Search for low-energy neutrino radiation accompanying gamma-ray bursts on the Baksan subterranean scintillation telescope

E. N. Alekseev, L. N. Alekseeva,*) V. N. Zakidyshev, and V. Ya. Poddubnyĭ

Institute of Nuclear Research, Russian Academy of Sciences, 117312 Moscow, Russia
(Submitted 27 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 1921–1929 (December 1998)

An analysis aimed at finding possible neutrino radiation accompanying gamma-ray bursts in a 24-h period about them is performed on the basis of the data in the 4B BATSE Gamma-Ray Burst Catalog and data from the Baksan scintillation telescope according to a program for finding neutrinos from collapsing stars. Values significantly exceeding the background are not discovered. A lower bound for the distance to the source is established under the assumption that the anticipated radiation has characteristics similar to the characteristics of collapse neutrinos. It attests to the cosmological origin of gamma-ray bursts with a high degree of probability. © *1998 American Institute of Physics.* [S1063-7761(98)00112-7]

1. INTRODUCTION

Enormous interest in the phenomenon of gamma-ray bursts has arisen since April 1991, when the Compton Gamma-Ray Observatory was launched. On board this spacecraft there were instruments for studying gamma radiation in various energy ranges: >150 keV (the OSSE instrument), 20–300 keV (the BATSE instrument), 0.75–30 MeV (the COMPTEL instrument), 0.6–170 MeV and 30 MeV–30 GeV (the EGRET instrument) (see Ref. 1 and the references cited therein). Instead of the expected concentration of sources of gamma-ray bursts in the galactic plane, the BATSE data (the most informative and statistically significant data) showed that they are distributed isotropically over the sky, although the intensity distribution of the gamma-ray bursts is spatially inhomogeneous.[2]

In the past seven years of operation, the BATSE instrument has recorded more than 2000 gamma-ray bursts; however, the problem of determining their origin and the distances to them remains open.

As the experimental data were accumulated, the search for various anisotropies, including repeated events from a single source, continued. Some investigators found angular anisotropy for individual groups of bursts[3] and even possible matches between some gamma-ray bursts and x-ray sources,[4] but the analysis of the complete data led to a negative result.[5,6]

The question of the interpretation of the bimodality observed in the duration distribution of the bursts,[7] according to which there are ''short'' bursts with durations lying in the range 0.03–2 s and ''long'' bursts with durations in the range 2–1000 s, remains open.

The photon energies can reach very high values, as was demonstrated by recording the GRB 940217 event with an energy of 18 GeV on the EGRET instrument, which was delayed relative to the BATSE signal[8] and several other events.[9]

The situation became even more complicated when accompanying radiation in the x-ray and optical ranges was detected for several gamma-ray bursts by the VerroSAX satellite, which also monitors bursts, but with better angular accuracy, and a red shift equal to 0.835 was successfully determined for one of them, GRB 970508, from the observed absorption and emission lines (see Refs. 1 and 12 and the references cited therein).

The interpretation of such a large body of experimental data prompted the construction of many models, whose authors clearly adhered either to the galactic or cosmological nature of this phenomenon, although the diversity of observatory data can fully cover the differences in the nature of gamma-ray bursts.

The purely galactic models, which run into difficulties in accounting for the angular isotropy of gamma-ray bursts, introduce an extended galactic corona with a radius of 200–300 kpc, which is filled with neutron stars that left the Milky Way (see, for example Ref. 13). The mechanism for the formation of gamma-ray bursts should be similar to that for x-ray sources.

Conversely, the observed isotropy and inhomogeneity of the distributions of the bursts can easily be explained, if they are at cosmological distances.[14]

It should be noted that both approaches have difficulties in accounting for the entire body of experimental data associated with these phenomena. A detailed investigation of the BATSE signals alone likewise does not permit making an unequivocal choice between the models (see, for example, Ref. 15).

Most of the proposed cosmological models of the formation of gamma-ray bursts are associated with the gravitational collapse of a massive object, during which so much energy is released that the necessary amount of roughly 10^{51} erg per gamma-ray burst is fully provided. It can involve the merger of binary neutron stars or binary black holes etc.,[16–18] the collapse of white dwarfs due to accretion,[19] stellar collapse with the appearance of Type Ib supernovas,[20] the explosion of a very massive rapidly rotating star with the formation of a black hole following collapse

of the core (a so-called ''hypernova''),[21] the collapse of a ''naked'' stellar core in a binary system (the ''minisupernova'' model),[22] or the gravitational collapse of a supermassive star having a mass more than 50 000 times greater than the mass of the Sun into a black hole.[23,24]

Each of these models has its own merits and shortcomings for describing the complete quantitative characteristics of gamma-ray bursts, but they all share one feature, i.e., the mandatory existence of a burst of neutrino radiation. The properties of such neutrinos can be similar to the properties of the neutrinos formed upon core collapse in a massive star with the appearance of a Type II supernova, i.e., the mean energy of the particles ν_e and $\tilde{\nu}_e$ is 10–15 MeV, and the total radiated energy is $\sim 3\times 10^{52}$ erg, but the duration of a neutrino burst can be much shorter, less than 1 s, and the relative fraction of electron antineutrinos $\tilde{\nu}_e$ is several times higher than in the standard model.[17] A few other characteristics of neutrinos appear in the ''supermassive'' star model, viz., a mean energy of 4–8 MeV, a total energy of 10^{58} erg, and a duration less than a several seconds.[24] Nevertheless, the question of the existence of such objects remains open.

Although the radiation associated with gamma-ray bursts can arrive from great distances and the probability of detecting such accompanying radiation from a single event on modern detectors operating simultaneously with the BATSE instrument is very low, with consideration of the broad spectrum of possibilities opened up by the experimentally detected properties of the appearance of gamma-ray bursts just enumerated, several searches for such radiation have been undertaken. In addition, in view of the complexity of the phenomenon under discussion, any additional information is very desirable.

Searches have been undertaken both for neutrino radiation accompanying gamma-ray bursts on the underground IMB,[25] LSD,[26] and LVD[27] detectors and for high-energy photons with an energy above 1 GeV on large instruments for studying extensive atmospheric showers,[28,29] but no positive effect was discovered. It should be noted that the studies performed on the underground detectors did not pass statistical testing very well, since data obtained for 53 gamma-ray bursts observed before October 1990, i.e., before the launch of the Compton Gamma-Ray Observatory (CGRO), were used in the IMB detector; the mass of the target in the LSD instrument used in the analysis was relatively small (30 tons) to reduce the background, and the first part of the LVD instrument began to collect data in the middle of 1993.

The Baksan subterranean scintillation telescope has been performing continuous tracking of our Galaxy for neutrino radiation since 1980. The sensitivity of the instrument is such that it ''examines'' 95% of the galactic stars in searches for single neutrino bursts from the gravitational collapse of the cores of massive stars. The effectiveness of the detector was confirmed by recording the neutrino signal from supernova 1987A BMO along with the IMB and Kamiokande instruments.[30] The ''live-time'' collection efficiency is above 90%, permitting the use of essentially all the data on gamma-ray bursts in the search for possible neutrino radiation accompanying them. The methods for such an analysis of a combined set of telescope data and BATSE data published in the 4B catalog[31] are described in this paper, and a lower bound on the distance to the sources of the gamma-ray bursts is obtained.

2. DETECTOR

The Baksan subterranean scintillation telescope is located in a tunnel in the Northern Caucasus Mountains at a depth of 850 meters of water equivalent (mwe).[32] It consists of eight planes (four vertical planes and four horizontal planes, of which the three lowest are called ''inner'' planes), which form a closed figure that is continuously covered by standard counters filled with the organic scintillator C_9H_{20} and special additives. Each of the counters is viewed by one photomultiplier with a photocathode diameter equal to 15 cm. The total target mass is 330 tons, and the mass of the ''inner'' part is 130 tons. The detection energy threshold of the detectors on the horizontal layers is 8 MeV, the threshold of the detectors on the vertical planes is 10 MeV, the dead time of the instrument is 4 ms, and the accuracy of setting the absolute time is 1 ms.

Since the telescope is at a relatively small depth, to reduce the background from the main source, i.e., cosmic-ray muons which leave a trajectory of triggered counters when they traverse the telescope, only single triggerings of all the counters are recorded on the tape for subsequent treatment in the present program for finding collapse neutrinos. The data supplied from any part of the telescope can be used, depending on the purpose of the work. For example, the ''inner'' 130 tons of the scintillator are employed as a trigger mass in a program for finding collapse neutrinos, i.e., after a signal of interest is found in this mass, the information from the entire telescope is processed.

3. SELECTION OF EVENTS AND METHODS FOR PROCESSING THEM

Although high-intensity neutrino fluxes associated with gamma-ray bursts are predicted in some models, we still do not know *a priori* the detailed parameters of this radiation and the time relationship between these two bursts.

Two types of low-energy neutrinos can be positively detected in the Baksan telescope according to the following reactions with the scintillator material:

$$\tilde{\nu}_e + p \rightarrow n + e^+, \quad E_{\text{th}} = 1.8 \text{ MeV}, \tag{1}$$

$$\begin{aligned} \tilde{\nu}_e + {}^{12}\text{C} &\rightarrow {}^{12}\text{B} + e^+, \quad E_{\text{th}} = 16.827 \text{ MeV}, \\ &\downarrow \\ {}^{12}\text{B} &\rightarrow {}^{12}\text{C} + e^-, \quad t_{1/2} = 0.020 \text{ s}, \end{aligned} \tag{2}$$

$$\begin{aligned} \nu_e + {}^{12}\text{C} &\rightarrow {}^{12}\text{N} + e^-, \quad E_{\text{th}} = 13.880 \text{ MeV}, \\ &\downarrow \\ {}^{12}\text{N} &\rightarrow {}^{12}\text{C} + e^+, \quad t_{1/2} = 0.011 \text{ s}. \end{aligned} \tag{3}$$

In reaction (1) only the positron signal is recorded, i.e., single events accompanying gamma-ray bursts are expected in the detectors. In reactions (2) and (3) two signals, i.e., electron and positron signals, can be detected in time inter-

vals equal to three beta decay periods of ^{12}B and ^{12}N, which are equal to $\sim$100 ms and $\sim$50 ms, respectively. The search for double signals accompanying gamma-ray bursts significantly improves the signal-to-background ratio, especially with consideration of the upper energy edges of the decay spectra (13.37 MeV for ^{12}B and 16.4 MeV for ^{12}N).

The following constraints were introduced into the telescope data for the subsequent treatment.

1) Periods of operation of the instrument when all the systems functioned normally were selected. For this reason only 1198 of the 1736 bursts in the 4B catalog, among which there is information on the duration for only 1234, were used.

2) Only events from detectors on the ''inner'' planes, which are most strongly shielded from the background of cosmic-ray muons, were selected. It was decided that each of these detectors would be triggered no more than three times in a 24-h period (the averaged counting rate of the ''inner'' detectors is one count per 24 h). This condition eliminated any detectors with increased inherent background from the analysis.

3) Events whose amplitude did not exceed half of the energy released by a relativistic particle in a standard detector were selected. Thus, the range of allowed energies was 8–25 MeV.

As a result, the data supplied from a telescope target mass equal to 126 tons with a mean counting rate of single pulses roughly equal to 0.012 events per second were used in the subsequent work.

The following information from the 4B catalog was used.

1) The treatment was performed using durations equal to both τ_{50} and τ_{90} [τ_{50} (τ_{90}) is the time interval during which 50% (90%) of the gamma-quantum flux is detected].

2) The gamma-ray bursts were separated into four groups according to duration: 0–1 s, 1–10 s, durations $>$10 s, and all durations.

3) The bursts with a total energy $>$20 keV determined by summation over all the channels were separated into three groups according to intensity: ''strong'' with a flux $I>10^{-6}$ erg/cm^2, ''weak'' with a flux $I<10^{-6}$ erg/cm^2, and a mixed group.

Then the events selected and separated in this manner were processes by several methods.

Since only single pulses from positrons would be detected as a result of reaction (1) and since we know nothing *a priori* amount the time relationship between gamma-ray bursts and the postulated neutrino radiation, the combined data were treated on the basis of the following assumptions.

A. It was assumed that gamma-ray bursts are of the same nature, regardless of their individual parameters; therefore, it could be expected that if neutrino radiation exists, it is located at approximately the same time distance on the time axis from a burst. Then, we can look for statistically significant upward deviations from the background for both the total set of bursts and various selected groups by summing the corresponding telescope data in a 24-h period centered at the position of the burst onset and analyzing these data in different time windows with durations from 1 s to 10 min.

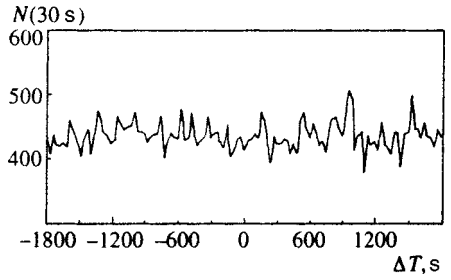

FIG. 1. Distribution of single telescope events in a 30-s window $N(30\ \text{s})$ in a $\pm$1800-s interval about 1198 gamma-ray bursts juxtaposed according to the onset of the signals: ΔT is the time from the onset of the bursts.

Such an analysis of the information was performed, but no values significantly exceeding the background were found. As an example, Fig. 1 shows the typical form of the distribution of the number of single telescope events in a 30-s window summed over all 1198 bursts (the vertical axis) in a 3600-s interval about a burst (the horizontal axis). The figure clearly demonstrates the absence of the expected signal.

B. It was assumed that gamma-ray bursts consist of signals of different origin. Then, if neutrino radiation exists, it can arrive at different times relative to a burst. In this case we can analyze the time intervals between pairs of telescope events and look for statistically significant upward deviations from the background by summing the corresponding telescope data for all the gamma-ray bursts or for various selected groups. The telescope data for 24-h intervals about the bursts were treated by such a method. Figure 2 shows the result of such a treatment of the data for the case of ''all bursts.'' The length of the interval between events is plotted along the horizontal axis, the zero point on the axis corresponding to the onset of the gamma-ray bursts. The two branches in the figure reflect the distributions found ''before'' and ''after'' the gamma-ray bursts.

It can be seen that the distribution of the time intervals between telescope events obtained by this method does not have any special features and that everything is described well by the corresponding Poisson formula.

The results presented were obtained with single telescope events, which are characterized by a relatively high background level. The signal-to-background ratio for the

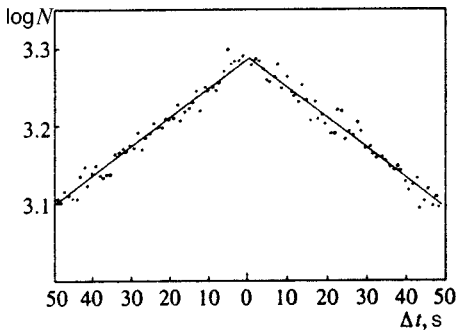

FIG. 2. Distribution of the time intervals between single telescope events $N(\Delta t)$ summed over 1198 gamma-ray bursts in a 24-h interval about them. The center of the horizontal axis corresponds to the position of the onset of the gamma-ray signals; $\Delta t=t_{i+1}-t_i$. Straight lines — result of a calculation using Poisson's formula.

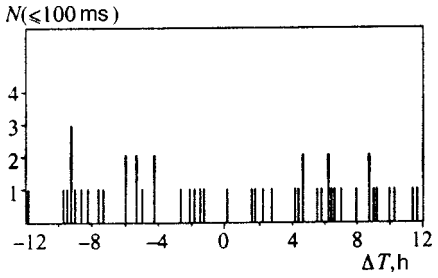

FIG. 3. Distribution of pairs of telescope events with a time interval $\leqslant 100$ ms between them in a 600-s window $N(\leqslant 100$ ms) summed over 1198 gamma-ray bursts during a 24-h period about them. The position of the onset of the gamma-ray bursts is at $\Delta T=0$.

products of reactions (2) and (3) is considerably better. A search for a possible neutrino signal was performed using paired telescope events separated by a time interval <100 ms according to the method described in paragraph A. Figure 3 shows the result of such a treatment of the data for the case of ''all bursts.'' The time on both sides of the bursts is plotted along the horizontal axis, and the number of paired events in a <100-ms window, which is summed over 600-s intervals because of their small number, is plotted along the vertical axis. The paired events are distributed evenly about the bursts, and there are values exceeding the background.

Thus, all the proposed methods for processing the information in a search for neutrino radiation in the Baksan telescope that can be associated with gamma-ray bursts from the BATSE 4B catalog yielded negative results.

4. CONCLUSION

The data obtained permit finding a lower bound on the distance to the sources of gamma-ray bursts within the model of accompanying neutrino radiation considered. Since the duration of the neutrino radiation and its position on the time axis relative to the gamma-ray bursts for each case are unknown, to find the bound we must start out from very general assumptions regarding these parameters, i.e., we must compare the total numbers of experimentally recorded and theoretically expected telescope events in different time intervals from the onset of the gamma-ray bursts.

Figure 4 shows the integral distribution of single background telescope events summed over all 1198 bursts (the horizontal axis) as a function of the time interval ''before'' (the left-hand curve) and ''after'' (the right-hand curve) the onset of the gamma-ray bursts, which is located at zero (the vertical axis). It can be seen in this figure that the telescope background is absolutely symmetric (does not have any features) relative to the center of the axis. Positron signals from the interaction of electron antineutrinos in reaction (1), which is detected with high reliability in the telescope, would be expected specifically among these events. Such signals were not discovered; therefore, the data shown in Fig. 4 were used to estimate the distances to the sources, and it was assumed that the expected neutrino radiation has the same characteristics as neutrinos from collapsing cores of massive stars, i.e., the total neutrino energy equals 3×10^{53} erg, the particles have thermal Fermi–Dirac spectra

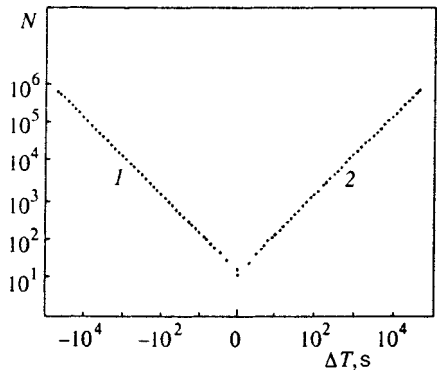

FIG. 4. Integral distribution of single telescope events in a 24-h period about 1198 gamma-ray bursts $N(\Delta T)$ as a function of the time interval ''before'' (*1*) and ''after'' (*2*) the onset of the gamma-ray signals. The zero point corresponds to the position of the onset of the gamma-ray bursts.

with a temperature of 3.5 MeV (the cases of 4 MeV and 4.5 MeV were also considered), and all the types of neutrinos are produced in equal numbers. The latter assumption somewhat worsens the estimate sought for the telescope in comparison to the large fraction of electron antineutrinos predicted in Ref. 17, but this is not very significant in our case of a more general treatment.

The lower bound on the distance to the sources of gamma and neutrino radiation obtained at the 90% confidence level is shown in Fig. 5, where the distance is plotted on the vertical axis and time intervals ''before'' (''after'') the onset of the gamma-ray bursts from 1 s to 12 h are plotted along the horizontal axis. Curve *1* represents the case where the ''temperature'' of the thermal neutrinos is equal to 3.5 MeV, curve *2* is for a neutrino ''temperature'' equal to 4 MeV, and curve *3* is for the case of 4.5 MeV per burst.

As can be seen in this figure, the bound obtained for any position of the neutrino signal relative to the gamma-ray burst attests to the extragalactic nature of the sources of the gamma-ray bursts, and the bound on the distance for neutrino and gamma radiation that are close in time essentially reaches 1 Mpc.

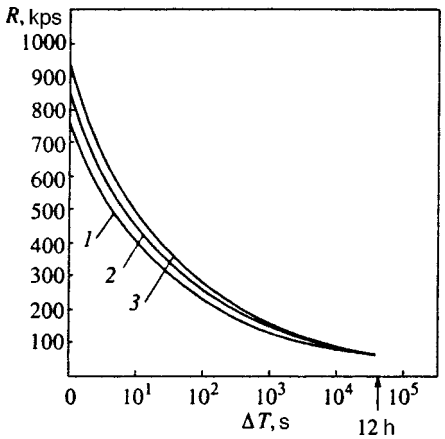

FIG. 5. Dependence of the lower bound on the distance R to the sources of gamma-ray bursts and neutrinos on the time interval ΔT before the onset of the gamma-ray signals. The different curves corresponds to different ''temperatures'' of the expected neutrinos: *1* — 3.5 MeV, *2* — 4 MeV, *3* — 4.5 MeV.

Thus, the analysis performed on data from the Baksan subterranean scintillation telescope that were recorded according to a program for searching for neutrino bursts from collapsing stars together with data in the 4B catalog on gamma-ray bursts recorded by the BATSE instrument on the CGRO did not yield a positive effect. The results of this work show that if the formation of the gamma radiation in sources of gamma-ray bursts is, in fact, accompanied by high-intensity neutrino radiation or if such neutrino radiation precedes the formation of gamma radiation, these sources have an extragalactic nature and are located at distances greater than 1 Mpc.

