which is similar to the N-soliton formula for the Korteweg–de Vries equation. Expression (64) derived by the Hirota method (see, e.g., [14]) is used in the next section to construct similar formulas for all densities R_n.

The properties of the multisoliton solutions described by formulas (59)–(61) are illustrated in Fig. 2. We took $\rho_0 = 1$. In Fig. 2a, $|u|$ is plotted against x; the arrows indicate the direction of soliton motion. Solitons *4* and *5* are in collision. The flow velocity distribution is shown in Fig. 2b. Figure 2c shows the hodograph of the complex function $u(x)$, i.e., the parametrically specified dependence of $\mathrm{Im}\,u(x)$ on $\mathrm{Re}\,u(x)$. Since $|u(x)|^2 < \rho_0$ for solitons, the hodograph is located with the circle $|u| < 1$. Each isolated soliton is indicated by a rectilinear segment with the ends on the circumference $|u| = 1$. The curve corresponds to the colliding solitons *4* and *5*.

Let us return to the problem of the statistical limit for the N-soliton solution (59)–(63). In accordance with the solution of the scattering problem (Section 4), we assume the parameters $s_1, \ldots, s_N$ determined by the soliton initial phases to be random, i.e., uniformly and independently distributed over the interval $(-L_s/2, L_s/2)$. Expressions (59)–(63) establish a one-to-one correspondence between the set of points in S space with the coordinates $s_n(t) = s_n + \lambda_n t$, which, according to (61), move uniformly, i.e., do not interact, and the set of interacting solitons. It would be natural to describe the former set by its mean concentration $C_s = \varepsilon N/L_s$, where L_s is the length of the interval in which points s_n are located. Thus, we arrive at a problem of the statistical properties for function (59)–(63) similar to that considered in [4] in which the shift parameters $s_1, \ldots, s_N$ are random variables uniformly and independently distributed over the interval of length L_s; $L_s \longrightarrow \infty$ is proportional to N, so the ratio $\varepsilon N/L_s = C_s$ is constant.

The problem is also to describe the asymptotic state that arises when $t \longrightarrow \infty$ through the evolution of the initial conditions (9) in terms of the random functions $u_N(x, t)$ in the statistical limit. It may be called soliton turbulence. The soliton spectrum is specified by the known function $d(\lambda)$ from (47); the soliton spatial density C (46) is also known. At the same time, the linear concentration C_s of points in the S set and their distribution in λ, i.e., the density $\rho_s(\lambda)$ of the set of λ_n in (61)–(63), are unknown; these have to be calculated. We define the normalization of $\rho_s(\lambda)$, as in (42), by the condition

$$\int_{-\omega}^{\omega} \rho_s(\lambda) d\lambda = 1. \tag{65}$$

This implies that the densities C_s and C are related by

$$C_s = C\frac{L}{L_s}, \quad CL = \varepsilon N,$$

where L is the scale on the x axis on which N interacting solitons are distributed. Here, we took into account the conservation laws for the number of solitons and the one-to-one correspondence between the solitons and the points of S space.

By analogy with the N-soliton solution of the Korteweg–de Vries equation [4], we seek a maximum of the exponent Φ_μ in μ. The corresponding quantity without allowance for the interacting solitons,

$$\varepsilon\Phi_\mu^{(0)} = \sum_{n=1}^{N} \mu_n \nu_n (a_n - s_n), \quad a_n = x - \lambda_n t,$$

is exactly maximized for

$$\mu_n = \eta(a_n - s_n), \quad \eta(x) = \begin{cases} 0, & x < 0, \\ 1, & x > 0. \end{cases} \tag{66}$$

A good approximate maximizer $\tilde{\mu}$ of the total Φ_μ at large N in which the "invalid" elements $\tilde{\mu}_n$ constitute only a small fraction (see [4] for a detailed estimation of the error) may be sought in a form similar to (66):

$$\tilde{\mu}_n = \eta(b_n - s_n), \tag{67}$$

where b_n are to be determined.

If only one nth index varies in the multi-index $\tilde{\mu}$,

$$\begin{aligned} \tilde{\mu}_p \longrightarrow \mu_p &= \tilde{\mu}_p + \sigma_n \delta_{np}, \\ \sigma_n &= 1 - 2\tilde{\mu}_n = (-1)^{\tilde{\mu}_n}, \end{aligned} \tag{68}$$