We thank Prof. I. G. Mitrofanov for some useful discussions regarding the subject of this research.

This work was supported by the Russian Fund for Fundamental Research (Project 95-02-04660-a).

*)E-mail: alexeyev@ms1.inr.ac.ru, alexeev@niutr.novoch.ru

[1] J. Greiner, M. Sommer, N. Bade *et al.*, Astron. Astrophys. **302**, 121 (1995).
[2] G. J. Fishman, Publ. Astron. Soc. Pac. **107**, 1145 (1995).
[3] B. Link and R. Epstein, E-print archive, astroph/9704244 (1997).
[4] K. Hurley, P. Li, A. Smette *et al.*, Astrophys. J. **464**, 342 (1996).
[5] M. Briggs, W. S. Paciesas, and G. N. Pendleton, Astrophys. J. **459**, 40 (1996).
[6] R. M. Kippen, J. M. Ryan, A. Connors *et al.*, Astrophys. J. **492**(1), (1998).
[7] C. Kouveliotou, C. A. Meegan, G. J. Fishman *et al.*, Astrophys. J. Lett. **41**, L101 (1993).
[8] K. Hurley, Nature (London) **372**, 652 (1994).
[9] B. L. Dingus, J. R. Catelli, and E. J. Schneid, in *Proceedings of the 25th ICRC, Durban, South Africa (1997)*, Vol. 3, p. 29.
[10] F. Frontera, E. Costa, D. Dal Fiume *et al.*, in *Proceedings of the 25th ICRC, Durban, South Africa (1997)*, Vol. 3, p. 25.
[11] M. G. Baring, in *Proceedings of the 5th Workshop on TeV Gamma-Ray Astrophysics*, Kruger National Park, South Africa (1997).
[12] D. E. Reichart, E-print archive, astroph/9801139, submitted to Astrophys. J. Lett.
[13] D. Q. Lamb, Publ. Astron. Soc. Pac. **107**, 1152 (1995).
[14] B. Paczynski, Publ. Astron. Soc. Pac. **107**, 1167 (1995).
[15] I. G. Mitrofanov, M. L. Litvak, and D. A. Ushakov, E-print archive, astroph/9707128 (1997), Astrophys. J (in press).
[16] S. I. Blinnikov, I. D. Novikov, T. V. Perevodchikova, A. G. Polnarev, Pis'ma Astron. Zh. **10**, 422 (1984) [Sov. Astron. Lett. **10**, 177 (1984)].
[17] M. Ruffert, H.-Th. Janka, K. Takahashi, and G. Schafer, Astron. Astrophys. **319**, 122 (1997).
[18] J. R. Wilson, J. D. Salmonson, and G. J. Matthews, in *Proceedings of the 4th Huntsville Gamma-Ray Burst Symposium*, Huntsville, September (1997).
[19] V. V. Usov, Nature (London) **357**, 472 (1992).
[20] S. E. Woosley, Astrophys. J. **405**, 273 (1993).
[21] B. Paczynski, in *Proceedings of the 4th Huntsville Gamma-Ray Burst Symposium*, Huntsville, September (1997); E-print archive astroph/9712123 (1997).
[22] S. I. Blinnikov and K. A. Postnov, Mon. Not. R. Astron. Soc. **293**, L29 (1998).
[23] G. M. Fuller and X. Shi, E-print archive, astroph/9711020 (1997), submitted to Astrophys. J.
[24] X. Shi and G. M. Fuller, E-print archive, astroph/9801106 (1998), submitted to Astrophys. J.
[25] R. Becker-Szendy, C. B. Bratton, J. Breault *et al.*, in *Proceedings of the 23th ICRC*, Calgary, USA (1993), Vol. 4, p. 476.
[26] M. Aglietta, P. Antonioli, G. Badino *et al.*, Astrophys. Space Sci. **231**, 355 (1995); in *Proceedings of the 24th ICRC*, Rome, Italy (1995), Vol. 2, p. 73.
[27] M. Aglietta, B. Alpat, E. D. Alyea *et al.*, in *Proceedings of the 24th ICRC*, Rome, Italy (1995), Vol. 1, p. 662.
[28] L. Padella, B. Funk, H. Krawczynski *et al.*, in *Proceedings of the 25th ICRC*, Durban, South Africa (1997), Vol. 3, p. 57.
[29] S. D. Bathelmy, P. Butterworth, R. Cabrera *et al.*, in *Proceedings of the 25th ICRC*, Durban, South Africa (1997), Vol. 3, p. 73.
[30] L. N. Alexeyeva, Highlights Astron. **8**, 229 (1989).
[31] Catalog 4B, http://www.batse.msfc.nasa.gov/data/grb/4bcatalog.
[32] E. N. Alekseev, L. N. Alekseeva, V. I. Volchenko *et al.*, Zh. Eksp. Teor. Fiz. **104**, 2897 (1993) [JETP **77**, 339 (1993)].

Translated and edited by P. Shelnitz

Coherent states of potentials of soliton origin

B. F. Samsonov*)

Tomsk State University, 634050 Tomsk, Russia
(Submitted 10 June 1998)
Zh. Éksp. Teor. Fiz. **114**, 1930–1943 (December 1998)

The spectral properties of the most general time-dependent potentials of the soliton type described by a self-adjoint operator acting in Hilbert space are discussed. The spectral decomposition for these potentials and the quasispectral decomposition for the Darboux transformation operators are obtained. The coherent states of such systems are examined. Finally, the measure realizing the decomposition of the identity operator in the projectors on the coherent states is calculated. © *1998 American Institute of Physics.* [S1063-7761(98)00212-1]

1. INTRODUCTION

The first to introduce coherent states (for the case of the harmonic oscillator) as nonspreading wave packets was Schrödinger.[1] Today such states have found wide application in various areas of physics and mathematics.[2–4] The different aspects of coherent states have been discussed in numerous reviews (see, e.g. Refs. 5–7 and the monographs by Perelomov,[2] Malkin and Man'ko,[3] and Klauder and Skagerstam[4]).

However, it must be noted that at present there is no unified definition of a coherent state, and different authors mean different things when speaking about such states. An analysis of the existing definitions shows that there are several characteristic properties of such states (described by vectors $\psi_z(x,t)$, which can be taken as their definitions) that can be used to define them:[8] (1) the vectors $\psi_z(x,t)$ are elements of the Hilbert space H of the states of the system; (2) the parameter z takes continuous values from a domain $\mathscr{D}$ of an n-dimensional complex space; (3) the vectors $\psi_z(x,t)$ are stable in time; and (4) there is a measure $\mu=\mu(z,\overline{z})$ (the bar over the symbol indicates complex conjugation) realizing the decomposition of the identity operator $\mathscr{T}$ acting in H,

$$\int_{\mathscr{D}} d\mu\,|\psi_z\rangle\langle\psi_z|=\mathscr{T}. \tag{1}$$

By temporal stability we mean that the states described by the vectors $\psi_z(x,t)$ remain coherent at all times, i.e., satisfy the properties 1, 2, and 4. To satisfy this condition, we will assume that the functions $\psi_z(x,t)$ are solutions of the Schrödinger equation

$$(i\partial_t-h_0)\psi_z(x,t)=0,$$

where h_0 is the Hamiltonian (not necessarily a time-independent one) acting in H.

If a quantum system has nontrivial symmetry properties, the study of coherent states for such a system simplifies, since group-theoretic methods prove to be highly effective.[2,6] Lately there has been an upsurge of interest in quantum systems whose differential symmetry operators do not form a closed algebra. They can form quadratic,[9,10] polynomial,[11,12] and q-deformed[13,14] algebras. In the latter case the concept of q-coherent states is introduced.[15,16] A characteristic feature of many such systems is that they are generated by a chain (possibly infinite) of Darboux transformations[17,18] from simpler systems, such as free particles, the Coulomb potential, a harmonic oscillator, etc.[19,20] The coherent states of such systems have been studied very little. In this connection the work of Fernandez *et al.*[21,22] deserves mention. These researchers studied the coherent states of isospectral Hamiltonians with an equidistant spectrum. Spiridonov[23] studied systems of the coherent states of q-deformed potentials of the harmonic oscillator and a class of self-similar potentials. Finally, Man'ko *et al.*[24] studied a generalization of the q-oscillator, which was called an f-oscillator, and obtained the coherent states for such an oscillator (f-coherentstates).

In those cases when the system considered may be generated by using the technique of the Darboux transformation operators, the coherent states for this system can be obtained by applying a transformation operator to the coherent states of the initial system (if they are known).[25,26] Here, to satisfy condition 3, we must have a generalization of this transformation to the time-dependent Schrödinger equation. One variant of such a generalization has been developed in soliton theory.[27] However, often this approach fails because it may lead to complex-valued potentials, which would violate condition 1. The method developed in Ref. 28 and thoroughly described in Ref. 19 is free of this drawback. It has been used to obtain and study coherent states of anharmonic oscillator Hamiltonians with a quasiequidistant spectrum,[25] which were first obtained in Refs. 11 and 12, coherent states of anharmonic oscillator Hamiltonians with an equidistant spectrum,[26] which were first obtained in Ref. 29, coherent states of anharmonic potentials with a singularity at zero of the form γx^{-2} (Ref. 30), and coherent states of the one-soliton time-independent[31] and time-dependent[32] potentials. Note that in the cases of equidistant and quasiequidistant spectra such states are described by nonspreading (in time) wave packets,[11] which earlier were known to exist only in systems with a Hamiltonian quadratic in the spatial coordinate.

A characteristic feature of time-independent soliton potentials[27] is their transparency. A particle scattered by such a potential is never reflected by it. Another important prop-

erties of such potentials is that each level in the discrete spectrum always occupies a preassigned position, which is controlled by the parameters of the potential. Because of this property, they can be used as model potentials is pseudopotential theories. For instance, they were used to describe relaxation processes in Fermi liquids.[33] A generalizations of such potentials is the time-dependent soliton potential introduced in order to solve the Kadomtsev–Petviashvili equation.[34] Note, however, that such potentials of the general type[27,34] are complex-valued and, therefore, no self-adjoint operator in Hilbert space can be associated with them.

In this paper we will use the time-dependent Darboux transformation[19] to derive fairly simple expressions for real-valued time-dependent many-soliton potentials. We will establish the relationship that exists between the Darboux transformation operators and the spectral decomposition of the momentum operator and will for the first time obtain the decompositions of these operators in the quasiprojectors on the eigenvectors of the momentum operator. Two types of coherent states will be examined. The states of the first type, obtained via an integral transformation operator, allow for the decomposition (1) of the identity operator in which the measure is a continuous function. In the second case, the coherent states are determined via a differential transformation operator, and the measure in (1) is determined by a generalized function.

2. TIME-DEPENDENT MANY-SOLITON POTENTIALS

One of the methods used in building time-dependent many-soliton potentials is described in Ref. 35, which also contains a description of many physical applications. Generally, such potentials are complex-valued, and no self-adjoint operator in the Hilbert space can be associated with them. In this section we will describe a method more suitable from the quantum mechanical viewpoint. The method is a direct generalization of the well-known constructs used in the theory of the time-independent Schrödinger equation.[36–38] The main idea can be illustrated by the example of a two-soliton potential, while for the general case we give only the final expressions.

To obtain a two-soliton time-dependent potential, we must perform two Darboux transformation in succession.[19] The operator L of the Darboux transformation is determined by a solution $u=u(x,t)$ of the initial Schrödinger equation (in our case this is an equation with a zero potential), $L=-u_x/u+\partial_x$ ($u_x\equiv\partial u/\partial x$ and $\partial_x\equiv\partial/\partial x$). If assume that

$$u=u_1=\exp\eta_1\cosh\theta_1,$$

$$\eta_1=i(\mu_1^2-\lambda_1^2)t-i\lambda_1 x,\quad \theta_1=\mu_1 x+2\lambda_1\mu_1 t,$$

where λ_1 and μ_1 are arbitrary real numbers, we arrive at the well-known one-soliton potential $V^{(1)}=-2\mu_1^2\operatorname{sech}^2\theta_1$. All the solutions of the Schrödinger equation with the potential $V^{(1)}$ except the one that belongs to the kernel of the operator $L^\dagger=-\overline{u_x}/\overline{u}-\partial_x$ can be obtained by applying L to a solution of the Schrödinger equation with a zero potential. In particular, let us take the following solution of the equation with a zero potential: $u_2=\exp\eta_2\sinh\theta_2$, where η_2 and θ_2 differ from η_1 and θ_1 by λ_1 and λ_2 being replaced by μ_1 and μ_2. Then the function

$$v_2=Lu_2=i(\lambda_1-\lambda_2)u_2+W_2\exp\eta_2\operatorname{sech}\theta_1,$$

where

$$W_2=\tfrac{1}{2}[(\mu_2+\mu_1)\cosh(\theta_1-\theta_2)+(\mu_2-\mu_1)\cosh(\theta_1+\theta_2)]$$

is the Wronskian of the functions $\tilde{u}_1=\cosh\theta_1$ and $\tilde{u}_2=\sinh\theta_2$, is a solution of the Schrödinger equation (which is not square-integrable) with the potential $V^{(1)}$. If we take v_2 as the transformation function for the second Darboux transformation, we arrive at the most general (complex-valued) two-soliton potential. (This is simply another form of the well-known two-soliton solution described in Ref. 34.) To obtain a real-valued potential, we must require that the transformation function v_2 obey the reality condition specified in Ref. 19: $[\ln(v_2/\overline{v_2})]_{xxx}=0$. This condition is met if we put $\lambda_1=\lambda_2=\lambda$. In this case the two-soliton potential becomes real:

$$V^{(2)}=-2(\ln W_2)_{xx}=-2(\mu_2^2-\mu_1^2)W_2^{-2}\times(\mu_2^2\cosh^2\theta_1+\mu_1^2\sinh^2\theta_2).$$

If in addition we require $\mu_2>\mu_1>0$, the Wronskian W_2 will retain its sign for all real values of x and t, and the potential $V^{(2)}$ will be a regular function for all x and t. Note that at $\lambda=0$ this potential becomes the well-known two-soliton time-independent potential.[27]

We will now take a chain of N Darboux transformations similar to the one considered for the case with $N=2$. The action of such a chain of Darboux transformations is equivalent to the action of an order-N transformation operator:[19]

$$L=W^{-1}(u_1,u_2,\ldots,u_N)\begin{vmatrix} u_1 & u_2 & \cdots & 1\\ u_{1x} & u_{2x} & \cdots & \partial_x\\ \vdots & \vdots & \ddots & \vdots\\ u_{1x}^{(N)} & u_{2x}^{(N)} & \cdots & \partial_x^N\end{vmatrix},\quad (2)$$

where $u_{kx}^{(N)}=\partial^N u_k/\partial x^N$, and the operator determinant is interpreted as a differential operator obtained by decomposing the determinant with respect to the last column with function coefficients in front of the differentiation operators. The functions u_k in our case are

$$u_1=\exp\eta_1\cosh\theta_1,\quad u_2=\exp\eta_2\sinh\theta_2,$$

$$u_3=\exp\eta_3\cosh\theta_3,\quad u_4=\exp\eta_4\sin\theta_4\ldots,$$

$$\eta_k=i(\mu_k^2-\lambda^2)t-i\lambda x,\quad \theta_k=\mu_k x+2\lambda\mu_k t,$$

$$\lambda,\mu_k\in\mathbb{R},\quad 0<\mu_{k-1}<\mu_k.\quad (3)$$

The N-soliton potential is given by the expression

$$V^{(N)}=-2[\ln W(u_1,u_2,\ldots,u_N)]_{xx}.$$

Clearly,

$$[\ln W(u_1,u_2,\ldots,u_N)]_{xx}=[\ln W(\tilde{u}_1,\tilde{u}_2,\ldots,\tilde{u}_N)]_{xx},$$

where $\tilde{u}_1=\cosh\theta_1$, $\tilde{u}_2=\sinh\theta_2$, $\tilde{u}_3=\cosh\theta_3$, ...

Note that the Wronskian $W(\tilde{u}_1,\tilde{u}_2,\ldots,\tilde{u}_N)$ is expressed by the well-known formula[38]

$$W(\tilde{u}_1,\tilde{u}_2,\ldots,\tilde{u}_N)=2^{1-N}\sum_{k=1}^{2^{N-1}}B_k\cosh\gamma_k. \tag{4}$$

All the coefficients γ_k are of the form

$$\gamma_k=\sum_{i=1}^{N}\vartheta_i,\quad \vartheta_i=\varepsilon\theta_i,\quad \varepsilon=\pm 1.$$

The various terms in the sum (4) break down into groups. For N even we have $N/2+1$ groups and for N odd, $(N+1)/2$ groups. Each group except the last contains $\binom{N}{k}$ terms, where $k=0,1,\ldots,N/2-1$ for N even and $k=0,1,\ldots,(N-3)/2$ for N odd. The last group contains $\frac{1}{2}\binom{N}{N/2}$ terms for N even and $\binom{N}{(N-1)/2}$ for N odd. One group differs from another by the number of negative values of the parameter ε. In the first group (it contains only one term) all $\varepsilon=1$, in the second group (N terms) there is only one $\varepsilon=-1$ in each term, in the third group there are two $\varepsilon=-1$ in each term, etc. For the coefficients B_k we have

$$B_k=\prod_{i>j}^{N^2}|m_i-m_j|,\quad m_i=\varepsilon\mu_i.$$

Note that after we have calculated the absolute value of the differences m_i-m_j in B_k for $\mu_i>\mu_j$, formula (4) becomes valid for arbitrary values of μ_i. We also note that in the particular case $\lambda=0$ we have the N-soliton time-independent potential.[27]

In Sec. 4 we will show that such a potential can be associated with a self-adjoint operator in Hilbert space.

3. HILBERT SPACE OF THE STATES OF A FREE PARTICLE

In accordance with property 1 formulated in the Introduction, we must set up a Hilbert space of the states of a soliton potential. According to Ref. 19, to achieve this we can use the Hilbert space of the states of the initial system (in our case, a free particle) and apply the Darboux transformation operator to its elements. (Here we will not discuss the exact domains of the operators involved.)

Using the solutions of the Schrödinger equation for a free particle to build a Hilbert space is a well-known procedure (see, e.g., the monograph by Miller[39] and the references cited therein). In this section we will only mention the properties needed in our further discussion.

The symmetry algebra of the Schrödinger equation with a zero potential is a six-dimensional Lie algebra, known as the Schrödinger algebra.[39] In particular, the operators

$$a=(i-t)\partial_x+ix/2,\quad a^\dagger=(i+t)\partial_x-ix/2$$

form a Heisenberg–Weyl subalgebra of this algebra. Square-integrable solutions of the Schrödinger equation can be obtained by separation of variables[39] if for the symmetry operator separating the variables we take $K_0=aa^\dagger+a^\dagger a$. Then the orthonormalized wave functions have the form

$$\psi_n(x,t)=(-i)^n(n!\,2^n\sqrt{2\pi})^{-1/2}(1+it)^{-1/2}\times\exp\left[-in\arctan t-\frac{x^2}{4(1+it)}\right]H_n\left(\frac{x}{\sqrt{2+2t^2}}\right),$$

where $H_n(z)$ are the Hermite polynomials. For these wave functions the operators $a^\dagger$ and a are the shift operators in the variable n: $a\psi_n=\sqrt{n}\,\psi_{n-1}$, $a^\dagger\psi_n=\sqrt{n+1}\,\psi_{n+1}$, and $a\psi_0=0$.

By $\mathscr{L}_0$ we denote the linear space comprised of various finite linear combinations of the functions $\psi_n(x,t)$ with complex-valued coefficients. In view of their linearity, the operators a and $a^\dagger$ are defined in the entire space $\mathscr{L}_0$ and map this space into itself. Note that the Hamiltonian $h_0=p_x^2=-\partial_x^2$ and the momentum operator $p_x=-i\partial_x$ can be expressed in terms of a and $a^\dagger$, e.g., $p_x=-(a+a^\dagger)/2$. Hence these operators are also defined in $\mathscr{L}_0$ and map this space into itself.

We use the ordinary Lebesgue integral to define the scalar product $\langle\psi_a|\psi_b\rangle$ on $\mathscr{L}_0$. The Hilbert space H is the completion of $\mathscr{L}_0$ in the norm defined by this scalar product. It is well known that the operators p_x and h_0 are self-adjoint in H, have common eigenfunctions (in the sense of generalized functions), and that their spectrum is purely continuous. We denote these common eigenfunctions by $\psi_p=\psi_p(x,t)$, i.e., $p_x\psi_p=p\psi_p$ and $h_0\psi_p=p^2\psi_p$, $p\in\mathbb{R}$. The functions ψ_p have been thoroughly studied and we will not give their explicit form here. Note, however, that they are orthogonal to each other and normalized (in the sense of generalized functions), $\langle\psi_p|\psi_q\rangle=\delta(p-q)$, and form a complete set in H. Symbolically the completeness property can be expressed in terms of projectors onto these functions as on elements of H:

$$\int dp\,|\psi_p\rangle\langle\psi_p|=1.$$

We do not indicate the limits of integration in integrals along the entire real axis. Note that all these constructs can by substantiated with sufficient mathematical rigor if one uses the concept of a nested Hilbert space.[40,41]

The coherent states of our system can be obtained by applying the operator of translation in the Heisenberg–Weyl group to the vacuum vector ψ_0 (see Ref. 2):

$$\psi_z(x,t)=\exp(za^\dagger-\bar{z}a)\,\psi_0(x,t),\quad z\in\mathbb{C}.$$

These vectors are the eigenvectors of the operator a, i.e. $a\psi_z=z\psi_z$. To describe these vectors we need something more than the space $\mathscr{L}_0$, since they belong to a broader domain densely filling H. The decomposition of these vectors in the basis $\{\psi_n\}$ has the form[3]

$$\psi_z=\Phi\sum_n a_nz^n\psi_n,\quad \Phi=\Phi(z,\bar{z})=\exp\left(-\frac{z\bar{z}}{2}\right),$$

$$a_n=\frac{1}{\sqrt{n!}},\quad z\in\mathbb{C}. \tag{5}$$

The vectors satisfy all the properties enumerated in the Introduction. In particular, the measure $d\mu = d\mu(z,\overline{z})$ realizing the decomposition (1) of the identity operator in theses states is well known:[3] $d\mu = dx\,dy/\pi$, $z = x+iy$.

4. THE SPECTRAL PROPERTIES OF SOLITON POTENTIALS

The structure of the spectrum of the Hamiltonians connected by the Darboux transformation has been studied by many researchers. Probably the most thorough investigation of this problem is that of Veselov and Shabat,[42] where the reader can also find references to the most important on the subject. In the present section we will study the properties of the Darboux transformation operators for a free particle and soliton Hamiltonians as operators acting in the Hilbert space of the states of the system that have not been examined before.

As noted earlier, a many-soliton Hamiltonian can be constructed via a chain of Darboux transformations or the N-order Darboux transformation operator (2). We denote the operator corresponding to the first link in the chain by L_1 and the operator that is the formal conjugate of L_1 by $L_1^\dagger$. Note that in the definition of the scalar product in terms of the Lebesgue integral, the operator that is the conjugate of ∂_x with respect to the scalar product is $-\partial_x$. Hence the operator $L_1^\dagger$ is the Hermitian conjugate of L_1. Here we will not clutter up our discussion with mathematical detail and describe the domains of these operators. Suffice it to say that they are defined not in the entire space H and the range of their values also differs from H. Moreover, it can be shown that if the initial domain, e.g., L_1, is the space $\mathscr{S}_0$, it always has a closed extension into H.

Simple calculations show that

$$L_1^\dagger L_1 = (p_x+\lambda)^2+\mu_1^2 \equiv g_{01}(\mu_1).$$

The operator $g_{00}=(p_x+\lambda)^2$ is the symmetry operator of the Schrödinger equation for a free particle, and the transformation function u_1 ($\notin H$) is an eigenfunction of this operator with an eigenvalue equal to $-\mu_1^2$. Being an operator in Hilbert space H, it is defined in a certain dense domain (which, if necessary, can be determined more exactly), is self-adjoint in this space, and has a purely continuous spectrum. Its eigenfunctions coincide with ψ_p, i.e., $g_{00}\psi_p=(p+\lambda)^2\psi_p$. We also note that $g_{00}=h_0$ at $\lambda=0$.

Clearly, all the transformation functions u_k in (3) are eigenfunctions of g_{00}, i.e., $g_{00}u_k=-\mu_k^2 u_k$. This is true only if we select all the parameters λ of the transformation functions u_k equal. Here the N-order operator $L^\dagger L$ is also a self-adjoint operator in H. This property is ensured by the fact that the operators L and $L^\dagger$ have the following remarkable factorization property:[19] $L^\dagger L = f(g_{00})$, where $f(x)$ is a polynomial, i.e., $f(x)=(x+\mu_1^2)\times(x+\mu_2^2)\cdots(x+\mu_2^N)$.

For the operator $g_0 = L^\dagger L = g_{01}(\mu_1)g_{01}(\mu_2)\cdots g_{01}(\mu_N)$ we can easily obtain its matrix representation in the basis $\{\psi_n\}$. Since the operator $g_{01}(\mu_1)$ can be expressed in terms of the ladder operators a and $a^\dagger$,

$$g_{01}(\mu_1)=\left[\lambda-\frac{a+a^\dagger}{2}\right]^2+\mu_1^2,$$

the matrix

$$S_0(\mu_1)=\|S_{nk}^0\|, \quad S_{nk}^0=\langle\psi_n|g_{00}|\psi_k\rangle$$

is a pentadiagonal symmetric real-valued matrix with the following finite elements:

$$S_{nn}^0=\lambda^2+\mu_1^2+\frac{n}{2}+\frac{1}{4}, \quad S_{n\,n+1}^0=-\lambda\sqrt{n+1},$$

$$S_{n\,n+2}^0=\frac{1}{4}\sqrt{(n+1)(n+2)}.$$

Then the matrix of the operator g_0, $S=\|S_{nk}\|$, has the form $S=S_0(\mu_1)S_0(\mu_2)\cdots S_0(\mu_N)$, and each row (and hence column) of S contains a finite number of nonzero elements $S_{nk}=S_{kn}$. This implies that the function $g_0\psi_n=\Sigma_k S_{kn}\psi_k$ belongs to $\mathscr{S}_0$ and that $\mathscr{S}_0$ is the invariant space for g_0. Note that the operator g_0 has a purely continuous spectrum and the ψ_p are its eigenfunctions, $g_0\psi_p=N_p^2\psi_p$, where $N_p^2=f((p+\lambda)^2)>0$.

In view of the relation $g_0=L^\dagger L$, the functions $\varphi_p=N_p^{-1}L\psi_p$ are eigenfunctions of the operator $g_1=LL^\dagger$, i.e., $g_1\varphi_p=N_p^2\varphi_p$ and $\langle\varphi_p|\varphi_q\rangle=\delta(p-q)$. The spectrum of the operator g_1 in Hilbert space H is mixed. The discrete part of this spectrum coincides with the numbers $-\mu_k^2$, $k=1,\ldots,N$. We denote the eigenfunctions belonging to the discrete spectrum by φ_{-k}. Obviously, the operator g_1 is self-adjoint in H and is defined in a dense domain, which can be represented by the orthogonal sum $H=H_0\oplus H_1$, where H_0 is the N-dimensional Hilbert space with the basis φ_{-k}, $k=1,\ldots,N$. We will not discuss the properties of this space here. The space H_1 is the invariant space for g_1, and the restriction of g_1 on H_1 has a purely continuous spectrum with eigenfunctions φ_p. In what follows we assume that g_1 is only a operator acting in H_1 and hence keep its current notation. Note that the functions φ_{-k} cannot be obtained by applying the operator L to a function belonging to H.

The operator L of the form (2) can be applied to all functions from H that are included in the domain of the operator g_0, since

$$\langle L\psi_a|L\psi_b\rangle=\langle\psi_a|L^\dagger L\psi_b\rangle=\langle\psi_a|g_0|\psi_b\rangle$$

is finite for all the functions $\psi_{a,b}$ from this domain. (More precisely, this implies that for the domain of the operator L we can take the domain of the operator $\sqrt{g_0}$, which is broader than the domain of g_0.) Reasoning along the same lines, we can say that $L^\dagger$ can be applied to any function belonging to H_1 that is included in the domain of g_1.

The operators g_{00} and g_0 can be expressed in terms of the projectors on their eigenfunctions via the following spectral decompositions:

$$g_{00}\int dp\,(p+\lambda)^2|\psi_p\rangle\langle\psi_p|, \quad g_0=\int dp\,N_p^2|\psi_p\rangle\langle\psi_p|.$$

The operator g_1 can be expressed in terms of the projectors on the φ_p:

$$g_1=\int dp\,N_p^2|\varphi_p\rangle\langle\varphi_p|.$$

Then g_1 can be associated with an operator g_{11} such that $g_1=f(g_{11})$ by defining it in H_1 via the spectral decomposition

$$g_{11}=\int dp\,(p+\lambda)^2|\varphi_p\rangle\langle\varphi_p|.$$

Obviously, g_{11} is a self-adjoint operator with a purely continuous spectrum from the interval $[\lambda^2,\infty)$. At $\lambda=0$ the operator g_1 coincides with the Hamiltonian of the time-independent N-soliton potential, whose action is restricted to H_1.

The wave functions ψ_p and φ_p constitute orthonormalized bases in H_0 and H_1, respectively. Let us examine an operator that associating with each function ψ_p a function φ_p:

$$U=\int dp\,|\varphi_p\rangle\langle\psi_p|.$$

The inverse map is performed by the adjoint operator

$$U^{-1}=U^\dagger=\int dp\,|\psi_p\rangle\langle\varphi_p|.$$

Obviously, U conserves the value of a scalar product and hence is an isometric operator. Consequently, for the operators L and $L^\dagger$ we have

$$L=\int dp\,N_p|\varphi_p\rangle\langle\psi_p|,\quad L^\dagger=\int dp\,N_p|\psi_p\rangle\langle\varphi_p|.$$

This leads to a representation of L and $L^\dagger$ in terms of U, g_0, and g_1:

$$L=Ug_0^{1/2}=g_1^{1/2}U,\quad L^\dagger=g_0^{1/2}U^\dagger=U^\dagger g_1^{1/2}.$$

Such a representation of operators is known as a polar decomposition (see, e.g., Ref. 43).

We now turn to the operator $M=(L^\dagger)^{-1}=Ug_0^{-1/2}=g_1^{-1/2}U$. Since

$$g_0^{\pm 1/2}=\int dp\,N_p^{\pm 1}|\psi_p\rangle\langle\psi_p|,$$

we have

$$M=\int dp\,N_p^{-1}|\varphi_p\rangle\langle\psi_p|,\quad M^\dagger=\int dp\,N_p^{-1}|\psi_p\rangle\langle\varphi_p|.$$

The operator M coincides with the integral operator introduced in Ref. 25 (see also Ref. 27). The operators M and $M^\dagger$ factorize the operators g_0^{-1} and g_1^{-1}, i.e., $M^\dagger M=g_0^{-1}$ and $MM^\dagger=g_1^{-1}$.

In view of the isometric nature of the operator U, the functions $\zeta_n=U\psi_n$, $n=0,1,2,\ldots$, form an orthonormalized basis in H_1. There is no way in which simple explicit expressions can be derived for these functions, since U is expressed in terms of the nonlocal operator $g_0^{-1/2}$. The functions $\varphi_n(x,t)=L\psi_n(x,t)$ are much simpler, since, according to (2), applying the operator L amounts to calculating derivatives and doing simple arithmetic. Obviously, $\varphi_n=g_1^{1/2}\zeta_n$, and the operator g_1 has a zero kernel in H_1 (and so does $g_1^{1/2}$). This implies that the functions φ_p form a nonorthogonal basis in H_1, i.e., $\langle\varphi_n|\varphi_k\rangle=S_{nk}$ (what is known as a basis equivalent to an orthonormalized one, or a Riesz basis; see, e.g., Ref. 44). The functions $\eta_n=g_1^{-1/2}\zeta_n=M\psi_n$ form a base set that is biorthogonal with φ_n, i.e., $\langle\varphi_n|\eta_k\rangle=\delta_{nk}$ and $\langle\eta_n|\eta_k\rangle=S_{nk}^{-1}$, where $S_{nk}^{-1}=S_{kn}^{-1}$ are the of the matrix S^{-1}, the reciprocal of S. The matrix elements S_{nk}^{-1} can be calculated by using the spectral decomposition of the operator g_0^{-1}:

$$g_0^{-1}=\int dp\,N_p^{-2}|\psi_p\rangle\langle\psi_p|. \tag{6}$$

Then

$$S_{nk}^{-1}=\langle\psi_n|M^\dagger M\psi_k\rangle=\langle\psi_n|g_0^{-1}|\psi_k\rangle$$
$$=\int dp\,N_p^{-2}\langle\psi_n|\psi_p\rangle\langle\psi_p|\psi_k\rangle.$$

Concluding this section, we note that not every function φ belonging to H_1 can be represented in the form $\varphi=L\psi$, where ψ is a function belonging to H_0. The reason is that H_1 is the closure of the linear span of the vectors $\varphi_n=L\psi_n$ relative to the scalar product in H, which product is restricted to the space of these vectors. On the other hand, the set of the functions $\varphi=L\psi$ is a complete Hilbert space relative to the scalar product $\langle\varphi_a|\varphi_b\rangle_1=\langle L\psi_a|L\psi_b\rangle_1=\langle\psi_a|g_0|\psi_b\rangle$, which is embedded in H_1.

5. COHERENT STATES OF SOLITON POTENTIALS

According to the ideas expressed in Refs. 25, 26, and 31, to obtain coherent states of Darboux-transformed systems we need only apply the transformation operator to the coherent states of the initial system. In this case the properties 1–3 formulated in the Introduction are sure to be satisfied. Thus, to be able to interpret the states obtained in this manner as coherent states, it is sufficient to have a measure that realizes the decomposition (1) of the identity operator in these states. Note that this decomposition was not discussed in Refs. 25, 26, and 31. In the present section we will derive explicit expressions for the corresponding measures for two types of coherent states of a time-dependent many-soliton potential, states obtained via the operators L and M. The particular case of $\lambda=0$ corresponds to the ordinary time-independent many-soliton potential.

Let us consider states described by the following wave functions:

$$\varphi_z=L\psi_z=\Phi\sum_n a_n z^n\varphi_n,\quad \eta_z=M\psi_z=\Phi\sum_n a_n z^n\eta_n.$$

We wish to show that the measures $\mu_\varphi=\mu_\varphi(z,\overline{z})$ and $\mu_\eta=\mu_\eta(z,\overline{z})$, which realize the decomposition the identity operator in these states do indeed exist.

We begin with the states η_z and the measure μ_η. Since the functions $\eta_p=N_pM\psi_p$, $\langle\eta_p|\eta_q\rangle=\delta((p-q)$, $p\in\mathbb{R}$, are basis states (in the sense of generalized functions) in H_1, Eq. (1) is equivalent to

$$\int d\mu_\eta\langle\eta_p|\eta_z\rangle\langle\eta_z|\eta_p\rangle=\delta(p-q).$$

This implies

$$N_p^{-1}N_q^{-1}\int d\mu_\eta\langle\psi_p|\psi_z\rangle\langle\psi_z|\psi_q\rangle=\delta(p-q).$$

Note that the integral on the left-hand side is time-independent and hence can be calculated at $t=0$. In what follows we use this trick without further comments. Allowing for the expression for the function $\langle\psi_p|\psi_z\rangle$,

$$\langle\psi_p|\psi_z\rangle=(2\pi)^{1/4}\Phi\psi_p(z),$$

$$\psi_p(z)=\exp(-p^2+2zp-z^2/2),\quad z=x+iy,$$

we look for a measure μ_η such that $\delta\mu_\eta=\omega_\eta(x)\,dx\,dy$. After integrating with respect to y, we arrive at an equation for $\omega_\eta(x)$:

$$\int dx\,\omega_\eta(x)F_p(x)=N_p^2(2\pi)^{-1/2}\exp(2p^2),$$

$$F_p(x)=\exp(4px-2x^2).$$

Since N_p^2 is a polynomial in p, the function $\omega_\eta(x)$ is a polynomial in x, whose coefficients are determined solely by the coefficients of the polynomial N_p^2. For example, for $N_p^2=(p-\lambda)^2+\mu_1^2$ (a one-soliton potential) we have

$$\omega_\eta=[(x-\lambda)^2+\mu_1^2-1/4]/\pi.$$

We see that the states η_z satisfy all the properties of coherent states. Using the factorization properties of the operators M and $M^\dagger$ and the spectral decomposition (6) of the operator g_0^{-1}, we calculate the normalization coefficient for these states:

$$N_z^{-2}=\langle\eta_z|\eta_z\rangle=\langle\psi_z|g_0^{-1}|\psi_z\rangle=\int dp\,N_p^{-2}|\langle\psi_z|\psi_p\rangle|^2.$$

Since N_p^2 is a polynomial in $\tau=(p+\lambda)^2$, or

$$N_p^2=(\tau+\mu_1^2)(\tau+\mu_2^2)\cdots(\tau+\mu_N^2),$$

we have

$$N_p^{-2}=\sum_{k=1}^N\frac{A_k}{\tau+\mu_k^2},\quad A_k=[(dN_p^2/d\tau)_{\tau=-\mu_k^2}]^{-1}. \tag{7}$$

Hence we arrive at an expression for the normalization integral:

$$N_z^{-2}=\sum_{k=1}^N A_kF_k,\quad z=x+iy,$$

$$F_k=\frac{\sqrt{2\pi}}{\lambda}\exp[2(\lambda^2-x^2)]\mathrm{Re}[\exp[4i\lambda(x-\mu_k)]\times\mathrm{erfc}(\lambda\sqrt{2}+i\sqrt{2}\,(x-\mu_k))].$$

Now we turn to the states φ_z and the measure μ_φ. We look for a measure μ_φ such that $d\mu_\varphi=dy\,d\varphi_\varphi(x)$. Then for $\omega_\varphi(x)$ we have the equation

$$\int d\omega_\varphi(x)\,F_p(x)=N_p^{-2}(2\pi)^{-1/2}\exp(2p^2). \tag{8}$$

We wish to show that the solution of this equation is a generalized function specified on a certain linear space. First we note that $|F_p(x+iy)|\leqslant\exp(-dx^2+by^2)$ holds for $2\leqslant d\leqslant b$. This implies $F_p(x)\in S_{1/2}^{1/2}$, where $S_{1/2}^{1/2}$ is the space of entire functions F such that $|F(x+iy)|\leqslant\exp(-dx^2+by^2)$, $0\leqslant d\leqslant b$ (Ref. 45). We look for φ_φ in the form of a functional over $S_{1/2}^{1/2}$. As is known, positive definite generalized functions over $S_{1/2}^{1/2}$ are specified by their Fourier transforms. If we denote the Fourier transform of the generalized function ω_φ by $\tilde\omega_\varphi$, the left-hand side of Eq. (8) must be interpreted in the sense of the equality

$$\int d\omega_\varphi(x)\,F_p(x)=\int d\tilde\omega_\varphi(t)\,\tilde F_p(t),$$

where $\tilde F_p(t)=\sqrt{\pi/2}\exp(2p^2+ipt-t^2/8)$ is the Fourier transform of the function $F_p(x)$. As a result, Eq. (8) yields an equation for $\tilde\omega_\varphi$:

$$\pi\int d\tilde\omega_\varphi(t)\exp\left(-\frac{t^2}{8}+ipt\right)=N_p^{-2}.$$

If we now look for a $\tilde\omega_\varphi$ such that $d\tilde\omega_\varphi(t)=\rho_\varphi(t)\,dt$, then with allowance for the expression (7) for N_p^{-2} we obtain a formula for $\rho_\varphi(t)$:

$$\rho_\varphi(t)=\frac{1}{2\pi}\sum_{k=1}^N\frac{A_k}{\mu_k}\exp\left(i\lambda t-\mu_k|t|+\frac{t^2}{8}\right). \tag{9}$$

Note that for functions $\rho_\varphi(t)$ of the form (9) the integral $\int d\tilde\omega_\varphi(t)\,\tilde F(t)$ does not converge for all $F(x)\in S_{1/2}^{1/2}$. Clearly, the convergence condition for this integral imposes a restriction on the decrease of the function $F(x)$ as $|x|\to\infty$. The integral converges only if $|F(x)|\geqslant\exp(-2x^2-Ax)$, where A is a nonnegative constant, one for each function $F(x)\in S_{1/2}^{1/2}$. We denote the set of functions satisfying this condition (the set is obviously a linear space) by $S^{\circ 1/2}_{1/2}(\in S_{1/2}^{1/2})$.

Thus, we have established that the decomposition (1) of the identity operator holds for states φ_z if the measure $d\mu_\varphi=dy\,d\omega_\varphi(x)$ is in terms of the Fourier transform $\tilde\omega_\varphi$ of the measure ω_φ, which specifies a functional over the space $S^{\circ 1/2}_{1/2}$. Here an integral with respect to μ_φ must be calculated by the formula

$$\int d\mu_\varphi\,\langle\varphi_a|\varphi_z\rangle\langle\varphi_z|\varphi_b\rangle=\int dt\,\rho_\varphi(t)\tilde F_{ab}(t),$$

where $\tilde F_{ab}(t)$ is the Fourier transform of the function

$$F_{ab}(x)=\int dy\,\langle\varphi_a|\varphi_z\rangle\langle\varphi_z|\varphi_b\rangle,\quad z=x+iy.$$

The square of the norm of the function φ_z coincides with the expectation value of the operator g_0 in state ψ_z and can easily be calculated. For instance, for the one-soliton potential we have

$$\langle\varphi_z|\varphi_z\rangle=\langle\psi_z|g_0|\psi_z\rangle=1/4+\mu_1^2+(x-\lambda)^2,\quad z=x+iy.$$

6. CONCLUSION

We have established the relationship that exists between real-valued soliton potentials of the most general form, i.e., real-valued time-dependent many-soliton potentials, and the spectral problem for a self-adjoint operator in Hilbert space.