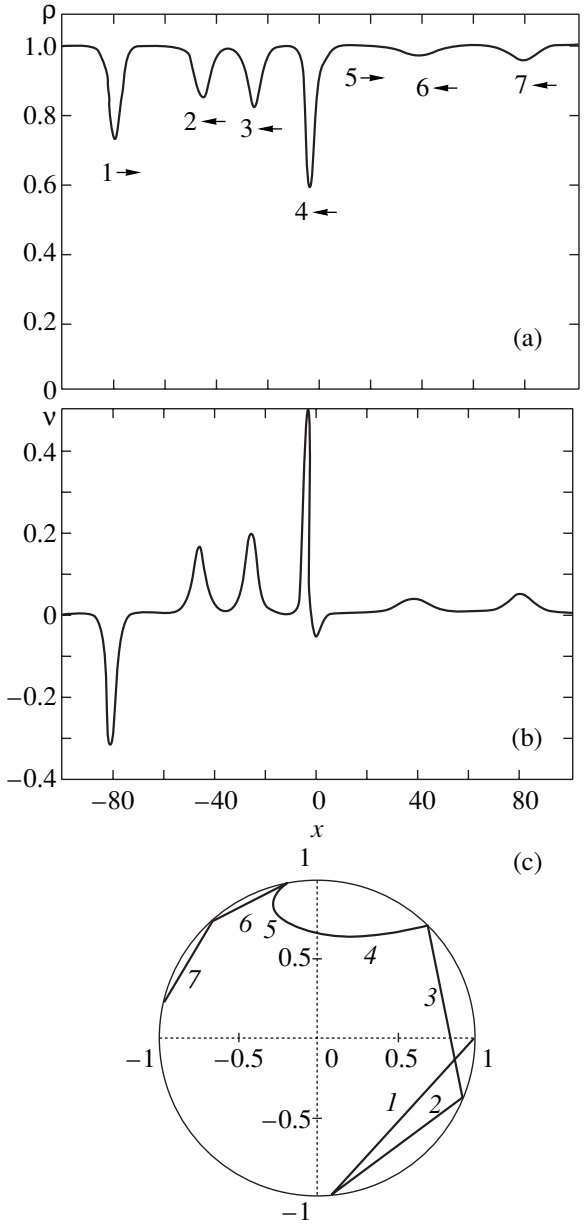

Fig. 2. An example of a multisoliton solution. (a) The density $\rho(x) = |u|^2$ and direction of soliton motion; the soliton amplitudes are 0.27 (*1*), 0.15 (*2*), 0.18 (*3*), 0.46 (*4*), 0.12 (*5*), 0.03 (*6*), and 0.04 (*7*); the soliton velocities are 1.46 (*1*), –1.70 (*2*), –1.65 (*3*), –1.08 (*4*), 1.75 (*5*), –1.94 (*6*), and –1.91 (*7*). (b) The hydrodynamic flow velocity $v(x)$. (c) The hodograph of the complex function $u(x)$.

then the increment of $\varepsilon\Phi_\mu$ is

$$\begin{aligned} \Delta_n \equiv \varepsilon(\Phi_\mu - \Phi_{\tilde{\mu}}) &= \sigma_n \left[\nu_n (a_n - s_n) + 2\varepsilon \sum_{p=1}^{N} l_{np} \tilde{\mu}_p \right] \\ &= \nu_n \sigma_n \left(a_n - s_n + \frac{2\varepsilon}{\nu_n} \sum_{p=1}^{N} l_{np} \langle \tilde{\mu}_p \rangle \right) + \tilde{\Delta}_n. \end{aligned}$$

Here, $\langle\tilde{\mu}_n\rangle$ is the mean value of $\tilde{\mu}_n$, which, in view of (67), depends on b_n,

$$\langle\tilde{\mu}_n\rangle = \frac{1}{L_s}\left[b_n - \left(-\frac{L_s}{2}\right)\right] = \frac{b_n}{L_s} + \frac{1}{2}, \tag{69}$$

and $\tilde{\Delta}_n$ is the random component of Δ_n with a zero mean. The following linear system of equations is the condition that the "systematic" part of the increment Δ_n be negative for all n:

$$a_n + \frac{2\varepsilon}{\nu_n}\sum_{p=1}^{N} l_{np}\left(\frac{b_p}{L_s} + \frac{1}{2}\right) = b_n, \quad n = 1, \ldots, N. \tag{70}$$

If we calculate $b_1, \ldots, b_n$ from this system, then, in view of (67) and (68), we have for any n

$$\Delta_n = \nu_n\sigma_n(b_n - s_n) + \tilde{\Delta}_n = -\nu_n|b_n - s_n| + \tilde{\Delta}_n,$$

i.e., $\Delta_n \le 0$ except for those cases where the random increment $\tilde{\Delta}_n$ is positive and exceeds the absolute value of the systematic part of Δ_n:

$$\tilde{\Delta}_n > \nu_n|b_n - s_n|.$$

Following [4], we can show that this condition for $N \longrightarrow \infty$ is satisfied only for an infinitesimal fraction of all N elements in the multi-index $\tilde{\mu}$ (67). In other words, condition (70) exactly determines the maximizer in the limit $N \longrightarrow \infty$.