In the particular case where the potential is time-independent, this operator is a polynomial in the Hamiltonian of the soliton potential. We have derived a quasi-spectral decomposition for the Darboux transformation operators. Two systems of coherent states of soliton potentials obtained via a Darboux transformation operator and via an integral operator have been discussed. In each case we have calculated the measure that realizes the decomposition of the identity operator in coherent states.

The existence of the decomposition (1) of the identity operator in coherent states makes it possible to obtain a holomorphic representation of the state space and the operators acting in this space. Using Berezin's technique of covariant symbols,[46] one can obtain a classical mechanical system corresponding to the given quantum system (see, e.g., Ref. 2). This line of reasoning was used in Ref. 30 to generate a classical mechanical analog of a quantum system obtained, via the Darboux transformation, from a system with a potential of the type $x^2+\gamma x^{-2}$, and to formulate the Darboux transformation in classical mechanical terms. The results of the present paper provide the means for building a classical mechanical analog of a quantum system with a soliton potential and for initiating a process in this system that is the inverse of quantization.

Partial support for this work was provided by the Russian Fund for Fundamental Research (Grant No. 9702-16279).

*)E-mail: samsonov@phys.tsu.ru

[1] E. Schrödinger, Naturwissenschaften **14**, 664 (1926) [English transl.: *Collected Papers on Wave Mechanics*, Blackie & Son, London (1928), pp. 41].
[2] A. M. Perelomov, *Generalized Coherent States and their Applications*, Springer, New York (1986) [Russ. orig. Nauka, Moscow (1987)].
[3] I. A. Malkin and V. I. Man'ko, *Dynamic Symmetries and Coherent States of Quantum Systems* [in Russian], Nauka, Moscow (1979).
[4] J. R. Klauder and B.-S. Skagerstam, *Coherent States; Applications in Physics and Mathematical Physics*, World Scientific, Singapore (1985).
[5] V. V. Dodonov and V. I. Man'ko, in *Proceedings (Trudy) of the P. N. Lebedev Physics Institute* [in Russian], Vol. 183, Moscow (1987), p. 71.
[6] Ya. A. Smorodinskiĭ, A. L. Shelepin, and L. A. Shelepin, Usp. Fiz. Nauk **162**(12), 1 (1992) [Sov. Phys. Usp. **35**, 1005 (1992)].
[7] V. V. Dodonov, V. I. Man'ko, and O. V. Malkin, in *Proceedings (Trudy) of the P. N. Lebedev Physics Institute* [in Russian], Vol. 204, Moscow (1988), p. 204.
[8] J. R. Klauder, J. Phys. A **29**, L293 (1996).
[9] Ya. I. Granovskiĭ, A. S. Zhedanov, and I. M. Lutsenko, Zh. Éksp. Teor. Fiz. **99**, 369 (1991) [Sov. Phys. JETP **72**, 205 (1991)].
[10] Ya. I. Granovskii, I. M. Lutzenko, and A. S. Zhedanov, Ann. Phys. (San Diego) **217**, 1 (1992).
[11] S. Yu. Dubov, V. M. Eleonskiĭ, and N. E. Kulagin, Zh. Éksp. Teor. Fiz. **102**, 814 (1992) [Sov. Phys. JETP **75**, 446 (1992)].
[12] V. M. Eleonskiĭ and V. G. Korolev, Zh. Éksp. Teor. Fiz. **110**, 1967 (1996) [Sov. Phys. JETP **83**, 1084 (1996)].
[13] V. Spiridonov, Phys. Rev. Lett. **69**, 398 (1992).
[14] S. Skorik and V. Spiridonov, Lett. Math. Phys. **28**, 59 (1993).
[15] A. S. Zhedanov, J. Math. Phys. **34**, 2631 (1993).
[16] V. P. Spiridonov, Lett. Math. Phys. **35**, 179 (1995).
[17] G. Darboux, *Leçons sur la théorie générale des surfaces et les application géométriques du calcul infinitésimale*, Deuxiéme partie, Gauthier-Villar et Fils, Paris (1889).
[18] E. L. Ince, *Ordinary Differential Equations*, Longmans-Green, London (1927).
[19] V. G. Bagrov and B. F. Samsonov, Fiz. Élem. Chastits At. Yadra **28**, 951 (1997) [Phys. Part. Nuclei **28**, 374 (1997)].
[20] A. B. Shabat, Inverse Probl. **6**, 303 (1992).
[21] C. D. J. Fernandez, V. Hussin, and L. M. Nieto, J. Phys. A **27**, 3547 (1994).
[22] C. D. J. Fernandez, L. M. Nieto, and O. Rosas-Ortiz, J. Phys. A **28**, 2963 (1995).
[23] V. Spiridonov, Phys. Rev. A **52**, 1909 (1995).
[24] V. I. Man'ko, G. Marmo, E. C. G. Sudarshan, and F. Zaccaria, Phys. Scr. **55**, 528 (1977).
[25] V. G. Bagrov and B. F. Samsonov, Zh. Éksp. Teor. Fiz. **109**, 1105 (1996) [JETP **82**, 593 (1996)].
[26] V. G. Bagrov and B. F. Samsonov, J. Phys. A **29**, 1011 (1996).
[27] V. Matveev and M. Salle, *Darboux Transformations and Solitons*, Springer, New York (1991).
[28] V. G. Bagrov and B. F. Samsonov, Phys. Lett. A **210**, 60 (1996).
[29] P. B. Abraham and H. A. Moses, Phys. Rev. A **22**, 1333 (1980).
[30] B. F. Samsonov, J. Math. Phys. **38**, 4492 (1997).
[31] B. F. Samsonov, Yad. Fiz. **59**, 753 (1996) [Phys. At. Nucl. **59**, 720 (1996)].
[32] V. G. Bagrov, B. F. Samsonov, and L. A. Shekoyan, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 1, 84 (1985).
[33] J. Vogel, E. Vogel, and C. Toepffer, Ann. Phys. (San Diego) **164**, 463 (1985).
[34] S. V. Manakov, S. P. Novikov, L. P. Pitaevskiĭ, and V. E. Zakharov, *Theory of Solitons: The Inverse Scattering Method*, Consultants Bureau, New York (1984).
[35] B. A. Dubrovin, T. M. Malanyuk, I. M. Krichever, and V. G. Makhan'kov, Fiz. Élem. Chastits At. Yadra **19**, 579 (1988) [Sov. J. Part. Nucl. **19**, 252 (1988)].
[36] V. P. Berezovoĭ and A. I. Pashnev, Teor. Mat. Fiz. **74**, 392 (1988).
[37] C. V. Sukumar, J. Phys. A **19**, 2297 (1986).
[38] B. F. Samsonov, J. Phys. A **28**, 6989 (1995).
[39] W. Miller, *Symmetry and Separation of Variables*, Addison-Wesley, Reading, MA (1981).
[40] I. M. Gel'fand and G. E. Shilov, *Theory of Differential Equations*, Academic Press, New York (1967).
[41] F. A. Berezin and M. A. Shubin, *The Schrödinger Equation*, Kluwer Academic, Boston (1991) [Russ. orig. Moscow Univ. Press, Moscow (1983)].
[42] A. P. Veselov and A. B. Shabat, Funktsional. Anal. Prilozh. **27**, 1 (1993).
[43] M. Reed and B. Simon, *Methods of Modern Mathematical Physics*, Vol. 1: *Functional Analysis*, Academic Press, New York (1978).
[44] I. C. GoKhberg and M. G. Kreĭn, *Introduction to the Theory of Linear Nonselfadjoint Operators*, American Mathematical Society, Providence, R.I. (1969; reprinted 1978).
[45] I. M. Gel'fand and N. Ya. Vilenkin, *Applications of Harmonic Analysis*, Academic Press, New York (1964).
[46] F. A. Berezin, *Methods of Second Quantization*, Academic Press. New York (1963).

Translated by Eugene Yankovsky

The effect of fluctuations of a classical exterior medium on tunneling nonradiative processes in molecules

A. V. Belousov*) and V. A. Kovarskiĭ

Institute of Applied Physics, Moldavian Academy of Sciences, 277028 Kishinev, Moldova
(Submitted 31 October 1998)
Zh. Éksp. Teor. Fiz. **114**, 1944–1953 (December 1998)

We examine nonradiative transitions in molecules with allowance for the effect of a classical polar exterior medium on tunneling charge transport. The approach allows for the vibrational frequencies of a molecule in the electron transition. In the case of slow fluctuations, the theory predicts a low-temperature (non-Arrhenius) increase in the tunneling nonradiative transition rate, and the results agree qualitatively with the experimental data. When the fluctuations of the exterior medium are rapid, at certain values of the molecular parameters the tunneling decay rate is found to decrease with increasing temperature because the conditions needed for resonant tunneling are violated. © *1998 American Institute of Physics.* [S1063-7761(98)00312-6]

1. INTRODUCTION

Tunneling processes in molecules (tunneling charge transport, tunneling chemical reactions, the quantum yield of luminescence, etc.) have been thoroughly studied in the past decades (see, e.g., Refs. 1 and 2). However, there is one problem that still requires clarification, the problem of the temperature dependence of the rates of tunneling processes. The traditional view of tunneling as a temperature-independent process is now being reconsidered and is usually linked with the effect of the exterior medium on these processes.[3–7]. Actually, the exterior medium affects the tunneling nonradiative process and the shape of the potential barrier of the reaction (note that Korst and Osherov[8,9] were the first to study the effect of the exterior medium on the rate of activation nonradiative processes). On the other hand, a well-known effect in optical transitions in molecules is the dynamic narrowing of the absorption lineshape, when, e.g., under certain conditions allowing rapid fluctuations of the exterior medium leads not to broadening but to effective narrowing of the absorption line as the correlation time τ_c of the medium decreases.[10–12]. As we will show in this paper, this effect plays an important role in nonradiative transitions, since it strongly affects the constant-energy and resonance conditions under which tunneling processes operate and ensures the characteristic temperature dependence of such processes. We limit ourselves to the approximation of two electronic terms with close energy values.

We assume that the corresponding energy difference Δ_{21} is much larger than kT, so that the probabilities of activation processes are negligible. The presence of the electric fields of a polar exterior medium may lead to fluctuations for which the effective energy gap Δ_{21} disappears because of Stark shifts of the levels, and a tunneling nonradiative transition occurs. If the characteristic time of variation of the fluctuations of the exterior medium is much longer than the tunneling time, the nonradiative transition rate increases with the intensity of the fluctuations. If, however, the time of variation of the fluctuations of the exterior medium is very short (i.e., much shorter than the tunneling time), the energy gap Δ_{21} remains essentially the same because of the balance of the electric fields generated by the dipoles of the polar exterior medium at the point occupied by the impurity molecules in time intervals proportional to the tunneling time. It is this ''switching off'' of the fluctuations of the exterior medium that destroys the constant-energy conditions necessary for the tunneling process to proceed and ensures a decrease in the nonradiative transition rate, despite the increasing intensity of the fluctuations.

These two cases are studied in the present article as limits of a general theory that is being developed. In contrast to Ref. 6, which considers the basic model (i.e., the model of displaced parabolas with equal frequencies in the electronic states) of the impurity molecule, here we use a more realistic model of the impurity molecule, a model that allows for variations of the vibrational frequencies in the electron transition. As noted in Refs. 13–15, this may change the nonradiative transition rate substantially. We assume that the polar exterior medium is characterized by a Gaussian Markov autocorrelation function,

$$\varphi(t_1,t_2)=B_0^2\exp(-\gamma|t_1-t_2|), \tag{1}$$

where B_0^2 is the (temperature-dependent) noise intensity, and $\gamma=1/\tau_c$. In many respects the parameter of the theory, b/κ (see below), which determines the temperature dependence of the nonradiative transition rate, is specified by the choice of the model describing the exterior medium. For the low-frequency classical exterior medium, $B_0^2\sim T$. Below we establish that for slow fluctuations of the exterior medium, the rate of the tunneling process is proportional to $\exp(cT)$, where c is a constant, and its increase as a function of temperature is not governed by the Arrhenius law. For rapid fluctuations ($\kappa\gg 1$), the specially selected values of the parameters ($\Delta_{21}\gtrsim\hbar\omega$), and a wide barrier, the transition rate decreases with increasing temperature as $1/\sqrt{T}$.

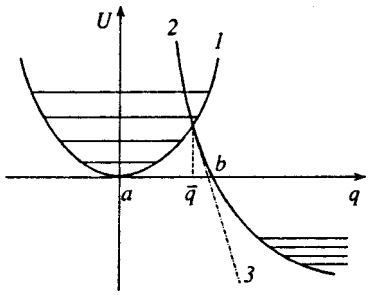

FIG. 1. Curves *1* and *2* correspond to different electronic terms of the molecule, the straight line *3* corresponds to the repulsive term of the molecule, $\bar{q}$ is the point of intersection, and ab is the tunneling path.

2. THE GENERAL EXPRESSION FOR THE RATE OF A TUNNELING NONRADIATIVE TRANSITION IN A MOLECULE WITH ALLOWANCE FOR THE EFFECT OF THE EXTERIOR MEDIUM

Let us consider a two-term molecule whose adiabatic potentials take into account both the positions of the nuclei and their vibrational frequencies in the electronic states 1 and 2. We are interested in the tunneling transition of the system from state 1 into state 2 along path ab (Fig. 1). Let us assume that the vibrational frequencies in state 1 (ω_1) and in state 2 (ω_2) obey the condition $\omega_1 \gg \omega_2$ ($\hbar\omega_1 \gg kT$). This condition leads to a situation in which the low-frequency polar exterior medium interacts more strongly with the low-frequency (ω_2) molecular vibrations in the final state 2, and to simplify matters we ignore the effect of the vibrations of the exterior medium on the vibrational mode with frequency ω_1 in the initial electronic state 1. The vibrational Hamiltonians in the electronic states 1 and 2 are

$$H_{1,2} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial q^2} + U_{1,2}(q),$$

$$U_1(q) = \frac{m\omega_1^2}{2}q^2,$$

$$U_2(q) = \frac{m\omega_2^2}{2}q^2 - v(\bar{q})(q-\bar{q}) - f(t)q, \tag{2}$$

where $\bar{q}$ is the point at which the vibrational terms cross (see Fig. 1). The force $f(t)$ is a Gaussian Markov process with the correlation function (1). The rate of the tunneling nonradiative process can be written in the following form (see the Appendix):

$$W_{21} = \frac{V_{21}^2}{\hbar^2}\,2\mathrm{Re}\int_0^\infty d\tau\, I_{21}(\tau)\exp\frac{iE_0\tau}{\hbar}, \tag{3}$$

$$I_{21}(\tau) = \int dq \int dq'\ \phi_0(q)\langle K(q\tau|q')\rangle\phi_0(q'). \tag{4}$$

Here V_{12} is the matrix element of the $1\to 2$ transition, $\varepsilon_0 = (1/2)m\omega_1^2\bar{q}^2$, $E_0 = \varepsilon_0 + v(\bar{q})\sqrt{\varepsilon_0}\sqrt{2/m\omega_1^2}$, $\phi_0(q)$ is the wave function of the ground state of an oscillator with frequency ω_1, and $K(q\tau'|q')$ is the Green's function determined by the Hamiltonian H_2. The angle brackets indicate averaging over the realization of the random process $f(t)$. We write the Green's function $K(q\tau|q')$ as a path integral:

$$K(q\tau|q') = \int Dq(\tau)\exp\left\{\frac{i}{\hbar}S(q\tau|q')\right\}, \tag{5}$$

where

$$S(q\tau|q') = \int_0^\tau dt\left[\frac{m}{2}\dot{q}^2 - \frac{m\omega_2^2}{2}q^2 + (v+f(t))q\right]$$

is the classical action. The path integral (5) is calculated along the paths $q(t)$ that satisfy the boundary conditions $q(0)=q'$ and $q(\tau)=q$. Calculating the average over the realizations of the random process $f(t)$ yields

$$\left\langle \exp\left\{\frac{i}{\hbar}\int_0^\tau dt\, f(t)q(t)\right\}\right\rangle = \exp\left\{\frac{1}{2}\left(\frac{i}{\hbar}\right)^2\int_0^\tau dt_1\int_0^\tau dt_2\, q(t_1)\varphi(t_1,t_2)q(t_2)\right\}. \tag{6}$$

The averaged Green's function $\langle K(q\tau|q')\rangle$ can be written

$$\langle K(q\tau|q')\rangle = \int Dq(\tau)\exp\left\{\frac{i}{\hbar}S_{\mathrm{eff}}(q\tau|q')\right\}. \tag{7}$$

The effective action $S_{\mathrm{eff}}(q\tau|q')$ is given by the formula

$$S_{\mathrm{eff}} = \int_0^\tau dt\left[\frac{m}{2}\dot{q}^2 - \frac{m}{2}\omega_2^2q^2 + vq + \frac{iB_0^2}{2\hbar}\times\int_0^\tau ds\exp(-\gamma|t-s|)\,q(t)q(s)\right]. \tag{8}$$

The extremal path $q(t)$, which minimizes the action $S_{\mathrm{eff}}(q\tau|q')$, satisfies the equation

$$\ddot{q} + \omega_2^2 q = \frac{iB_0^2}{\hbar m}\int_0^\tau ds\exp(-\gamma|t-s|)\,q(s) + \frac{v}{m}. \tag{9}$$

The extremal action on the extremal path (9) assumes the form

$$S_{\mathrm{eff}}^{(\mathrm{cl})}(q\tau|q') = \frac{m}{2}q\dot{q}\big|_0^\tau + \frac{1}{2}v\int_0^\tau dt\, q(t). \tag{10}$$

To exponential accuracy, the path integral (7) can be written[16]

$$\langle K(q\tau|q')\rangle = \left(-\frac{1}{2\pi\hbar i}\frac{\partial^2 S_{\mathrm{eff}}^{(\mathrm{cl})}}{\partial q\,\partial q'}\right)^{1/2}\exp\left[\frac{i}{\hbar}S_{\mathrm{eff}}^{(\mathrm{cl})}(q\tau|q')\right]. \tag{11}$$

Note that the pre-exponential factor is independent of q and q', since the extremal path $q(t)$ is a linear form in q and q'. Hence the expression (11) for the averaged Green's function $\langle K(q\tau|q')\rangle$ is exact. The integro-differential equation (9) can be reduced to a fourth-order differential equation:

$$Q^{(4)} + (\omega_2^3 - \gamma^2)Q^{(2)} - (2i\gamma D + \gamma^2\omega_2^2)Q = 0, \tag{12}$$

$$Q = q - A, \quad A = \frac{F_0\gamma^2}{2i\gamma D + \gamma^2\omega_2^2}, \quad D = \frac{B_0^2}{\hbar m}, \quad F_0 = \frac{v}{m}.$$

The two addition integration constants can be found from Eq. (9) at the points $t=0$ and $t=\tau$. By plugging (11) into the expression for the generating function (4) and calculating the elementary Gaussian integrals with respect to q and q' we find $I_{21}(\tau)$. Then we determine the rate of the nonradiative process by formula (3). The solution of the boundary value problem (9), (12) is extremely cumbersome. For this reason we examine the limits of slow and rapid fluctuations of the exterior medium.

(a) The case of slow fluctuations of the exterior medium

We replace the parabolic term for the coordinate of the reaction near the intersection point $\overline{q}$ by the linear term $v(q-\overline{q})$ (see Fig. 1). In the limit of slow fluctuations of the random force $f(t)$ (the quasistatic case), $b/\kappa \gg 1$ ($b=B_0^2/\hbar m\omega_1^3$ and $\kappa=\gamma/\omega_1$), the solution for the extremal path (9) has the form

$$q(t)=\frac{q-q'}{\tau}t+q'+\frac{1}{2}\frac{F_0+(1/2)D(q+q')\tau}{1+D\tau^3/12}(t^2-t\tau). \quad (13)$$

Combining this with (10) and (11), we arrive at a formula for the generating function:

$$I_{21}(x)=\frac{1}{\sqrt{1+ix/2}}\frac{\sqrt{1-ibx}}{\sqrt{1+(bx^2/2)(1+ix/6)}} \times\exp\left[-\frac{V_0^2}{4}x^2\frac{1+ix/6}{1+(bx^2/2)(1+ix/6)}\right], \quad (14)$$

$$x=\omega_1\tau, \quad V_0^2=\frac{v^2}{\hbar m\omega_1^3}.$$

Formula (14) was obtained in the limit of slow fluctuations of the exterior medium, $k\to 0$.

(b) The case of rapid fluctuations of the exterior medium

In the limit of rapid fluctuations of the exterior medium ($\kappa\gg 1$, $b/\kappa\ll 1$), the extremal path has the form

$$q(t)=\frac{1}{\sin\Omega\tau}[q\sin\Omega t+q'\sin(\Omega(\tau-t))] + \frac{v}{m\Omega^2}\left[1-\frac{\sin(\Omega(\tau-t))+\sin\Omega t}{\sin\Omega\tau}\right], \quad (15)$$

with $\Omega\equiv\omega_2^2(1-id)$ and $d\equiv D\tau_c/\omega_2^2$. Combining (15) with (10) and (11), we arrive at a formula for the generating function in the case of rapid fluctuations of the exterior medium:

$$I_{21}(x)=\left[\cos\lambda x+\frac{i}{2}\left(\frac{\omega_1}{\omega_2}\frac{1}{\lambda}+\frac{\omega_2}{\omega_1}\lambda\right)\sin\lambda x\right]^{-1/2} \times\exp\left[\frac{iV_0^2}{2}\frac{x}{\lambda^2}-\frac{iV_0^2}{\lambda^3} \times\frac{\sin(\lambda x/2)}{\cos(\lambda x/2)+i\lambda(\omega_2/\omega_1)\sin(\lambda x/2)}\right], \quad (16)$$

where $x\equiv\omega_2\tau$ and $\lambda\equiv\Omega/\omega_2$.

3. PARTICULAR CASES. DISCUSSION

The foregoing expressions for the rates of tunneling processes in the cases of slow and rapid fluctuations can be calculated numerically by Eqs. (3) and (4). Here we limit ourselves to a qualitative analysis of some of the corollaries of these formulas.

(a) The case of slow fluctuations

To exponential accuracy, the generating function $I_{21}(x)$ given by Eq. (14) can be represented in the form

$$I_{21}(x)=\frac{1}{\sqrt{\pi}\,2B_0}\int_0^\infty df\,\exp\left(-\frac{f^2}{2B_0^2}\right)I_{21}^{(0)}(x;f), \quad (17)$$

where we have introduced the notation

$$I_{21}^{(0)}(x)=\frac{1}{\sqrt{1+ix/2}}\exp\left[-\frac{(V_0+f)^2}{4}x^2\left(1+\frac{ix}{6}\right)\right]. \quad (18)$$

The expression (18) is the generating function for the process of a nonradiative transition to the repulsive term at a fixed value of the external field acting on the molecule. In accordance with (18) we have

$$W_{21}=\frac{1}{\sqrt{\pi}\,2B_0}\int_0^\infty df\exp\left(-\frac{f^2}{2B_0^2}\right)W_{21}^{(0)}(f). \quad (19)$$

Thus, the nonradiative transition rate W_{21} is the result of averaging the expression $W_{21}^{(0)}(f)$ for the transition rate in Ref. 17 over the Gaussian distribution of the random force f. If we use the approximation of a wide potential barrier, the tunneling transition rate (19) can be written

$$W_{21}\approx W_{21}^{(\mathrm{tun})}\exp\left(\frac{2B_0^2}{V_0^2}\xi^2\right). \quad (20)$$

The rate of a tunneling nonradiative transition in the absence of an exterior medium, $W_{21}^{(\mathrm{tun})}$ has the form[17]

$$W_{21}^{(\mathrm{tun})}\approx\omega\frac{\sqrt{2\pi}}{s_0}\exp(-2s_0)\exp\left\{-\frac{4}{3}\frac{\varepsilon_0^{3/2}}{v}\right\}, \quad s_0=\frac{E_0}{\hbar\omega_1}, \quad (21)$$

where ω is the frequency of the transition rate, which is equal in order of magnitude to the vibrational frequency in the ground electronic state, and the constant ξ in the approximation in which $V_0^2\ll 1$ is given by the expression

$$\xi=\frac{1}{3}\frac{\varepsilon_0^{3/2}}{V_0(\hbar\omega_1)^{3/2}}-\frac{\varepsilon_0}{\hbar\omega_1}(\sqrt{2}-1). \quad (22)$$

One example of temperature-dependent tunneling is known from low-temperature experiments in oxidation of low-potential cytochrome.[18]. Without going into a discussion about the numerous attempts to explain the low-temperature dependence of this process by the common methods of the theory of multiphoton processes (see the review in Ref. 19), we note the existence of an alternative approach, which is based on allowing for the effect of a polar exterior medium

(including a protein medium) on the rate of nonradiative transition of cytochrome from reduced to oxidized form due to the growth of classical fluctuations with temperature in accordance with Eq. (20). Instead of the Arrhenius law $W_{21}\propto\exp\{-E/kT\}$ (E is the activation energy of the process), at low temperatures we have, according to (20), $W_{21}\approx\exp(\text{const}\times T)$, since $B_0^2\sim T$. For values of the cytochrome parameters $\hbar\omega_1\approx 0.05$ eV, $V_0\sim 0.5$, $\varepsilon_0/\hbar\omega_1\sim 3$, $\bar{\omega}\sim 10^{14}\,\text{s}^{-1}$, and $\xi^2\approx 4$, we have

$$W_{21}\approx 3\times 10^2\exp\frac{T}{43\,\text{K}}. \tag{23}$$

This formula provides a good approximation to the data in the temperature interval in which the experiment was conducted. Note that actually the linear-term model adopted in our calculations is closer to the experimental situation than the displaced-parabola model because of anharmonicity effects inherent in the highly excited states of oxidized cytochrome. Here we limit ourselves to the above example, since almost all of its parameters are known. The theory developed here can be applied to other processes of the same type via direct numerical calculations by Eqs. (3) and (14).

(b) The case of rapid fluctuations

This case ($\kappa\gg 1$) was studied by Averbukh *et al.*[6] using the basic model ($\omega_1=\omega_2=\omega$). Below we give their results (see Ref. 6) for the case of small constants $V_0^2\ll 1$, which was not analyzed in Ref. 6. If we remain within the static model that describes the behavior of the transition rate, then for the transition operator we can take the electron–vibrational interaction (with frequency ω), which mixes the electronic states 1 and 2. Let us assume that $\Delta_{21}\gtrsim\hbar\omega$. This leads to an expression for the tunneling transition rate ($\hbar\omega\gg kT$):

$$W_{21}=W_{21}^{(0)}\frac{\hbar\Gamma+(b/\kappa)\hbar\omega}{(\Delta_{21}-\hbar\omega)^2+(\hbar\Gamma+(b/\kappa)\hbar\omega)^2}, \tag{24}$$

where $W_{21}^{(0)}$ is a constant factor and Γ is the width of the electron level in the absence of an exterior medium. For intense fluctuations Eq. (24) implies that $(b/\kappa)\omega\gg\Gamma$. The width may become larger than the detuning from resonance, $\Delta_{21}/\hbar-\omega$ and can determine the tunneling nonradiative transition rate. The parameter b/κ varies with temperature ($b\propto T$), and the tunneling decay rate decreases with increasing temperature, i.e., the transition rate falls. In the more general model developed in the present paper the result is similar. For the case of large constants $V_0^2\gg 1$, $\Delta_{21}\gtrsim\hbar\omega$, the general expression (3) for the tunneling nonradiative transition rate becomes

$$W_{21}=W_{21}^{(0)}J,$$

$$J=2\,\text{Re}\int_0^\infty dt\,\frac{\exp\{i(\Delta_{21}-\hbar\omega)t/\hbar-(\Gamma+\sqrt{b/\kappa}\,\omega_1)t\}}{\sqrt{1+i\omega_1 t/2}}, \tag{25}$$

where $W_{21}^{(0)}$ is a temperature-independent constant, and the integral J reduces to

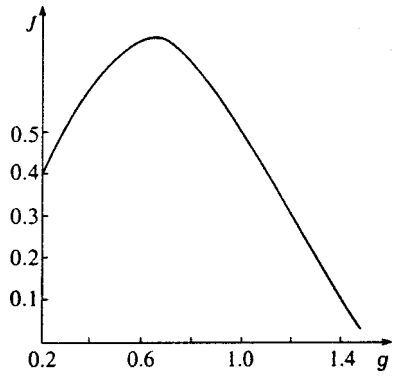

FIG. 2. J vs. parameter $g=\sqrt{b/\kappa}\,(1/\Delta)$ for $\Delta\sim 1$.

$$J=\sqrt{\frac{b}{\kappa}}\frac{1}{\Delta^{3/2}}\int_{-\infty}^{\infty}dx\,\frac{\exp(-x^2\Delta)}{(x^2/2-1)^2+(b/\kappa)(1/\Delta^2)}. \tag{26}$$

The temperature dependence of the transition rate is determined by (26). Figure 2 depicts the dependence of the tunneling transition rate on the parameter b/κ at a fixed value of $\Delta_{21}-\hbar\omega$. We see that in the model with $b\propto\infty$ the nonradiative transition rate decreases with increasing temperature, as it does in the main model considered earlier.

Our analysis of the temperature dependence of the rates of a tunneling transition in an electron-vibrational system for $\hbar\omega\gg kT$ shows that a polar exterior medium with frequencies $\nu(\hbar\nu\gtrsim kT)$ strongly affects the nonradiative transition rates. In the case of slow fluctuations of the exterior medium we arrive at a non-Arrhenius law governing the increase in the transition rate with temperature, while in the case of rapid fluctuations we have, for certain values of the parameters of the system, a decrease in the transition rate with increasing temperature.

APPENDIX

We write the molecular wave function as follows:

$$|\psi(t)\rangle=S(t,t_0)|\psi(t_0)\rangle. \tag{A1}$$

Next we write the molecular state at time $t_0\to-\infty$ in the form

$$|\psi(t_0)\rangle=|1\rangle|0\rangle,$$

where $|1\rangle$ is the wave function of the electronic state 1, and $|0\rangle$ is the wave function of the vibrational state of the electronic term 1. The scattering matrix $S(t,t_0)$ is determined by the Hamiltonian

$$H=H_0+V,$$

$$H_0=\begin{pmatrix}H_1 & 0\\ 0 & H_2\end{pmatrix},\quad V=\begin{pmatrix}0 & V_{12}\\ V_{21} & 0\end{pmatrix}, \tag{A2}$$

where $H_{1,2}$ are the vibrational Hamiltonians in the electronic states 1 and 2, V is the operator responsible for the transition (e.g., V can be related to the promoting modes of the molecules causing the transition). The final state is $|\psi_f\rangle=|2\rangle|\varphi_p\rangle$, where $|2\rangle$ is the wave function of the electronic state 2, and φ_p is the vibrational wave function of the repul-

sive term 2. The amplitude of the $1\rightarrow 2$ transition in the first order in the perturbation operator V can be written

$$A_{12}=-\frac{i}{\hbar}V_{12}\int_{t_0}^{t}d\tau\int dq\;\phi_0(q)\psi_p(q,\tau), \tag{A3}$$

where

$$\phi_0(q)=\langle q|0\rangle,\quad \psi_p(q,\tau)=\langle q|U_2(\tau,t_0)|\varphi_p\rangle,$$

with $U_2(\tau,t)$ the evolution operator determined by the vibrational Hamiltonian H_2. The differential transition rate is expressed by the formula

$$dw_{12}(p)=\frac{V_{12}^2}{\hbar^2}2\,\mathrm{Re}\int_{-\infty}^{t}d\tau\int dq_1\,dq_2\exp\left(-\frac{iE_0(\tau-t)}{\hbar}\right)$$
$$\times\phi_0(q_1)\psi_p^*(q_1t)\psi_p(q_2\tau)\phi_0(q_2). \tag{A4}$$

Integrating (A4) with respect to ϕ, we arrive at an expression for the differential transition probability:

$$W_{21}=\frac{V_{12}^2}{\hbar^2}2\,\mathrm{Re}\int_0^{\infty}d\tau\,I_{21}(\tau)\exp\frac{iE_0\tau}{\hbar},$$

$$I_{21}(\tau)=\int dq_1\,dq_2\;\phi_0(q_1)K(q_1t|q_2t-\tau)\phi_0(q_2). \tag{A5}$$

In deriving (A5) we used the representation

$$\int dp\;\psi_p^*(q_1t)\psi_p(q_2\tau)=K(q_1t|q_2\tau)$$

for the Green's function and changed the variable $t-\tau$ to τ. Averaging (A5) over the realization of the random process $f(t)$ and using the fact that the Gaussian Markov process is time-independent, we arrive at formula (3).

*)E-mail: exciton@cc.acad.md

[1] H. Eyring, S. H. Lin, and S. M. Lin, *Basic Chemical Kinetics*, Wiley, New York (1980).
[2] V. I. Gol'danskiĭ, L. I. Trakhtenberg, and V. N. Flerov, *Tunneling Phenomena in Chemical Kinetics* [in Russian], Nauka, Moscow (1986).
[3] I. Rips and J. Jortner, J. Chem. Phys. **87**, 2090 (1987).
[4] L. D. Zusman, Teor. Eksp. Khim. **15**, 227 (1979).
[5] A. I. Burstein and A. G. Kofman, Chem. Phys. **40**, 289 (1979).
[6] I. Sh. Averbukh, V. A. Kovarskiĭ, A. A. Mosyak, and N. F. Perel'man, Teor. Mat. Fiz. **81**, 271 (1989).
[7] A. O. Caldeira and A. J. Leggett, Phys. Rev. Lett. **46**, 211 (1981).
[8] N. N. Korst and V. I. Osherov, Zh. Éksp. Teor. Fiz. **51**, 825 (1966) [Sov. Phys. JETP **24**, 550 (1967)].
[9] V. I. Osherov, J. Chem. Phys. **47**, 3885 (1967).
[10] R. H. Dicke, Phys. Rev. **89**, 472 (1953).
[11] S. A. Akhmanov, Yu. E. D'yakov, and A. S. Chirkin, *Introduction to Statistical Radiophysics and Optics* [in Russian], Nauka, Moscow (1981).
[12] A. I. German, V. A. Kovarskiĭ, and N. F. Perel'man, Zh. Éksp. Teor. Fiz. **106**, 801 (1994) [JETP **79**, 439 (1994)].
[13] M. D. Frank-Kamenetskiĭ and A. V. Lukashin, Usp. Fiz. Nauk **116**, 193 (1975) [Sov. Phys. Usp. **18**, 391 (1975)].
[14] E. S. Medvedev and V. I. Osherov, *Radiationless Transitions in Polyatomic Molecules*, Springer, New York (1995).
[15] R. Engelman, *Non-Radiative Decay of Ions and Molecules in Solids*, North-Holland, Amsterdam (1979).
[16] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals*, McGraw-Hill, New York (1965).
[17] V. Kovarskii, L. Chernysh, and A. Belousov, Phys. Status Solidi B **123**, 345 (1984).
[18] D. De Vault, J. H. Parkes, and B. Chance, Nature (London) **215**, 642 (1967).
[19] É. G. Petrov, *Physics of Charge Transfer in Biosystems* [in Russian], Naukova Dumka, Kiev (1984).

Translated by Eugene Yankovsky

Mixing of bound states with electron transport by a radiation field in waveguides

E. N. Bulgakov

Kirensky Institute of Physics, Siberian Branch of Russian Academie of Sciences, 660036 Krasnoyarsk, Russia

A. F. Sadreev*)

Kirensky Institute of Physics, Siberian Branch of Russian Academie of Sciences, 660036 Krasnoyarsk, Russia; Abo Akademi, Institutionen för Fysik, Department of Physics SF-20500, Abo, Finland
(Submitted 3 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1954–1970 (December 1998)

Electron transmission in the two-, three-, and four-terminal nanostructures is considered under the influence of a radiation field. The frequency of the radiation field is tuned to the transition between the energy of a bound state and the Fermi energy of the incident electrons. The radiation induced resonant peaks and dips of the electron transport are exhibited for zero and low magnetic fields. It is shown that rotation of the radiation field polarization can effectively control the electron transport into different electrodes attached to the structures because of the symmetry of the structures. The resonant anomalies of the Hall resistance are found in a weak magnetic field. © *1998 American Institute of Physics.* [S1063-7761(98)00412-0]

1. INTRODUCTION

For several decades the transport of electrons in structures of low dimensionality and complicated geometry has been the focus of extensive theoretical and experimental study. Electrons can be confined to very narrow regions fabricated on an interface of an AlGaAs/GaAs heterostructure. Since the electrons in such regions can have high mobilities in the two dimensions available to them, such systems are called two-dimensional electron gases (2DEGs). The study of electronic transport properties of 2DEGs is of great current interest not only from the standpoint of the basic quantum effects involved but also for potential engineering applications. An idealized sample becomes an electron waveguide, wherein the quantum transport properties are solely determined by the geometry of the structure and the wavelike nature of the electrons. A remarkable manifestation of the successful achievement of quantum ballistic transport through a semiconductor nanostructure is the appearance of quantized steps on the conductance through a narrow structure as the number of one-dimensional channels is successively varied,[1,2] the quenching of the Hall effect, and the last plateau and the negative bend resistance in the cross geometry.[3–5]

Ford *et al.*[5] presented a systematic investigation of the influence of cross geometry on the Hall effect. They fabricated various differently shaped cross sections based on GaAs–$Al_xGa_{1-x}As$, which demonstrated that near zero magnetic field the Hall resistance can be quenched, enhanced over its classical value, or even negative. This effect has been considered in detail theoretically by Schult *et al.*[6] and Amemiya and Kawamura.[7] Another interesting feature of the cross geometries is a bound state found numerically by Schult *et al.*[8] and by Peeters.[9]

The question of the existence of electromagnetic modes trapped by special geometries has been a classic one in the theory of waveguides.[10] It has been realized that the introduction of bends into waveguides generally leads to bound states, or localized modes, which exist below the cut-off frequency for the waveguide. Carini *et al.*[11,12] have demonstrated theoretically and experimentally the presence of bound *TE* modes for rectangular bent waveguides and shown that the number of bound *TE* modes is bend-dependent. For the two-dimensional Schrödinger equation it was proved that any curved two-dimensional waveguide of constant width and infinite length posses bound states.[13–18] Bound states were found in the same year (1989) in a four-terminal junction of narrow wires by Schult[6,8] and independently by Peeters[9] (see also Refs. 19 and 20).

For the stationary processes of the energy conservated electron transmission only quasi-bound states with energies within the conduction subbands are important.[6,7] In particular, it was shown that the quasi-bound states in the Hall junction result in resonant dips of the resistance in high magnetic fields when the magnetic length is comparable with the size of the junction. The Hall resistance sensitively depends on the geometry of a junction and can become negative for a smoothed junction for small magnetic fields.

Although the bound states below the lowest subband threshold do not participate in stationary transmission, the possibility of observing of them, at least in principle, was shown by Berggren and Ji for the case of two intersecting electron waveguides with finite electrodes.[20] In that case bound states can be probed by resonant tunneling through the electrodes below the subband. However, it is possible to mix the bound state with electron transmission through electron waveguides with infinite electrodes directly by application of a radiation field, provided that the dipole matrix elements between the bound state and propagating ones are not equal to zero. Such a possibility was demonstrated for the four-terminal's Hall junction.[21] Let E_1 be the energy of the bound

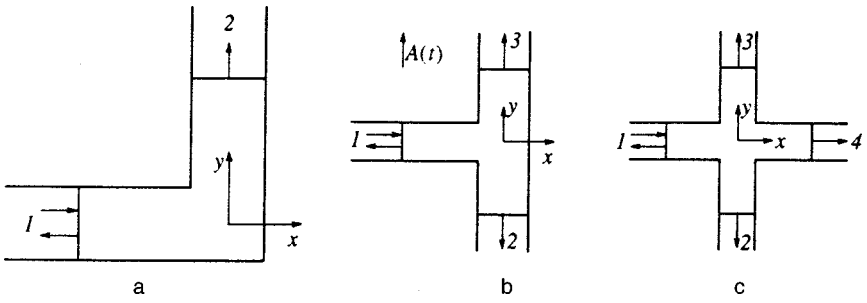

FIG. 1. The types of structures considered: *L*-structure (a), the *T*-structure (b), and *X*-structure (c).

state below the subbands which for zero magnetic field can be specified as

$$E_n(k)=\frac{\hbar^2}{2m^*d^2}(k^2+\pi^2n^2),$$

where d is the width of the electrodes, n is the number of the subband, and k is the wave number of the incident electron. Tuning a perturbation frequency near the resonance $\hbar\omega=E_n(k)-E_0$, one can expect resonant anomalies in the electron transmission through the many-terminal junctions.[21]

The aim of the present article is to consider the electron transmission effected by mixing of bound states with the propagating solutions in the *X*-, *T*-, and *L*-types of electron waveguides which are shown in Fig. 1.