Given relation (69), the system of equations (70) can be replaced with an equivalent system for the unknowns $y^{(n)} = \langle\tilde{\mu}_n\rangle$. We will also use an explicit expression of a_n in terms of x, t and the relation of ν_n to λ_n. As a result, we obtain the system of equations

$$y^{(n)} = \frac{1}{2} + \frac{x}{L_s} - \lambda_n\frac{t}{L_s} + \frac{2\varepsilon}{L_s\sqrt{\omega^2 - \lambda_n^2}}\sum_{p=1}^{N} l_{np}y^{(p)}. \tag{71}$$

Substituting the integral for the sum when $N \longrightarrow \infty$ [the number of spectral levels λ_n in the interval $(\lambda, \lambda + d\lambda)$ is $N\rho_s(\lambda)d\lambda$], we derive an integral equation of the second kind ($x/L_s = \xi$, $t/L_s = \tau$, $N/L_s = \varepsilon^{-1}C_s$) from (71):

$$y(\lambda; \xi, \tau) = \frac{1}{2} + \xi - \lambda\tau + \frac{2C_s}{\nu(\lambda)}\int_{-\omega}^{\omega} l(\lambda, \lambda')\rho_s(\lambda')y(\lambda'; \xi, \tau)d\lambda', \tag{72}$$

where

$$\nu(\lambda) = \sqrt{\omega^2 - \lambda^2},$$

$$l(\lambda, \lambda') = \frac{1}{2}\ln\frac{\omega^2 - \lambda\lambda' - \nu(\lambda)\nu(\lambda')}{\omega^2 - \lambda\lambda' + \nu(\lambda)\nu(\lambda')}.$$

In the next section, we will need to generalize the multisoliton solution (59)–(63) to the entire hierarchy of higher NSEs (30) to calculate the invariant densities $\langle R_n\rangle$.

To obtain the "multitime" N-soliton solution, it will suffice to substitute the combination $v_1(\lambda_n)t_1 + v_2(\lambda_n)t_2 + \ldots$ for $\lambda_n t$ in (61). Here, $v_m(\lambda)$ is the velocity of the soliton of the mth equation in hierarchy (30) that corresponds to the discrete spectral level λ. These velocities are determined by the dynamics of the transition coefficients of the discrete spectrum for a finite density [9, pp. 240, 241] and can be calculated from the formulas ($k = 0, 1, 2, \ldots$)

$$v_{2k+1}(\lambda) = (-1)^k\lambda\nu^{2k},$$
$$v_{2k}(\lambda) = (-1)^k\sum_{p=0}^{k}(-1)^p b_p^+\omega^{2p}\nu^{2(k-p)}, \tag{73}$$

where $\nu = \nu(\lambda)$ and b_p^+ are the coefficients for the expansion of $(1 + x)^{1/2}$ in powers of x. In particular, $v_1(\lambda) = \lambda$ is the soliton velocity for Eq. (19). The set of soliton velocities $v_1, v_2, \ldots$ is supplemented here with the element $v_0 = 1$, which is convenient in writing the general formulas.

All the reasoning that leads to the integral equation (72) remains valid and gives an equation for the generalized averaged maximizer $y(\lambda; \xi, \tau_1, \tau_2, \ldots)$ that depends on many times:

$$y(\lambda) = \frac{1}{2} + \xi - v_1(\lambda)\tau_1 - v_2(\lambda)\tau_2 - \ldots + \frac{C_s}{\nu(\lambda)}\int_{-\omega}^{\omega} l(\lambda, \lambda')\rho_s(\lambda')y(\lambda')d\lambda'. \tag{74}$$

At $\tau_1 = \tau$ and $\tau_2 = \tau_3 = \ldots = 0$, we have the special case (72). The solution of Eq. (74) is a linear combination,

$$y(\lambda) = \left(\frac{1}{2} + \xi\right)y_0(\lambda) - \sum_{n>0}\tau_k y_k(\lambda), \tag{75}$$

of the solutions to the equations

$$y_k(\lambda) = v_k(\lambda) + \frac{2C_s}{\nu(\lambda)}\int_{-\omega}^{\omega} l(\lambda, \lambda')\rho_s(\lambda')y_k(\lambda')d\lambda', \quad k = 0, 1, 2, \ldots. \tag{76}$$

The identity used below follows from relations (76):

$$\int_{-\omega}^{\omega}\nu(\lambda)\rho_s(\lambda)y_k(\lambda)d\lambda = \int_{-\omega}^{\omega}\nu(\lambda)\rho_s(\lambda)v_k(\lambda)y_0(\lambda)d\lambda, \quad k = 1, 2, \ldots. \tag{77}$$

To prove it, we must multiply Eq. (76) by $\nu(\lambda)\rho_s(\lambda)y_k(\lambda)$ at $k = 0$ and by $\nu(\lambda)\rho_s(\lambda)y_0(\lambda)$ for nonzero k and, subsequently, integrate the derived relations over λ from $-\omega$ to ω and subtract one from the other.