2. CONDUCTANCE ANOMALIES INDUCED BY THE RADIATION FIELD

2.1. Numerical method

In this section we consider single electron transmission through the rectangular structures, the geometries of which are shown in Fig. 1 and specified below as *L*-, *T*-, and *X*-structures. They share the property of having at least one bound state. The Schrödinger equation for an electron of a mass m^* subjected to a magnetic field B applied normal to the junction and to a radiation field $\mathbf{A}_1(t)$ directed in the plane of the junction can be written

$$i\hbar\frac{\partial\psi(\mathbf{r},t)}{\partial t}=\frac{\hbar^2}{2m^*}\left(i\nabla+\frac{e}{\hbar c}(\mathbf{A}_0(\mathbf{r})+\mathbf{A}_1\cos\omega t)\right)^2\psi(\mathbf{r},t). \tag{1}$$

Here we use the gauge $\mathbf{A}_0(\mathbf{r})=(-By,0,0)$. The radiation field is considered in the long-wavelength approximation, in which the wavelength of the radiation field is much greater than the size of the junction. We use the following dimensionless transformations:

$$t\to\frac{\hbar t}{2m^*d^2},\quad \mathbf{r}\to\frac{\mathbf{r}}{d},\quad \epsilon=\frac{2m^*d^2E}{\hbar^2},$$

$$\omega\to\frac{2m^*d^2\omega}{\hbar},\quad \mathbf{a}=\frac{2\pi d\mathbf{A}_1}{\phi_0},\quad \gamma=\frac{2\pi d^2B}{\phi_0}, \tag{2}$$

where $\phi_0=ch/e$ is the magnetic flux quantum. In terms of the dimensionless variables (2) the Schrödinger equation (1) takes the form

$$i\frac{\partial\psi(\mathbf{r},t)}{\partial t}=(i\nabla+(\mathbf{a}_0(\mathbf{r})+\mathbf{a}\cos\omega t))^2\psi(\mathbf{r},t), \tag{3}$$

where $\mathbf{a}_0(\mathbf{r})=(-\gamma y,0,0)$. We map this equation onto a square lattice with elementary unit w. The lattice site is specified as (m,l). The total vector potential $\mathbf{a}_0+\mathbf{a}\cos\omega t$ is accounted for by a Peierls phase factor.[22] Then Eq. (3) transforms as follows:

$$\begin{aligned}iw^2\frac{\partial\psi(m,l)}{\partial t}=&4\psi(m,l)-\exp(i\tilde\gamma l)\psi(m+1,l)\\&-\exp(-i\tilde\gamma l)\psi(m-1,l)\\&-\exp(-i\tilde a\cos\omega t)\psi(m,l+1)\\&-\exp(i\tilde a\cos\omega t)\psi(m,l-1),\end{aligned} \tag{4}$$

where $\tilde\gamma=\gamma w^2$, $\tilde a=aw$. In the four-terminal junction we use also a different gauge $\mathbf{a}_0(\mathbf{r})=(0,\gamma x,0)$ for which the Schrödinger equation (3) will map onto a square lattice as follows:

$$\begin{aligned}iw^2\frac{\partial\tilde\psi(m,l)}{\partial t}=&4\tilde\psi(m,l)-\tilde\psi(m+1,l)-\tilde\psi(m-1,l)\\&-\exp(-i\tilde a\cos\omega t-i\tilde\gamma m)\tilde\psi(m,l+1)\\&-\exp(i\tilde a\cos\omega t+i\tilde\gamma m)\tilde\psi(m,l-1),\end{aligned} \tag{5}$$

where $\psi(m,l)=\exp(-i\tilde\gamma ml)\tilde\psi(m,l)$. Because of the processes of absorption and emission of photons, we write the wave function in the electrodes[23,24]

$$\psi(\mathbf{r},t)=\sum_n\exp[-i(\epsilon+n\omega)t]\psi_n(\mathbf{r}). \tag{6}$$

Substitution of (6) into (4) gives

$$\begin{aligned}w^2(\epsilon+n\omega)\psi_n(m,l)=&4\psi_n(m,l)-\exp(i\tilde\gamma l)\psi_n(m+1,l)\\&-\exp(-i\tilde\gamma l)\psi_n(m-1,l)\\&-\sum_s\Gamma_{ns}\psi_s(m,l-1)\\&-\sum_s\Gamma^*_{ns}\psi_s(m,l+1),\end{aligned} \tag{7}$$

where

$$\Gamma_{ns}=i^{s-n}J_{s-n}(\tilde a).$$

Here we used the standard expansion of an exponential in Bessel functions[25]

$$\exp(i\tilde{a}\cos\omega t)=\sum_m i^m J_m(\tilde{a})\exp(im\omega t). \tag{8}$$

Let us introduce column vectors for each site of the square lattice describing the amplitudes of the wave function (6)

$$\Psi(m,l)=\mathrm{col}(\ldots,\psi_1(m,l),\psi_0(m,l),\psi_{-1},(m,l),\ldots).$$

Then Eq. (7) takes more compact form

$$\begin{aligned}(w^2\epsilon+\Omega)\Psi=&4\Psi(m,l)-\exp(i\tilde{\gamma}l)\Psi(m+1,l)\\&-\exp(-i\tilde{\gamma}l)\Psi(m-1,l)\\&-\Gamma\Psi(m,l-1)-\Gamma^*\Psi(m,l+1),\end{aligned} \tag{9}$$

where we have introduced two matrices $\Omega=\mathrm{diag}(w^2n\omega)$ and $\Gamma=\{\Gamma_{ns}\}$. Following Ando[26] we take the electrodes *1* and *4* (Fig. 1c) to be infinitely long in the x direction and consisting of M lattice sites in the y direction. We introduce a generalized vector

$$C_m=\mathrm{col}(\Psi(m,M),\Psi(m,M-1),\cdots,\Psi(m,1)).$$

The dimension of this vector is $M\times L$ where L is the dimension of the vector $\Psi(m,l)$. In computer simulations the dimension L, which is the number of amplitudes of the wave function (6), was taken to be a finite number chosen by numerical accuracy.[27] We also introduce a diagonal matrix

$$P_{ll'}=\delta_{ll'}\exp(-i\tilde{\gamma}l),$$

the unit matrices I_0 of dimension $M\times M$ and I of dimension $L\times L$, and the up-diagonal matrix

$$D=\begin{pmatrix}0&1&0&0&\cdots\\0&0&1&0&\cdots\\0&0&0&1&\cdots\\\vdots&&&&\end{pmatrix}$$

of dimension $M\times M$.

Then Eq. (9) takes the form presented by Ando[26]

$$(w^2\epsilon-H_0)C_m+P_\parallel C_{m-1}+P_\parallel^* C_{m+1}=0, \tag{10}$$

where

$$H_0=4I_0\otimes I-D\otimes\Gamma^*-D^+\otimes\Gamma-I_0\otimes\Omega,\quad P_\parallel=P\otimes I.$$

To obtain the linearly independent modes of Eq. (10) we set[26]

$$C_m=\lambda^m C_0,$$

which gives

$$\lambda\begin{pmatrix}C_1\\C_0\end{pmatrix}=\begin{pmatrix}P_\parallel(H_0-w^2\epsilon)&-P_\parallel^2\\1&0\end{pmatrix}\begin{pmatrix}C_1\\C_0\end{pmatrix}. \tag{11}$$

In order to find similar modes in the perpendicular electrodes (*2* and *3* in Fig. 1c) we write the Schrödinger equation (5) as follows:

$$\begin{aligned}(w^2\epsilon+\Omega)\tilde{\Psi}(m,l)=&4\tilde{\Psi}(m,l)-\exp(-i\tilde{\gamma}m)\Gamma^*\tilde{\Psi}\\&\times(m,l+1)-\exp(i\tilde{\gamma}m)\Gamma\tilde{\Psi}(m,l-1)\\&-\tilde{\Psi}(m+1,l)-\tilde{\Psi}(m-1,l).\end{aligned} \tag{12}$$

Introducing again the column vector $\tilde{C}_l$ which describes the amplitudes of the lth line along the x-direction, we obtain from Eq. (12)

$$(w^2\epsilon-\tilde{H}_0)\tilde{C}_l+P_\perp\tilde{C}_{l-1}+P_\perp^*\tilde{C}_{l+1}=0, \tag{13}$$

where

$$\tilde{H}_0=4I_0\otimes I-D\otimes I-D^+\otimes I-I_0\otimes\Omega,\quad P_\perp=P^*\otimes\Gamma.$$

Using the relation $\tilde{C}_l=\lambda^l\tilde{C}_0$ we obtain from (13) the linearly independent solutions in the electrodes *2* and *3*:

$$\lambda\begin{pmatrix}\tilde{C}_1\\\tilde{C}_0\end{pmatrix}=\begin{pmatrix}P_\perp(\tilde{H}_0-w^2\epsilon)&-P_\perp^2\\1&0\end{pmatrix}\begin{pmatrix}\tilde{C}_1\\\tilde{C}_0\end{pmatrix}. \tag{14}$$

From the Schrödinger equations (4) and (5) the following continuity equation for the probability density follows:

$$-\frac{w^2}{2}\frac{\partial\rho}{\partial t}=j^{(x)}_{m,l}-j^{(x)}_{m-1,l}+j^{(y)}_{m,l}-j^{(y)}_{m,l-1}, \tag{15}$$

where $\mathbf{j}_{m,l}=(j^{(x)}_{m,l},j^{(y)}_{m,l})$ is the probability current density. In particular, for the gauge $\mathbf{a}_0=(-\gamma y,0,0)$ in the electrodes *1* and *4* we have

$$\begin{aligned}j^{(x)}_{m,l}&=\mathrm{Im}\{\exp(i\tilde{\gamma}l)\psi^*_{m,l}\psi_{m+1,l}\},\\j^{(y)}_{m,l}&=\mathrm{Im}\{\exp(-i\tilde{a}\cos\omega t)\psi^*_{m,l}\psi_{m,l+1}\}.\end{aligned} \tag{16}$$

For the gauge $\mathbf{a}_0=(0,\gamma x,0)$ in the electrodes *2* and *3* we have

$$\begin{aligned}j^{(x)}_{m,l}&=\mathrm{Im}\{\tilde{\psi}^*_{m,l}\tilde{\psi}_{m+1,l}\},\\j^{(y)}_{m,l}&=\mathrm{Im}\{\exp(-i\tilde{a}\cos\omega t-i\tilde{\gamma}m)\tilde{\psi}^*_{m,l}\tilde{\psi}_{m,l+1}\}.\end{aligned} \tag{17}$$

From these expressions for the probability current density it is easy to find the stationary current carryed by the propagating mode with $|\lambda|=1$ in the x direction through the section m in the electrodes *1* and *4*:

$$J^{(x)}_m=\mathrm{Im}(\lambda\langle C_0|P_\parallel^*|C_0\rangle). \tag{18}$$

As with the mode propagating in the y direction in the electrodes *2* and *3* we have

$$J^{(y)}_l=\mathrm{Im}(\lambda\langle\tilde{C}_0|P_\perp^*|\tilde{C}_0\rangle). \tag{19}$$

Now let us consider the scattering region (Fig. 2) connected to four electrodes. Following Ando[26] we define

$$U(\pm)=(\mathbf{u}_1(\pm),\mathbf{u}_2(\pm),\cdots\mathbf{u}_{LM}(\pm)).$$

$$\Lambda(\pm)=\mathrm{diag}(\lambda_1(\pm),\lambda_2(\pm),\cdots\lambda_{LM}(\pm)).$$

Here $\mathbf{u}_i(\pm)$ are the solutions of Eq. (11) which correspond to the eigenvalue $\lambda_i(\pm)$. The signs ''$\pm$'' refer to the propagating and evanescent modes in the positive (negative) x direction in the electrodes *1* and *4*. Similar matrices $\tilde{U},\tilde{\Lambda}$ can be defined for the electrodes *2* and *3*.

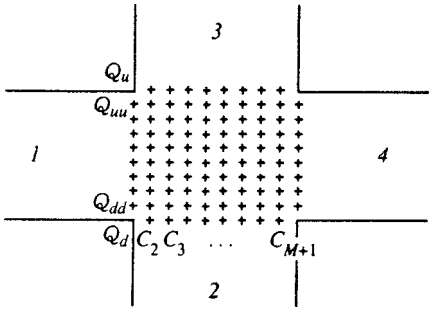

FIG. 2. Configuration of the lattice model for the scattering region.

For the modes which are superpositions only of the ''+'' type (or of the ''−'' type) we can write simple recurrence formulas[26]

$$C_{m+1}(\pm)=F(\pm)C_m(\pm)$$

with

$$F(\pm)=U(\pm)\Lambda(\pm)U^{-1}(\pm).$$

The same formulas take place for the electrodes *2* and *3*. These relations will be explored to define boundary conditions in the scattering region.

Next, we consider the solutions inside the scattering region which are shown in Fig. 2. At the boundary *1* there is the incident mode $C_1(+)$, and at the boundaries *2, 3*, and *4* there are only outgoing modes. Introduce vertical vectors $C_1, C_2, \ldots, C_{M+2}$ which describe the amplitudes of the wave function on a square lattice in the scattering region along the *y* direction as shown in Fig. 2, and a pair of horizontal vectors Q_u, Q_d which describe the amplitudes at the upper and down boundaries of the scattering region. The aim is to write closed equations for these vectors using the boundary conditions. The boundary conditions are that the wave is incident only through the left boundary *1* and is given as $C_1(+)$, and the other waves exit through all boundaries. Within the scattering region the equation for the amplitudes takes

$$(w^2\epsilon-H_0)C_m+P_\parallel C_{m-1}+P_\parallel^* C_{m+1}=0, \tag{20}$$

where $m=2,3,\ldots,M+1$.

In addition we consider the analogous equations at every boundary. At the boundary *1* we represent the vertical vector C_1 at site 1 as a superposition of the incident and reflected solutions:

$$C_1=C_1(+)+C_1(-).$$

The vector C_0 belonging to the electrode *1* can be expressed as

$$C_0=F^{-1}(+)C_1(+)+F^{-1}(-)C_1(-)$$
$$=F^{-1}(-)C_1+(F^{-1}(+)-F^{-1}(-))C_1(+).$$

Hence the solutions at the right edge of the electrode *1* are expressed in terms of the solutions at sites of the boundary of the scattering region and incident wave. As a result the equation for the amplitudes at the first vertical sites of the scattering region has the form

$$(w^2\epsilon-H_1)C_1+P_\parallel^* C_2=-P_\parallel[F^{-1}(+)-F^{-1}(-)]C_1(+),$$

$$H_1=H_0-P_\parallel F^{-1}(-).$$

At the boundary *4* we can write similar equations

$$(w^2\epsilon-H_4)C_{M+2}+P_\parallel C_{M+1}=0,$$

$$H_4=H_0-P_\parallel^* F(+).$$

Since at the boundaries *2* and *3* we have different gauges, we introduce two additional matrices which transform the primary gauge of the wave function:

$$\hat{E}_{u1}=\mathrm{diag}(\exp(i\tilde{\gamma}m(M+1)/2))\otimes I,$$

$$\hat{E}_{u2}=\mathrm{diag}(\exp(i\tilde{\gamma}m(M-1)/2))\otimes I.$$

If we take into account these gauge transformations, the equation for the vector Q_u at the boundary *3* can be written

$$(w^2\epsilon-H_3)E_{u1}Q_u+P_\perp E_{u2}Q_{uu}=0,$$

$$H_3=\tilde{H}_0-P_\perp^*\tilde{F}(+).$$

At the boundary *2*

$$(w^2\epsilon-H_2)E_{u1}^*Q_d+P_\perp^* E_{u2}^*Q_{dd}=0,$$

$$H_2=\tilde{H}_0-P_\perp\tilde{F}^{-1}(-).$$

Here Q_{uu}, Q_{dd} denote the horizontal vectors adjacent to the vectors Q_u, Q_d as shown in Fig. 2.

By means of these relations it is easy to write the Schrödinger equation for the amplitudes at the sites of the scattering region in closed form as $\hat{K}X=Y$ where $X=\mathrm{col}\times(C_1,C_2,\ldots,C_{M+2},Q_u,Q_d)$ with the known matrix K and vector Y. Numerical solution of this equation gives the solution inside the scattering region, at its boundary, and thereby at each electrode. For the simpler *L*- and *T*-structures shown in Fig. 1a and 1b the solutions are easily obtained if we set the solutions in the excluded electrodes equal to zero.

In conclusion we comment on the choice of the matrix Γ defined in (7). For an infinite matrix Γ we have the unitary condition $\Gamma\Gamma^+=1$. If we were to truncate directly the matrix Γ in the numerical calculations the unitary condition would be violated. In turn this would give rise to breakdown of the probability current conservation and what is more crucial, to appearance of undesirable exponentially growing and decaying propagating solutions with small exponents. In order to avoid this difficulty we introduce a new Hermitian matrix W which determines the matrix Γ as follows

$$\Gamma=\exp(i\tilde{\alpha}W/2), \tag{21}$$

where

$$W=\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 1 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 1 & \cdots \\ \vdots & & & & \end{pmatrix}.$$

Although in the computer simulations the matrix W is truncated to a finite dimension L, the relation (21) preserves the unitarity of the matrix Γ.

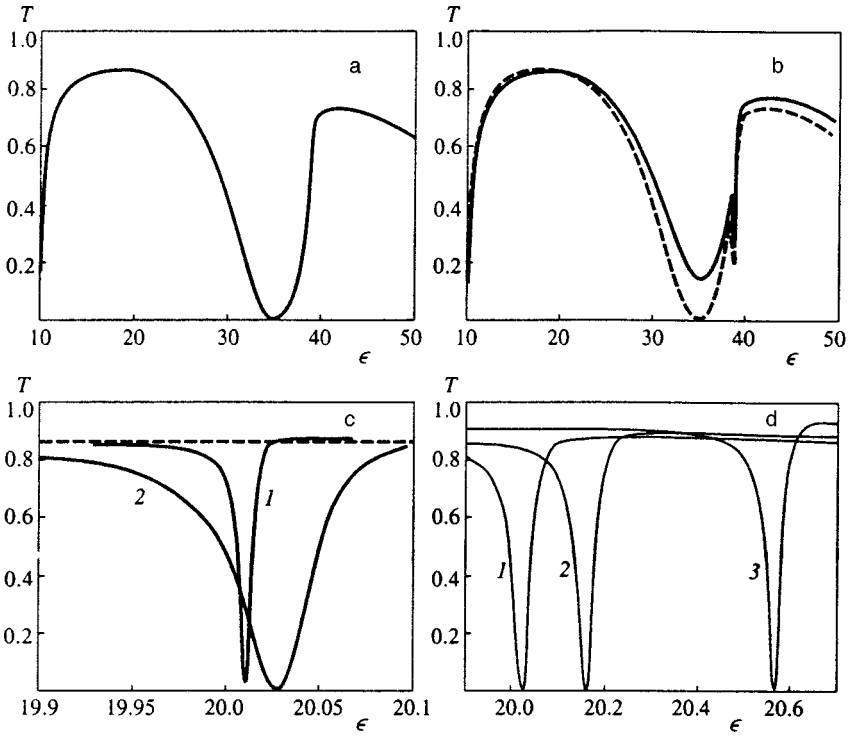

FIG. 3. The energy dependence of the transmission in the *L*-structure under the influence of the radiation field: a—stationary transmission $a=0$ for zero magnetic field; b—the frequency of a radiation field is resonant with transitions between the edges of the first subband and the second one $\omega=29.6$, $a=0.1$, $\gamma=0$. The solid line represents the case in which the polarization of the radiation field is perpendicular to the input electrode, the dashed line shows the case, in which the polarization is parallel to the input electrode. c—The frequency of the radiation field is tuned to transitions between the bound state and the first subband $\omega=10.82$, $\gamma=0$. The curve *1* corresponds to $a=0.2$, the curve *2* to $a=0.5$. The dashed line shows the steady case. d—Similar resonant dips with applied magnetic field produced by the radiation field: $\omega=10.82$, $a=0.5$. The curves *1, 2,* and *3* correspond to $\gamma=0,2,4$ respectively.

2.2. Numerical results

We begin by considering the simplest *L*-structure (Fig. 1a). It has only one bound state with energy $\epsilon_0=0.9291\pi^2$.[14] A magnetic field slightly increases this value. Consider at first the case when the frequency of the radiation field is tuned to transitions between the edges of the second and first subbands, $\omega\approx 3\pi^2$. When the polarization **a** of the radiation field is perpendicular to the input electrode of the structure the dipole matrix element mixing states of the second and first subbands equals

$$\langle 1|y|2\rangle=\int dy f_1(y) y f_2(y)\neq 0,$$

where $f_1(y)=\sqrt{2}\sin(\pi(y-1/2))$, $f_2(y)=\sqrt{2}\sin(2\pi(y-1/2))$. If the polarization of the radiation field is parallel to the input electrode the dipole matrix element $\langle k|x|k'\rangle$ calculated in terms of the incident modes $\exp(ikx)$ is less than $\langle 1|y|2\rangle$ because of the oscillatory behavior of the functions $\exp(ikx)$. Since the square of the dipole matrix element determines the radiation field effect, the electron transmission strongly depends on the polarization of the field as is in fact seen from Fig. 3b.

Second, consider the case when the frequency of the radiation field is resonant with transitions between the bound state energy and the first subband. For zero magnetic field the radiation field induces the deep narrow resonant dip shown in Fig. 3c. The width of the resonant dip depends sensitively on the amplitude of the radiation field. Figure 3d shows the shift of the resonant dip versus the applied magnetic field.

The structure intermediate between the *L*- and *X*-structure is the *T*-structure (Fig. 1b). Like the *L*-structure, it has only one bound state provided that the whole structure has the same width, but there are two ways to direct an incident electron: through the electrode *1* and through the electrode *2*. For the former case the transmissions T_{12} and T_{13} coincide, provided that $\gamma=0$. The polarization of a radiation field is chosen perpendicular to the input electrode along the *y*-axis (Fig. 1b). Consider the dipole matrix element between the bound state and the propagating one, $\langle\chi_1|y|\psi_{k,1}\rangle$, where $\chi_1(x,y)$ denotes the bound state with the energy $\epsilon_1=7.98$ and $\psi_{k,1}(x,y)$ is the propagating state for the steady case describing an electron incident on the first subband and the electrode *I*. Since both states are even relative to inversion $y\to -y$ around the symmetry line $y=0$ (for $\gamma=0$), this dipole matrix element vanishes and the radiation field produces no effect. In fact, our numerical calculations show that if the incident electron propagates in the first subband at zero magnetic field there are no resonant phenomena resulting from the radiation field.