6. CALCULATING THE CONSERVED DENSITIES IN A STATISTICAL STATE USING THE AVERAGED MAXIMIZER

In Section 4, we established relation (52) between the conserved quantities $\langle R_n\rangle$ and the spectral densities C and $d(\lambda)$. Let us derive similar formulas for $\langle R_n\rangle$ via the spectral parameters C_s and $\rho_s(\lambda)$ of the statistical soliton state described in the preceding section. To this end, we use the averaged maximizer constructed above.

As the starting point, we make use of the local relations between the density of the conserved quantities (23) and the sum D (60), which underlies the N-soliton solution. Denote

$$G_0 = -\partial_{xx}\ln D,$$

$$G_m = \frac{\partial^2 \ln D}{\partial t_m \partial x} \quad \text{for} \quad m \geq 1.$$

The following relation holds:

$$R_n = \frac{\varepsilon^2}{n}\sum_{0\leq 2p<n} b_p^- \omega^{2p} G_{n-1-2p}, \quad n = 1, 2, \ldots, \quad (78)$$

where the coefficients b_p^- are defined above [see (51)] and the summation is performed over all integer p that satisfy the above inequality. For brevity, the inessential additional terms that are the total derivatives were omitted in (78). The validity of the first formula in (78), i.e., expressions (64), has already been established. To prove (78) for $n > 1$, we use the conservation laws in higher times (32) at $n = 1$ and the generating functions (20).

Differentiating (64) with respect to t_m ($m = 1, 2, \ldots$) yields the relation

$$\frac{\partial R_1}{\partial t_m} + \varepsilon^2 \frac{\partial G_m}{\partial x} = 0,$$

which is the conservation law (32) for $n = 1$. In this case, the flux is given by formula (33), which allows G_m to be related to the invariant densities Q_k: to within the total derivative with respect to x, we have

$$\varepsilon^2 G_m = \sum_{k=1}^{m+1} k\beta_{k+1,m+1-k} Q_k, \quad (79)$$

where $m = 1, 2, \ldots$. When deriving this relation, we took into account the fact that the function $G_m(x)$, as well as the densities $Q_k(x)$, tends to zero as $|x| \longrightarrow \infty$.

It is important to note that (79) also holds for $m = 0$. This follows from (64), given that $R_1 = Q_1$ (22) and $\beta_{n0} = 1$ (21). In this case, system (79) becomes triangular and allows us to successively calculate Q_n as linear combinations of $G_0, G_1, \ldots, G_{n-1}$ and then derive formulas (78) for R_n using (22). To efficiently perform this procedure, let us define the generating function

$$G(\eta) = \sum_{m=0}^{\infty} G_m \eta^m.$$

The relation between $G(\eta)$ and $Q(\xi)$ that follows from the system of equations (79) with $m = 0, 1, 2, \ldots$ is

$$\xi^2 \frac{dQ}{d\xi} = \varepsilon^2 \eta^2 G(\eta). \quad (80)$$

This can be easily established by calculating the expansion of the derivative $dQ/d\xi$ in powers of η using expansions (21). Since $\xi = \eta(1 + \omega^2\eta^2)^{-1/2}$, we obtain

$$\frac{dQ}{d\xi} = \frac{\eta^3}{\xi^3}\frac{dR}{d\eta}$$

and derive the sought-for relation between the generating functions $R(\eta)$ and $G(\eta)$ from (80):

$$\frac{dR}{d\eta} = \varepsilon^2 (1 + \omega^2\eta^2)^{-1/2} G(\eta),$$

whence follows relation (78) for the coefficients of the expansions $R(\eta) = R_1\eta + R_2\eta^2 + \ldots$ and $G(\eta) = G_0 + G_1\eta + \ldots$.

Let us average the quantities (78). We make use of the equality established in [4] for the sums

$$D = \sum_{\mu} \exp(\Phi_\mu)$$

in the continuum limit when $N \longrightarrow \infty$ between the mean quantities

$$\left\langle \partial_x \ln \sum_{\mu} \exp(\Phi_\mu) \right\rangle = \left\langle \partial_x \max_{\mu} \Phi_\mu \right\rangle. \quad (81)$$

This relation allows the above averaged maximizer $\langle \tilde{\mu}_n\rangle = y^{(n)}$ to be used: according to (61),

$$\partial_x \max_{\mu} \Phi_\mu = \varepsilon^{-1} \sum_{n=1}^{N} \tilde{\mu}_n \nu_n,$$

and we have for $N \longrightarrow \infty$

$$\left\langle \partial_x \max_{\mu} \Phi_\mu \right\rangle = \varepsilon^{-1} \sum_{n=1}^{N} y^{(n)} \nu_n. \quad (82)$$

Substituting the integral for the latter sum in the continuum limit and combining (81) and (82) yields

$$\langle \partial_x \ln D\rangle = \varepsilon^{-1} N \int_{-\omega}^{\omega} y(\lambda)\nu(\lambda)\rho_s(\lambda)d\lambda. \quad (83)$$

Based on formulas (78) and the averaging rule for (83), we can derive an expression for the mean $\langle R_n \rangle$ that contains a linear combination of the derivatives of the generalized averaged maximizer $y(\lambda; x, t_1, t_2, \ldots)$ with respect to $x, t_1, t_2, \ldots$ under the integral. These derivatives can be easily calculated from (75):

$$\frac{\partial y}{\partial x} = L_s^{-1} y_0(\lambda), \quad \frac{\partial y}{\partial t_n} = -L_s^{-1} y_n(\lambda),$$

which allows the following expression for the mean densities to be derived by using identity (77):

$$\langle R_n \rangle = -\frac{C_s}{n} \times \int_{-\omega}^{\omega} \sum_{0 \le 2p < n} b_p^- \omega^{2p} v_{n-1-2p}(\lambda) y_0(\lambda) v(\lambda) \rho_s(\lambda) d\lambda. \quad (84)$$

The soliton velocities v_n and the conserved integrals for an isolated soliton j_n are related by the relations that follow from (51) and (73):

$$j_n(\lambda) = -\frac{v(\lambda)}{n} \sum_{0 \le 2p < n} \omega^{2p} b_p^- v_{n-1-2p}(\lambda), \quad n = 1, 2, \ldots.$$

Given these identities, formulas (84) reduce to

$$\langle R_n \rangle = C_s \int_{-\omega}^{\omega} j_n(\lambda) \rho_s(\lambda) y_0(\lambda) d\lambda. \quad (85)$$

These are the sought-for expressions for the conserved densities in terms of the soliton-gas parameters.

Let us compare expressions (85) with formulas (52), which give the conserved densities $\langle R_n \rangle$ in the initial state. Together with the condition for the satisfaction of all conservation laws for the mean densities (52) and (85), this allows the relationship between $C_s \rho_s(\lambda)$ and the analogous parameter of the initial spectrum to be unambiguously determined:

$$C d(\lambda) = C_s \rho_s(\lambda) y_0(\lambda). \quad (86)$$

One cannot but note the remarkable fact that the identical equality of an infinite set of conserved densities in the initial and statistical states is satisfied when only one relation (86) holds [using only the transformation function $y_0(\lambda)$].

Give relation (86), we can find the solution of Eq. (76) at $k = 0$, i.e., the equations

$$y_0(\lambda) = 1 + \frac{2C_s}{v(\lambda)} \int_{-\omega}^{\omega} l(\lambda, \lambda') \rho_s(\lambda') y_0(\lambda') d\lambda'.$$

We have

$$y_0(\lambda) = 1 + \frac{2C}{v(\lambda)} \int_{-\omega}^{\omega} l(\lambda, \lambda') d(\lambda') d\lambda', \quad (87)$$

where the linear level density C and the spectral density $d(\lambda)$ are specified by the initial conditions (9), i.e., these are fixed. Taking into account normalization (65) of the density $\rho_s(\lambda)$, we determine the linear concentration of soliton phases from (86):

$$C_s = C \int_{-\omega}^{\omega} \frac{d(\lambda)}{y_0(\lambda)} d\lambda. \quad (88)$$

Once C_s and $y_0(\lambda)$ have been calculated, the spectral density in soliton gas can be determined from (86):

$$\rho_s(\lambda) = \frac{C}{C_s} \frac{d(\lambda)}{y_0(\lambda)}. \quad (89)$$

The transformation function $y_0(\lambda)$ is the key element that relates the spectral parameters of the initial and asymptotic statistical states of the system. A similar formula was derived in [4] for the Korteweg–de Vries equation.

6.1. The Transformation Function

Expression (87) for $y_0(\lambda)$ can be simplified by directly expressing $y_0(\lambda)$ in terms of the functions $f(x)$ and $g(x)$ that describe the shape of the local hydrodynamic perturbations of which the initial condition (9) consists. Note that

$$d(\lambda) = AF'(\lambda),$$

where the function

$$F(\lambda) = \begin{cases} -S(\lambda) & \text{for } -\omega \le \lambda \le m_-, \\ 0 & \text{for } m_- \le \lambda \le m_+, \\ S(\lambda) & \text{for } m_+ \le \lambda \le \omega. \end{cases}$$