There are two ways to the dipole matrix element can be finite. The first one is to apply an external magnetic field, and the second one is to consider electron transport in the second subband. These possibilities are shown in Fig. 4a and 4b. The steady transmissions through the *T*-structure are shown by thin lines. One can see that a magnetic field makes the transmissions T_{12} and T_{13} nonequivalent. Application of the radiation field gives rise to resonant dips which are very narrow, with widths proportional to the square of the amplitude of the radiation field. In the vicinity of the resonance the transmission T_{13} exceeds the transmission T_{12} which gives rise to the anomalous Hall effect. This effect was first demonstrated for the four-terminal structure.[21]

If an electron is incident on the second electrode, the dipole matrix element is not zero, and we expect resonant behavior for the transmissions to both electrodes *1* and *3*. However, the radiation field produces a resonant dip only for the transmission T_{23} (Fig. 4c). The reason for the absence of

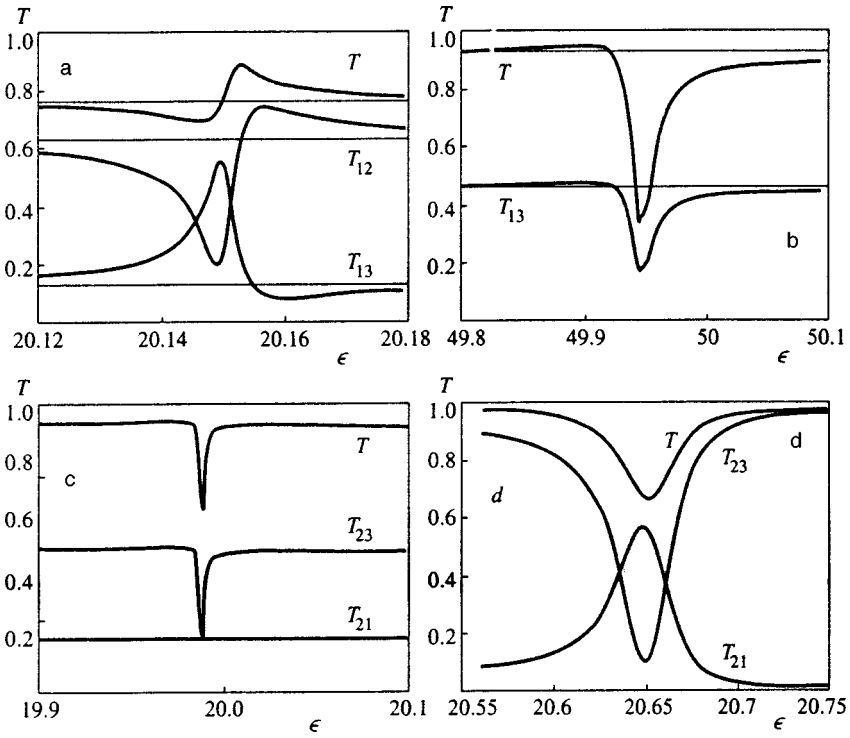

FIG. 4. Transmissions through the *T*-structure. a—Electron incidents on the electrode *1* (see Fig. 1b) and the first subband: $\omega=12$, $a=0.2$, $\gamma=2$. b—Electron incidents on the second subband with parameters $\omega=41.86$, $a=0.5$, $\gamma=0$. c—Electron incidents on the electrode *2*, $\omega=12$, $a=0.1$, $\gamma=0$. d—The same as in Fig. 4c with parameters $\omega=12$, $a=0.5$, $\gamma=4$.

a resonant dip for the transmission to the first electrode is related to the more complicated symmetry and will be given below. Application of an external magnetic field causes the resonant dips for all transmissions shown in Fig. 4d. Also because the bound state energy level is increased by an external magnetic field, the location of the resonant dips is slightly shifted as is seen from Fig. 4c and 4d.

Consider the four-terminal junction (Fig. 1c) which is an element of the Hall structures.[6,7,9,19] First, consider the radiation field effects for zero magnetic field, among which the most interesting is the resonant control of the electron transmissions by rotation of the radiation field polarization. As was mentioned for the *T*-structure this effect has a purely symmetric origin. However, the symmetry of the *X*-structure is higher than that of the *T*-structure. Moreover, the *X*-structure has two bound states. The one with the energy $\epsilon_1=6.55$ below the first subband is symmetrical relative to coordinate inversions $x\to -x$ or $y\to -y$, and the second with the energy $\epsilon_2=36.72$ below the second subband is antisymmetric.

As was observed for the *T*-structure, if the polarization of the radiation field is perpendicular to the input electrode and the frequency of the field is tuned to the transition between the first subband and the energy of the first symmetrical bound state, there are no field-induced resonant effects in the transmissions. The reason is that the propagating state is even, $\psi_{k,1}(x,y)=\psi_{k,1}(x,-y)$, and so we have $\langle\chi_1|y|\psi_{k,1}\rangle=0$, which means that the transmissions exhibit no resonant effect. On the other hand, there is no symmetry of the propagating state relative to $x\to -x$ due to electrons incident on the first electrode along the *x*-axis. Therefore, for the case of the polarization parallel to the input electrode *1* the dipole matrix element satisfies $\langle\chi_1|x|\psi_{k,1}\rangle\neq 0$. In fact, one can see from Fig. 5a a narrow resonant peak in the transmission T_{14}. However, as in the case of the *T*-structure, surprisingly, there are no resonant effects for the transmissions to the electrodes *2* and *3*.

To understand this following[21] we perform the gauge transformation

$$\psi(\mathbf{r},t)=\exp(i\mathbf{ar}\cos\omega t)\,\phi(\mathbf{r},t),$$

and substitute it into Eq. (3). As a result we obtain the following equation for the amplitudes $\phi_n(\mathbf{r})$ of expansion (6):

$$(\epsilon+n\omega)\,\phi_n=(i\nabla+\mathbf{a}_0(\mathbf{r}))^2\phi_n+\frac{i\omega}{2}(\mathbf{ar})(\phi_{n+1}-\phi_{n-1}). \tag{22}$$

Since we have assumed that the radiation field is resonant with transitions between the first bound state and the propagating one, for small perturbations we can restrict ourselves to two states ϕ_0 and ϕ_{-1} in Eq. (22) satisfying the following equations:

$$\nabla^2\phi_0+\epsilon\phi_0=-\frac{i}{2}(\mathbf{ar})\omega\phi_{-1}, \tag{23}$$

$$\nabla^2\phi_{-1}+(\epsilon-\omega)\,\phi_{-1}=\frac{i}{2}(\mathbf{ar})\omega\phi_0, \tag{24}$$

where the functions $\phi_0(\mathbf{r})$ and $\phi_{-1}(\mathbf{r})$ correspond to the propagating and the bound states, respectively. For the resonant case $\epsilon-\omega\approx\epsilon_1$ we can write the truncated Green's function for the left side of Eq. (24)

$$G_1(\mathbf{r},\mathbf{r}',\epsilon)\approx\frac{\chi_1(\mathbf{r})\chi_1^*(\mathbf{r}')}{\epsilon-\epsilon_1-i\delta} \tag{25}$$

where δ accounts for the finite width of the level because of coupling of the structure with the electrodes and the mixing

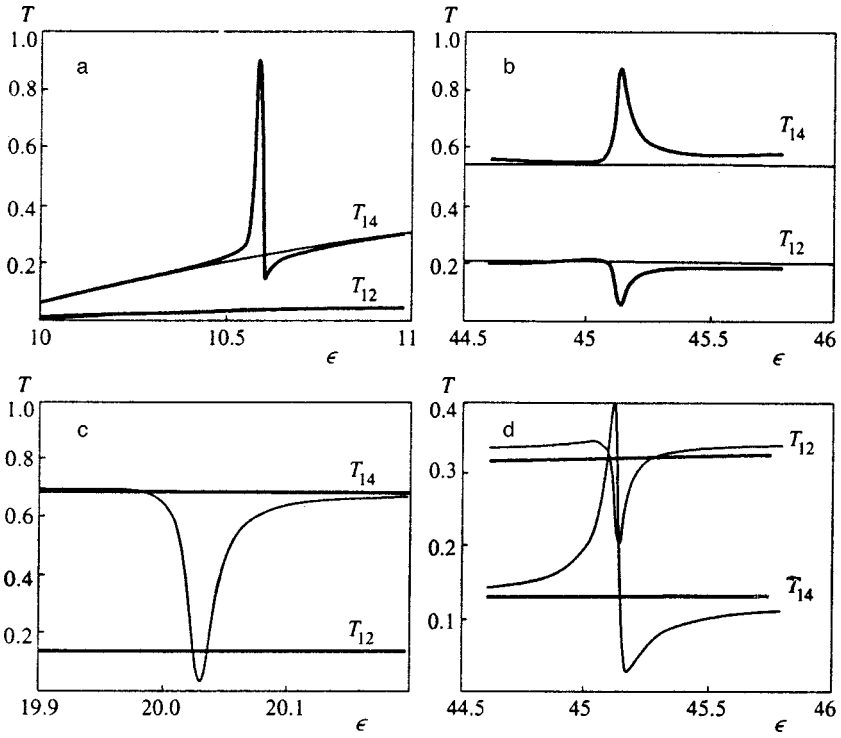

FIG. 5. Transmissions through the *X*-structure for zero magnetic field. In all pictures the electron is incident on the electrode *1*. a—The frequency of the radiation field is tuned to transitions between the energy of the first bound state and bottom of the first subband, $\omega=4$. The case of the polarization of the radiation field along the input electrode is shown by thick lines, and the case of the polarization perpendicular to the input electrode is shown by thin lines. In both cases the amplitude of the radiation field is $a=0.05$. The electron is incident on the first subband ($n=1$). b—$\omega=38.45$, $a=0.5$, the electron is incident on the first subband. c—The frequency is tuned to transitions between the second bound state and the first subband. $\omega=16.715$, $a=0.2$. d—As in Fig. 5b, but the electron is incident on the second subband ($n=2$).

of propagating states with the bound state by the radiation field. Then a solution of Eq. (24) can be expressed via the Green's function (25) as follows:

$$\phi_{-1}(\mathbf{r})=\frac{i\omega\chi_1(\mathbf{r})}{2(\epsilon-\omega-\epsilon_1-i\delta)}\times\int\chi_1^*(\mathbf{r}')(\mathbf{ar}')\phi_{0h}(\mathbf{r}')d^2\mathbf{r}'. \tag{26}$$

Substituting Eq. (23) and carrying out a similar procedure of expression in terms of the Green's function, we finally obtain

$$\phi_0(\mathbf{r})=\phi_{0h}(\mathbf{r})+\frac{\omega^2 d_{10}}{4(\epsilon-\epsilon_1-\omega-i\delta)}\times\int G(\mathbf{r},\mathbf{r}',\epsilon)(\mathbf{ar}')\chi_1(\mathbf{r}')d^2\mathbf{r}'. \tag{27}$$

Here $\phi_{0h}(\mathbf{r})$ is the solution for a switched off radiation field and

$$d_{10}=\int\chi_1^*(\mathbf{r})(\mathbf{ar})\phi_{0h}(\mathbf{r})d^2\mathbf{r}$$

is the dipole matrix element between the bound state and the propagating one.

Similar to (26) and (27) we can write a solution of Eq. (24) for the case when the frequency of the radiation field is resonant with transitions between the second bound state and the Fermi energy of the incident electron $\epsilon+\omega\approx\epsilon_2$

$$\phi_0(\mathbf{r})=\phi_{0h}(\mathbf{r})+\frac{\omega^2 d_{20}}{4(\epsilon-\epsilon_2+\omega-i\delta)}\times\int G(\mathbf{r},\mathbf{r}',\epsilon)(\mathbf{ar}')\chi_2(\mathbf{r}')d^2\mathbf{r}',$$

$$d_{20}=\int\chi_2^*(\mathbf{r})(\mathbf{ar})\phi_{0h}(\mathbf{r})d^2\mathbf{r} \tag{28}$$

In order to analyze the transmission on the basis of Eq. (27) we need the following symmetry properties of the Green's function in the *X*-structure:

$$G(x,y;x',y',\epsilon)=G(-x,y;-x',y',\epsilon)=G(x,-y;x',-y',\epsilon). \tag{29}$$

Now let us return to the transmission T_{12} (Fig. 1c) for the case when the radiation field polarization is parallel to the input electrode (*x*-axis). From Eqs. (27) and (29) we can see that the last resonant term in (27) is odd relative to $x\to-x$ in the electrodes *2* and *3*, provided that the bound state $\chi_1(\mathbf{r})$ is even. Thus, the last term in (27) is not able to contribute to the propagating mode in the electrodes *2* and *3* because for the electron transport in the first subband it should be even with respect $x\to-x$. Next, since the last term in (27) is even with respect $y\to-y$, it contributes to the even transport mode ϕ_0 in the electrodes *1* and *4*. As we see from Fig. 5a computer calculations completely confirm that conclusion. If the incident electron belongs to the first subband with $n=1$ (even state relative to $y\to-y$), but the outgoing mode can be represent as a superposition of states of the first and second subband ($n=1,2$), these symmetry restrictions are removed for the polarization parallel to the input electrode. As a result the radiation field induces resonant anomalies in the transmissions T_{12}, T_{13} (see Fig. 5b). Briefly, this symmetry rule can be formulated as follows. If the parity of the state excited by the dipole transition $(\mathbf{ar})\chi_1(\mathbf{r})$ does not conflict with the parity of the outgoing modes, then the transmission to the corresponding electrode can display resonant features,

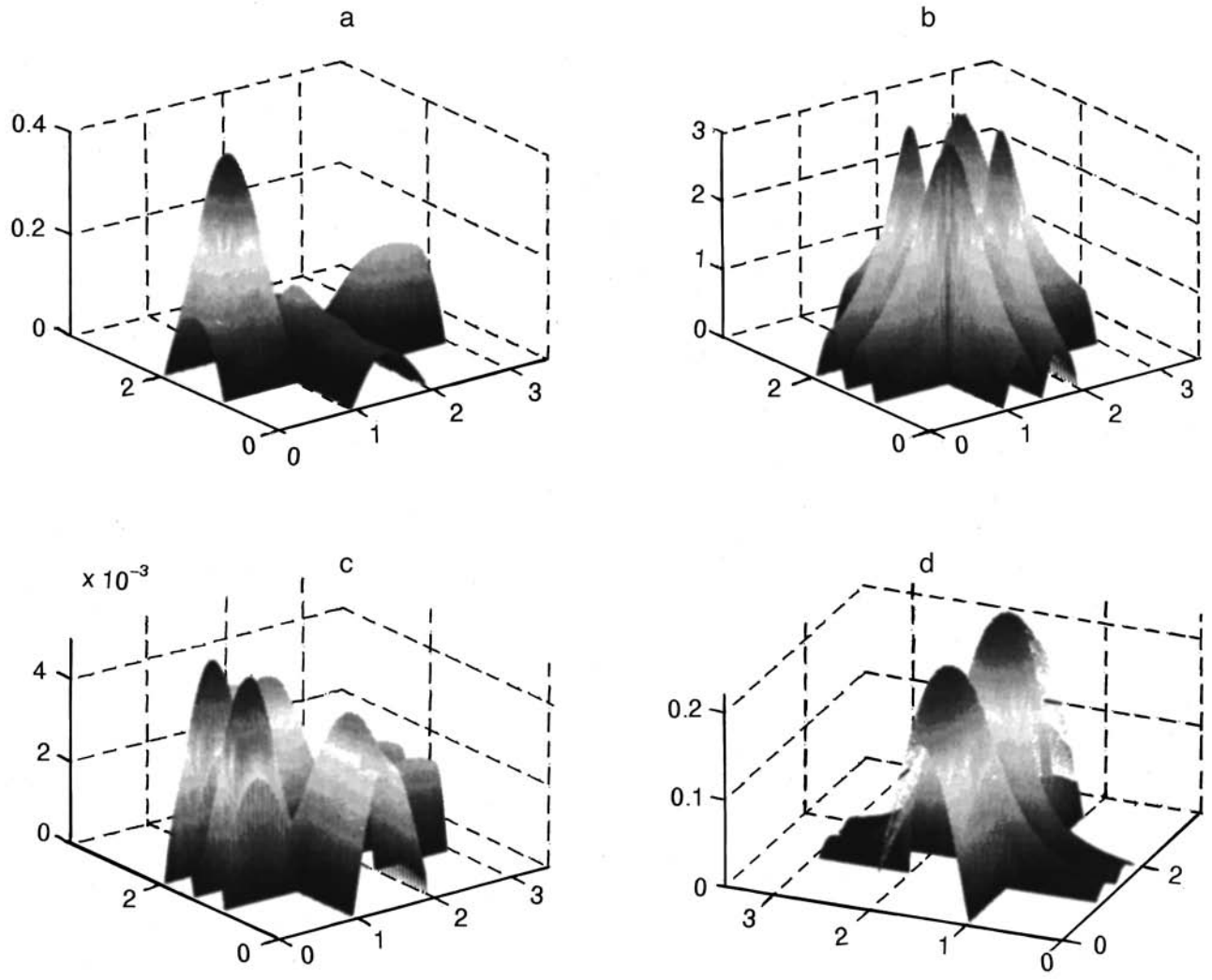

FIG. 6. Views of the amplitudes of the quasienergy wave function $\epsilon = 20.02$, $a = 0.2$, $\omega = 16.715$. a—$|\psi_0(\mathbf{r})|$, b—$|\psi_1(\mathbf{r})|$, c—$|\psi_{-1}(\mathbf{r})|$, d—$|\psi_0(\mathbf{r}) - \psi_{0h}(\mathbf{r})|$. Definitions of the amplitudes are given in (6).

and vice versa. Later we will demonstrate numerous examples of the application of this symmetry rule.

First, we apply the symmetry rule to the case of radiation field mixing of the second bound state, which is odd relative to $x \to -x$ or $y \to -y$. For the radiation field polarization directed parallel to the input electrode the dipole matrix element d_{20} vanishes, and consequently there are no radiation field induced effects. In the opposite case, when the polarization is perpendicular to the input electrode, the dipole matrix element is not equal to zero. However, the radiation field contribution to the electrodes *2* and *3* described by the last term in Eq. (26) is odd, opposite to the symmetry of the incident mode. So the transmissions to the electrodes *2* and *3* coincide with steady results, as shown in Fig. 5c. Finally, Fig. 5d shows the case when the incident electron belongs to the second subband. In contrast to the case in Fig. 5b, the field-induced effects take place when the polarization is perpendicular to the input electrode. Note that the same symmetry arguments explain the absence of radiation field effects in the electron transmission from electrode *2* to electrode *1* shown in Fig. 4c for the *T*-structure.

To confirm the approach using the quasi energy amplitudes (26) and (27) and to illustrate symmetry rules, in Fig. 6 we present numerical solutions of the full Schrödinger equation (4). In Fig. 6a, 6b, and 6c the amplitudes $\psi_n(\mathbf{r})$ with $n = 0, 1, -1$ respectively are shown for parameters corresponding to the case shown in Fig. 5c. One can see that, in fact, only two amplitudes, ψ_0, ψ_1, are important in the resonant case. Moreover, in agreement with Eq. (28) we see that the amplitude $\psi_1(\mathbf{r})$ exactly reproduces the second bound wave function χ_2 and that the next amplitude ψ_{-1} is negligible. Second, the difference between the radiation field perturbed solution ψ_0 and the steady solution ψ_{0h} is shown in Fig. 6d. One can see that symmetry of the outgoing part of this difference in the electrodes coincides with that predicted by the last term in Eq. (28). The parity is even in the electrodes *1* and *4* and is odd in the electrodes *2* and *3*. Also we can see from Fig. 6d that the odd contributions are decaying in the electrodes *1* and *3*.

As was mentioned above an external magnetic field breaks the symmetry of the structure, resulting in a more complicated picture of radiation field effects. Results of these calculations are presented in Fig. 7. Figure 7a corresponds to Fig. 5a, with the difference that we have $\gamma = 1$, and presents

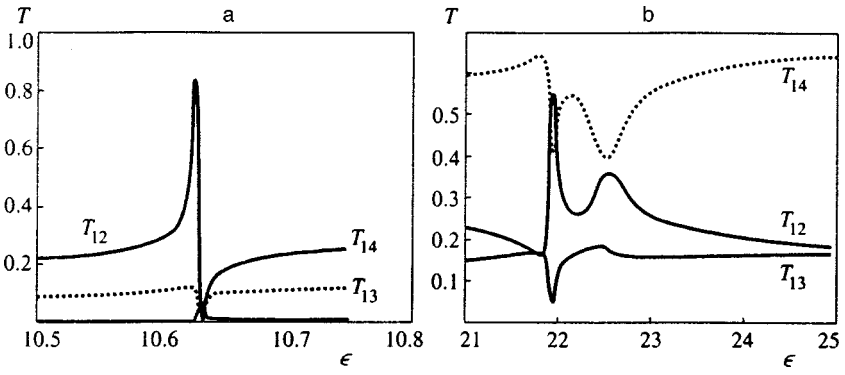

FIG. 7. The energy dependences of the transmissions in applied magnetic field in the *X*-structure. a—$\omega=4$, $a=0.05$, $\gamma=1.0$, the polarization of the radiation field is parallel to the input electrode: b—$\omega=15.08$, $a=0.5$, $\gamma=2$, polarization of the radiation field is perpendicular to the input electrode.

a case in which the radiation field excites the first bound state in the first subband. One can see all transmissions undergo resonant peaks or dips, in contrast to Fig. 5a. Figure 7b presents a case in which the frequency of the radiation field is tuned to $\omega=(\epsilon_2-\epsilon_1)/2$. One can see that exciting of two bound states results in two resonant peaks in the transmissions. The first bound state gives rise to sharp resonant peaks and dips, while the second produces wide peaks and dips.