Therefore, through integration by parts, (87) transforms to

$$y_0(\lambda) = 1 - 2CA \int_{-\omega}^{\omega} \frac{F(s)}{(s-\lambda)\sqrt{\omega^2 - s^2}} ds, \quad (90)$$

where the integral has the meaning of the Cauchy principal value. Before integrating by parts, we must exclude the δ neighborhood of the singularity $\lambda' = \lambda$ from the range of integration and then let $\delta \longrightarrow 0$. Using expression (41) in (90) with allowance for the difference in sign between the functions $F(\lambda)$ and $S(\lambda)$ and changing the order of integration, we obtain

$$y_0(\lambda) = 1 - 2CA \int_{-\infty}^{\infty} dx \int_{D_+(x)} \frac{\sigma(x, s)\sqrt{r_1(x, s)}}{(s-\lambda)\sqrt{\omega^2 - s^2}} ds, \quad (91)$$

where $r_1(x, s)$ is given by (38) and

$$\sigma(x, \lambda) = \begin{cases} -1 & \text{for } -\omega \le \lambda \le \lambda_-(x), \\ 1 & \text{for } \lambda_+(x) \le \lambda \le \omega. \end{cases}$$

The range of integration in the inner integral is the range where $r_1(x, s) > 0$:

$$D_+(x) = (-\omega, \lambda_-(x)) \cup (\lambda_+(x), \omega).$$

The inner integral in formula (91) can be calculated explicitly. To this end, we use the relation

$$\mathrm{Re}\Phi(\lambda + i0) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{Im}\Phi(s + i0)}{s - \lambda} ds$$

between the real and imaginary parts of the function of complex variable $\Phi(z)$ on the real axis, which is analytic in the upper half-plane and has the limit $\Phi(\infty) = 0$. If we are be able to choose a function $\Phi(z)$ for which

$$\mathrm{Im}\Phi(s + i0) = \frac{\sigma(s)\sqrt{r_1(s)}}{\sqrt{\omega^2 - s^2}}$$

at s in the domain D_+ and the imaginary part is zero outside D_+ (the argument x acting as a parameter here was omitted), then the inner integral in (91) will be

$$\pi \mathrm{Re}\Phi(\lambda + i0).$$

It is easy to see that

$$\Phi(z) = \frac{1}{2}\sqrt{\frac{(z - \lambda_-)(z - \lambda_+)}{(z - \omega)(z + \omega)}} + \frac{1}{2}$$

[here, we took into account the fact that $r_1 = (1/4)(\lambda - \lambda_-)(\lambda - \lambda_+)$]. The constant 1/2 was added to satisfy the condition $\Phi(\infty) = 0$ and the branch of the root must be chosen in such a way that the imaginary part of $\Phi(z)$ had the correct sign of $\sigma(s)$ on the real axis. We then obtain

$$\mathrm{Re}\Phi(\lambda + i0) = \frac{1}{2}$$

for λ in the domain D_+ and

$$\mathrm{Re}\Phi(\lambda + i0) = \frac{1}{2} - \frac{1}{2}\sqrt{\frac{(\lambda - \lambda_-)(\lambda_+ - \lambda)}{\omega^2 - \lambda^2}}$$

$$= \frac{1}{2} - \frac{\sqrt{-r_1(x, \lambda)}}{\sqrt{\omega^2 - \lambda^2}}$$

for $\lambda_- \le \lambda \le \lambda_+$.

Substituting the derived expressions for the inner integral in (91) yields the sought-for explicit expression for $y_0(\lambda)$. Finally, we have

$$y_0(\lambda) = 1 - 2\pi CA\int_{-\infty}^{\infty}\left(\frac{1}{2} - \frac{\sqrt{-r_1(x, \lambda)}}{\sqrt{\omega^2 - \lambda^2}}\right)dx$$

in the interval $m_- \le \lambda \le m_+$ and

$$y_0(\lambda) = 1 - \pi CA[x_+(\lambda) - x_-(\lambda)]$$
$$- 2\pi CA\int_{\Delta_-(\lambda)}\left(\frac{1}{2} - \frac{\sqrt{-r_1(x, \lambda)}}{\sqrt{\omega^2 - \lambda^2}}\right)dx \qquad (92)$$

outside the interval $m_- \le \lambda \le m_+$ (but for $-\omega \le \lambda \le \omega$), where $\Delta_-(\lambda)$ is the range outside the segment $[x_-(\lambda), x_+(\lambda)]$ in which $r_1 < 0$.

6.2. The Maximum Soliton Density

The existence of the upper limit C_m follows from relations (92) and (88). It cannot exceed the linear level density C, because, when C approaches C_m from below, $C_s(C)$ indefinitely increases. Indeed, the function $y_0(\lambda)$ reaches a minimum at the ends of the interval $|\lambda| \le \omega$, with its derivative being finite at $\lambda = \pm\omega$. The values of $y_0(\pm\omega)$ can easily be determined from (92) by taking into account the fact that $x_+(\pm\omega) - x_-(\pm\omega) = 1$, according to the definition of $r_1(x, \lambda)$ (38), and that the integral in (92) vanishes at $\lambda = \pm\omega$. We have

$$\min y_0(\lambda) = 1 - \pi CA.$$

As C increases, min y_0 decreases, which, in view of (88), causes the increase in $C_s(C)$ to accelerate. Finally, min y_0 becomes zero at $C = C_m = (\pi A)^{-1}$, the integral in (88) diverges, and $C_s(C) \longrightarrow \infty$ when $C \longrightarrow C_m$. Note that a similar divergence of $C_s(C)$ and, accordingly, the maximum density C_m also arise in the solution of the Korteweg–de Vries equation [4].

The limiting linear spectral density $C_m = (\pi A)^{-1}$ corresponds to a dense packing of wells in the initial state. It follows from the well nonoverlapping condition that the mean distance $\bar{l}$ between the well centers must be larger than the mean well width $\bar{\theta}$, i.e., $\gamma = \bar{\theta}/\bar{l} < 1$; the limiting value $\gamma = 1$ corresponds to a dense packing of wells. Taking into account relation (48), we see that γ tends to unity when $C \longrightarrow C_m$.

6.3. An Example of a Numerical Computation

We tested the equality of the invariant densities in the initial state (9) and in the statistical limit on the first three densities by means of a numerical experiment. As

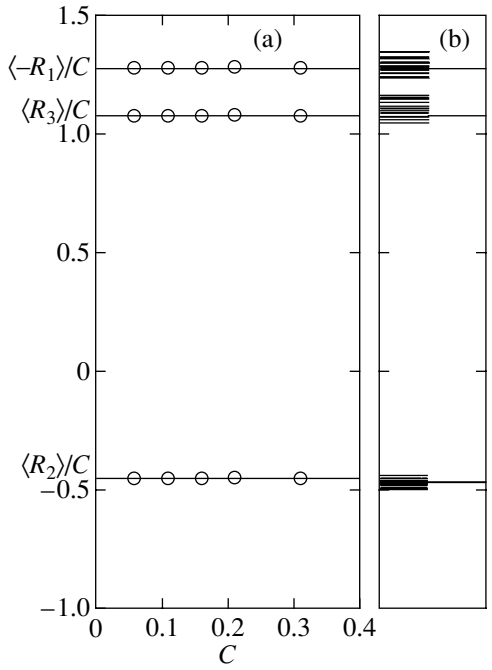

Fig. 3. (a) Comparison of the invariant densities in the initial (horizontal straight lines) and statistical (circles) states. (b) The scatter of means for one realization of $u(x)$ relative to the means for all realizations.

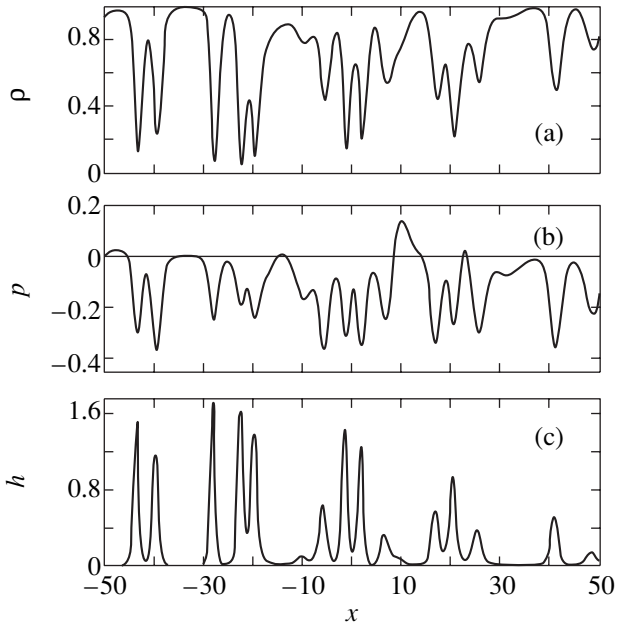

Fig. 4. A typical portion of a specific realization. The first three densities of the conserved quantities are shown: (a) ρ, (b) momentum density $p = \rho v$, and (c) energy density h. The coordinate x is normalized to $\varepsilon/\sqrt{\rho_0}$; the quantities ρ, p, and h are normalized to ρ_0, $\rho_0^{3/2}$, and ρ_0^2, respectively.

the specific model for our computation, we took the initial condition (9) with $\rho_0 = 1$ and the functions

$$f(x) = \begin{cases} a(1-4x^2) & \text{for } |x| \le \frac{1}{2}, \\ 0 & \text{for } |x| > \frac{1}{2}, \end{cases} \qquad (93)$$

$$g(x) = 2\alpha(1-\sqrt{1-f(x)}).$$

The nonnegative parameter $a < 1$ characterizes the well depth relative to the background level ρ_0. The parameter α describes the velocity perturbation; for $-1 \le \alpha \le 1$, condition (39), which provides the structure of the semiclassical spectrum described in Section 4, is satisfied.

According to (52), $\langle R_n \rangle$ are proportional to the linear spectral density C. The proportionality coefficients were computed for the above model (93) with $a = 0.8$ and $\alpha = -0, 8$ at $n = 1, 2, 3$. These are represented in Fig. 3a as the horizontal straight lines.

The circles in Fig. 3a indicate the ratios $\langle R_n \rangle_s / C$ at $n = 1, 2, 3$ at which $\langle R_n \rangle_s$ were determined by averaging the local densities R_n (27) for the N-soliton solution (59) with randomly chosen parameters s_n over x. The N-soliton solution itself was computed by using an algorithm that allows for the existence of a maximizer. Using the maximizer significantly reduces the computational time by discarding the overwhelming majority of the 2^N terms that constitute sum (60). The algorithm is detailed in [4]. In our computation, we took $N = 200$. By way of illustration, Fig. 4 shows plots of the densities $\rho = \rho_0 + R_1$, as well as p and h (29) at $C = 0.2$ for one specific set of numbers s_n. The correspondence of the statistical N-soliton solution (59)–(63) to the initial condition (9) and (93) was achieved by using formulas (88) and (89) to determine C_s and λ_n.

To obtain statistically significant results, we performed an additional averaging over the many (40 in Fig. 3a) realizations (59) determined by a random choice of the set of parameters $s_1, \ldots, s_N$. Figure 3b shows the scatter of means (normalized to C) computed for individual realizations relative to the means for all realizations.

We see from Fig. 3a that there is good agreement between the conserved densities computed for the statistical state and their values in the initial hydrodynamic state. The agreement holds up to C close to the limiting value $C_m = 0.4157$. This shows that using the maximizer to compute the statistical parameters of a turbulent state proves to be very efficient even at maximum densities of the gas of interacting solitons.

ACKNOWLEDGMENTS

We are grateful to S.P. Novikov for a helpful discussion. This study was supported by the INTAS (grant

no. 99-1068) and the Russian Foundation for Basic Research (project no. 00-15-96594l).

REFERENCES

1. V. E. Zakharov and A. B. Shabat, Zh. Éksp. Teor. Fiz. **61**, 118 (1971) [Sov. Phys. JETP **34**, 62 (1971)].
2. S. Novikov, S. V. Manakov, L. P. Pitaevskiĭ, and V. E. Zakharov, *Theory of Solitons: the Inverse Scattering Method* (Nauka, Moscow, 1980; Consultants Bureau, New York, 1984).
3. A. V. Gurevich, K. P. Zybin, and G. A. Él', Zh. Éksp. Teor. Fiz. **115**, 333 (1999) [JETP **88**, 182 (1999)].
4. A. V. Gurevich, N. G. Mazur, and K. P. Zybin, Zh. Éksp. Teor. Fiz. **117**, 797 (2000) [JETP **90**, 695 (2000)].
5. A. V. Gurevich and A. B. Shvartsburg, *The Nonlinear Theory of the Propagation of Radio Waves in the Ionosphere* (Nauka, Moscow, 1973); A. V. Gurevich, *Nonlinear Phenomena in the Ionosphere* (Springer-Verlag, New-York, 1978).
6. Y. Kodama, SIAM J. Appl. Math. **59**, 2162 (1999).
7. F. Dalfovo, S. Giorgini, L. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).
8. I. P. Kornfel'd, Ya. G. Sinaĭ, and S. V. Fomin, *Ergodic Theory* (Nauka, Moscow, 1980).
9. L. D. Faddeev and L. A. Takhtajan, *Hamiltonian Methods in the Theory of Solitons* (Nauka, Moscow, 1986; Sptinger-Verlag, Berlin, 1987).
10. Ya. G. Sinaĭ and E. I. Dinaburg, private communication (1998).
11. G. M. Zaslavsky, *Chaos in Dynamical Systems* (Nauka, Moscow, 1984; Harwood, Chur, 1985).
12. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 5: *Statistical Physics* (Nauka, Moscow, 1986; Pergamon, Oxford, 1980), Part 1.
13. I. Kay and H. E. Moses, J. Appl. Phys. **27**, 1503 (1956).
14. M. J. Ablowitz and H. Segur, *Solitons and the Inverse Scattering Transform* (SIAM, Philadelphia, 1981; Mir, Moscow, 1987).
15. N. G. Mazur, V. V. Geogdzhaev, A. V. Gurevich, and K. P. Zybin, *Proceedings of the Memorial Andronov Conference*, Nizhni Nongorod, 2001.

Translated by V. Astakhov