From Fig. 7b we can see that in some narrow region of energies the transmission T_{13} coincides with T_{12} and may even slightly exceed it. Obviously, it would give rise to the negative Hall resistance as was shown in Ref. 21. Moreover we can see from Fig. 7b that the transmission T_{12} undergoes peaks, while the transmission T_{13} does dips. As a result we may observe resonant peaks in the Hall resistance as is demonstrated in Fig. 8.

From Fig. 7a and 7b (see also Ref. 21) in the narrow region of resonance the transmission T_{13} can slightly exceed the transmission T_{12} in an external magnetic field. This means that the radiation field can even cause anomalies of the Hall resistance to be negative.[21] Figure 8 illustrates various resonant anomalies of the Hall resistance induced by the radiation field: dips (Fig. 8a) and peaks (Fig. 8b, 8c, and 8d). The resonance between the first bound state and the Fermi energy of the incident electron produces a resonant dip (Fig. 8a). In the case of the Fermi energy $\epsilon\approx(\epsilon_2+\epsilon_1)/2$ and $\omega\approx(\epsilon_2-\epsilon_1)/2$ the radiation field induces two wide peaks contributed by two bound states. Figure 8c shows that the radiation field transforms the dip in the Hall resistance[6] into a resonant peak. Finally, Fig. 8d shows a case like Fig. 8a, but the radiation field excites the second bound state.

3. CONCLUSION

The resonant behavior in the electron transmission arises because the radiation field resonantly substitutes the bound states into the state of the incident electron propagating through the scattering region of the structures to produce various interference phenomena. These phenomena are clearly seen in the current density patterns shown in Fig. 9. The resonant anomalies are very specific to the forms the structure and the type of bound state. The symmetry of the

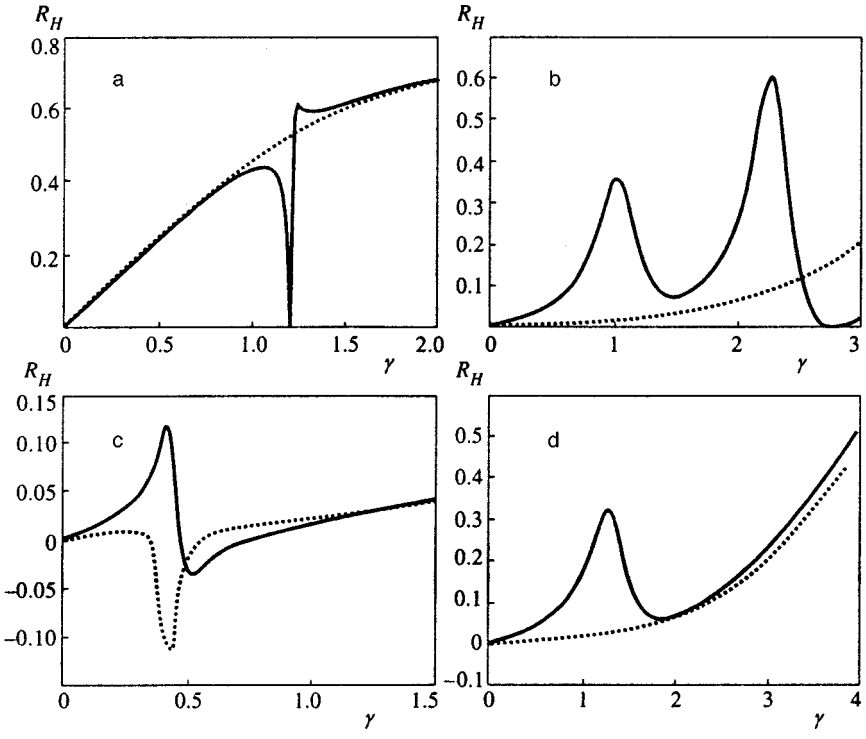

FIG. 8. The Hall resistence R_H in the *X*-structure versus an external magnetic field in response to the radiation field. The radiation field induced resistance is shown by solid lines, the steady resistance is shown by dotted line. a—$\epsilon=10.95$, $\omega=4.3$, $a=0.05$; b—$\epsilon=22$, $\omega=15.08$, $a=0.5$; c—$\epsilon=36.75$, $\omega=30.17$, $a=0.1$; d—$\epsilon=20.5$, $\omega=16.715$, $a=0.5$.

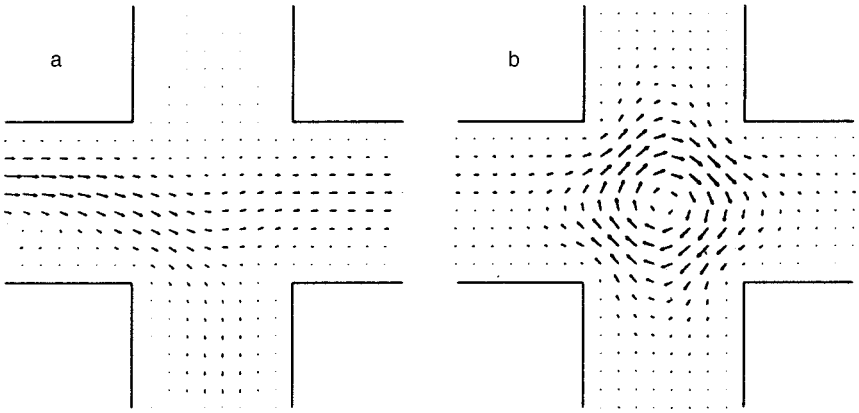

FIG. 9. Current density flowing in the *X*-structure when the radiation field is switched off (a) and switched on (b); $\omega=4.3$, $a=0.05$, $\epsilon=10.95$, $\gamma=1.21$.

structure and corresponding parity of the bound state plays an important role for the radiation field-induced effects because of symmetry rules for the dipole matrix element and for the resonant contribution, which is described by the right side of Eqs. (27) and (28). As a result the direction of the radiation field's polarization relative to the input electrode has strong effect on the resonant anomalies. This suggests an idea for controlling electron transmissions through the corresponding electrodes by simple rotation of the polarization of the radiation field.

In conclusion we give dimensional estimates for the radiation field which is able to produce resonant effects in the 2DEG. The Fermi energy in the semiconductor layered AlGaAs structures depends on the density of the electron gas and typically lies between 10 meV and 100 meV. The characteristic sizes d of the structures are between 100 nm and 1 μm. Accordingly, the frequency of the radiation field tuned to transitions between the bound states and the first electron subband will be roughly proportional to the Fermi energy. The amplitude of the radiation field is of order $E \approx E_F a\omega/ed \sim 10^3 - 10^4$ V/cm, where a and ω are dimensionless.

A. F. S. thanks K.-F. Berggren, P. Šeba and S. Wang for useful discussions of electron transport in waveguides. This work was partially supported by the INTAS-RFBR (Grant No. 95-IN-RU-657), RFFI (Grant No. 97-02-16305), Krasnoyarsk Science Foundation (Grant No. 7F0130) and the Foundation for Theoretical Physics in Slemeno, Czech Republic.

*)E-mail: zeos@zeos.krascience.rssi.ru

[1] B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhouhen, D. van der Marel, and C. T. Foxon, Phys. Rev. **60**, 848 (1988).
[2] D. A. Wrahamn, T. J. Thornton, R. Newbury, M. Pepper, H. Ajmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, J. Phys. C **21**, L209 (1988).
[3] M. L. Roukes, A. Sherer, S. J. Allen, Jr., H. G. Graighead, R. M. Ruthen, E. D. Beeke, and J. P. Harbison, Phys. Rev. Lett. **59**, 3011 (1987).
[4] Y. Takagaki, K. Gamo, S. Namba, S. Ishida, S. Takaoka, K. Murase, K. Ishibashi, and Y. Aoyagi, Solid State Commun. **68**, 1051 (1988).
[5] C. J. B. Ford, S. Washburn, M. Büttiker, C. M. Knoedler, and J. M. Hong, Phys. Rev. Lett. **62**, 2724 (1989).
[6] R. L. Schult, D. G. Ravenhall, and H. W. Wyld, Phys. Rev. B **41**, 12760 (1990).
[7] K. Amemiya and K. Kawamura, J. Phys. Soc. Jpn. **63**, 3087 (1994).
[8] R. L. Schult, D. G. Ravenhall, and H. W. Wyld, Phys. Rev. B **39**, 5476 (1989).
[9] F. M. Pecters, Superlattices Microstruct. **6**, 217 (1989).
[10] H. Sakaki, in *Proc. of the Intern. Symp. on Foundation of Quantum Mechanics in the Light of New Technology*, ed. by S. Kameluchi, Physical Society of Japan, Tokyo (1984), p. 94.
[11] J. P. Carini, J. T. Londergan, K. Mullen, and D. P. Murdock, Phys. Rev. B **46**, 15538 (1992).
[12] J. P. Carini, J. T. Londergan, K. Mullen, and D. P. Murdock, Phys. Rev. B **48**, 4503 (1993).
[13] P. Exner and P. Šeba, J. Math. Phys. **30**, 2574 (1989).
[14] P. Exner, P. Šeba, and P. Šťovíček, Czech. J. Phys., Sect. B **39**, 1181 (1989).
[15] Y. Avishai, D. Bessis, B. G. Giraud, and G. Mantica, Phys. Rev. B **44**, 8028 (1991).
[16] J. Goldstone and R. L. Jaffe, Phys. Rev. B **45**, 14100 (1992).
[17] P. Duclos and P. Exner, Rev. Math. Phys. **7**, 73 (1995).
[18] K. Lin and R. L. Jaffe, E-prints archive cond-mat/9601004.
[19] Y. Avishai, and Y. B. Band, Phys. Rev. Lett. **62**, 2527 (1989).
[20] K.-F. Berggren and Z.-L. Ji, Phys. Rev. **43**, 4760 (1991).
[21] E. N. Bulgakov and A. F. Sadreev, JETP Lett. **66**, 431 (1997).
[22] R. E. Peierls, Z. Phys. **80**, 763 (1933).
[23] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982).
[24] D. L. Haavig and R. Reifenberger, Phys. Rev. B **26**, 6408 (1982).
[25] *Handbook of Mathematical Functions*, ed. by M. Abramowitz and I. A. Stegun, U.S. GPO, Washington, DC (1964).
[26] T. Ando, Phys. Rev. B **44**, 8017 (1991).
[27] E. N. Bulgakov and A. F. Sadreev, J. Phys.: Condens. Matter **8**, 8869 (1996).

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

Enhancement of the quantum characteristics of light from an inversionless three-level Λ-laser in cascade schemes

Yu. M. Golubev*) and G. R. Ershov

Research Physics Institute at St. Petersburg State University, 193904 St. Petersburg, Russia
(Submitted 5 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 1971–1992 (December 1998)

We show that when a three-level laser is excited by Poisson light the lasing remains sub-Poisson, as asserted by H. Ritsch, M. A. M. Marte, and P. Zoller [Europhys. Lett. **19**, 7 (1992)], only if there is no reciprocal action on the exciting laser (weak coupling). In cascade schemes in which each three-level Λ-laser excites each succeeding Λ-laser with its light, any desired degree of shot-noise suppression can be achieved. © *1998 American Institute of Physics.* [S1063-7761(98)00512-5]

1. INTRODUCTION

A property of a laser medium consisting of three-level atoms[1] is that such a medium is capable of generating nonclassical (sub-Poisson) light without any additional effort on our part (e.g., regular pumping, as is the case with the two-level sub-Poisson laser examined in Ref. 2). This intrinsic property of the system can be connected with the fact that three-level atoms have been found to automatically form a kind of negative feedback loop, which in the final analysis stabilizes the radiation. Later in this paper we will see that in the variant of the Λ-scheme considered here (a weak pump field exciting the atom from the ground state to the highest state and a laser field that is in resonance with upper atomic level and the level in the middle), the pumping efficiency depends on the lasing power and decreases as a function of the power of the laser field.

Under these conditions the fluctuations of the pump field are to a certain extent balanced and do not give rise to corresponding fluctuations of the laser field, unlike the case without negative feedback. For instance, a positive fluctuation in the pump field first produces a positive fluctuation in the population of the upper laser level and then lasing. Next this gives rise to a negative fluctuation in the pump field and hence to a negative fluctuation in the laser field. Thus, within a certain time interval the fluctuations in lasing become balanced and do not depend on the fluctuations in the initial pump field. This mechanism explains why sub-Poisson lasing occurs even with Poisson pumping. This situation is very desirable, since realizing it requires no more efforts than building an ordinary laser. However, there are definite faults in it. Negative feedback follows only the intracavity field and cannot stabilize the light that has left the cavity, and since the process in which photons leave a cavity is stochastic, light outside the cavity cannot have the same good properties (in the sense of its quantum properties) as intracavity light.

Thus, a laser with three-level atoms as the active medium is probably the simplest system for demonstrating the possibility of generating sub-Poisson light in experiments. At the same time, in pure form this system has hardly any potential as a source of nonclassical light that could be used in highly precise and supersensitive optical measurements, since reducing shot noise by several orders of magnitude may prove very important in experimental practice. It is clear even on the qualitative level that here the limits are of a fundamental nature. If we examine this laser as an isolated entity, we will never be able to improve it without improving something related to it (e.g., by replacing Poisson pumping with sub-Poisson). Nevertheless, we believe that one obvious way of improving the quantum characteristics of the laser field should be analyzed. The possibility can be realized in cascade schemes, in which each preceding laser with moderate quantum characteristics excites the active medium of the succeeding laser, whose quantum characteristics of the light become more pronounced as a result of such excitation, etc.

In constructing the quantum theory of a three-level Λ-laser we will not study the most general case, discussed in Ref. 1, and will limit ourselves to the physical conditions in which there is no population inversion in the laser transition and yet there is gain. This choice of model is not very important and is related only to the fact that to demonstrate the effect of enhancement of the quantum properties of light in cascade schemes it is sufficient to consider a particular case.

The following problem arises in theoretical studies of cascade schemes. Formally, the process of excitation of the first laser in the cascade by coherent (Poisson) light can be taken into account by introducing a *c*-number exciting field into the interaction Hamiltonian. Here, as we have said earlier, lasing develops in such a way that the laser field becomes sub-Poisson, so that the second, third, etc., lasers in the cascade are excited by nonclassical light. Such pumping cannot be taken into account by *c*-numbers in the Hamiltonian. The theoretical approach must account for the fact that not only is the laser field nonclassical, but so is the pump field. In our investigation we use the method of a kinetic equation for the density matrix that describes all coherent fields participating in the process.

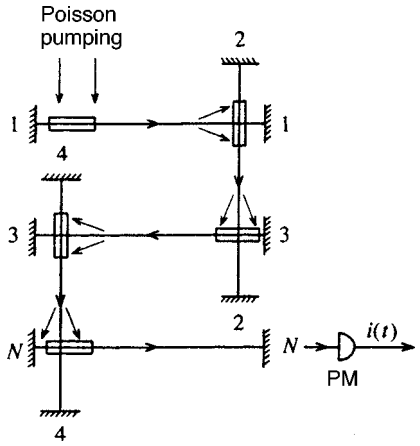

FIG. 1.

2. A CASCADE SCHEME WITH INTRACAVITY EXCITATION OF THE ACTIVE MEDIA

Figure 1 depicts a schematic of a thought experiment that we wish to discuss here. The experimental setup consists of a cascade of lasers, with the first laser, 1, excited by a prescribed Poisson electromagnetic field. Due to the excitation the laser emits light, which excites laser 2. The process is continued with laser 3, laser 4, etc. Here we have chosen the variant of intracavity excitation of the active media (often used in experiments) in which the media of the exciting and excited lasers occupy the same cavity. We must bear in mind that the excitation of the active medium may have a strong effect on the exciting laser. The most interesting case here is when the losses from the mode of the exciting laser are determined chiefly not by the escape of light through the mirror of the optical cavity of that laser but by the excitation of the active medium of the excited laser. Naturally, the conditions for lasing in the absence of an excited laser differ dramatically from those for lasing in the presence of such a laser. We call this the case of strong laser coupling. The other limit (weak coupling), where the presence or absence of the excited laser remains essentially unnoticed by the exciting laser, is also important and interesting from the physical viewpoint.

For a cascade scheme to operate, resonant conditions must be met. We can assume, for instance, that there are two types of active laser media (Fig. 2), which alternate from laser to laser. Clearly, from the mathematical viewpoint the two media are equivalent (all formulas for the two media coincide). Indeed, in both media coherent fields act on the

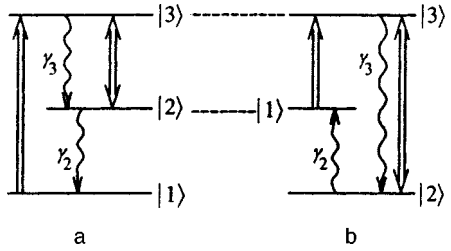

FIG. 2.

$|1\rangle-|3\rangle$ and $|2\rangle-|3\rangle$ transitions and incoherent processes act on the $|1\rangle-|2\rangle$ and $|2\rangle-|3\rangle$ transitions, processes that phenomenologically are described in the same way irrespective of their name (spontaneous decay or weak incoherent pumping).

In order to be specific, we assume that, say, the active medium of laser 1 (and all other lasers with odd numbers) consists of atoms with an energy level structure depicted in Fig. 2a, while the active media of even-numbered lasers consist of atoms in Fig. 2b. We also assume that in the first laser, excitation of atoms from the ground state to the upper laser state occurs on the $|1\rangle-|3\rangle$ transition. The transition is excited by an external Poisson electromagnetic field. Lasing occurs on the $|3\rangle-2\rangle$ transition. To simplify matters, we assume that all atomic transitions are allowed. However, in our calculations we will allow for spontaneous emission only on the $|3\rangle-2\rangle$ laser transition with a rate γ_3 and on the $|2\rangle-1\rangle$ laser transition with a rate γ_2. In the active media of laser 2 without coherent fields, due to spontaneous $|3\rangle-2\rangle$ emission and weak incoherent pumping (e.g., in a gas discharge or by a weak spectrally broad external electromagnetic field), under steady-state conditions all the atoms are in level $|1\rangle$. The coherent action of the light from laser 1 on the $|1\rangle-3\rangle$ transition populates the upper laser level and leads to lasing on the $|2\rangle-|3\rangle$transition.

The plan of the paper is as follows. In Sec. 3 we discuss the elements of semiclassical laser theory, important for our discussion. We derive auxiliary formulas for atomic populations and coherences in conditions of inversionless lasing and useful relationships for the lasing powers of the excited and exciting lasers.

Section 4 discusses the quantum statistical theory of lasing involving three-level atoms with the active medium excited by the light of an auxiliary Poisson or sub-Poisson laser. We develop the most important aspects of setting up the kinetic equation for the density matrix of the laser field and the pump field in the inversionless approximation and in the approximation of small photon fluctuations.

The photocurrent spectrum obtained in measurements of the light emitted by a single three-level laser excited by the light of a Poisson or sub-Poisson laser is discussed in Sec. 5.

In Sec. 6 we generalize the theory developed in Secs. 4 and 5 to the case of a cascade scheme consisting of N three-level lasers. The first laser in this scheme is excited by the light of a Poisson laser, its emission excites the next three-level laser, and so on.

In the Appendix we discuss in detail the method used in deriving the kinetic equation for the laser fields in the cascade.

3. SEMICLASSICAL THEORY OF LASING

3.1. Steady-state behavior of an atom in external pump and laser fields

In constructing a statistical theory we assume that the basic parameters of the problem fluctuate only weakly about the corresponding semiclassical values. This allows us to linearize the theory and hence simplify the mathematics. The explicit expressions for the semiclassical quantities, such as

the matrix elements of the atomic density matrix describing the behavior of the atom as it interacts with two classical fields, the laser field and the pump field, play an important role in this theory.

In this section we discuss the semiclassical theory of lasing with the medium consisting of atoms whose energy levels are depicted schematically in Fig. 2.

The behavior of the atoms is determined by a system of equations for the matrix elements of the atomic density matrix $\hat{\sigma}$, which with allowance for external fields and relaxation processes assumes the form

$$\begin{aligned}
&\dot{\sigma}_{11}=2\gamma_2\sigma_{22}-g_0(\alpha_0\sigma_{13}+\alpha_0^*\sigma_{31})=0,\\
&\dot{\sigma}_{33}=-2\gamma_3\sigma_{33}+g_0(\alpha_0\sigma_{13}+\alpha_0^*\sigma_{31})\\
&\qquad +g_1(\alpha_1\sigma_{23}+\alpha_1^*\sigma_{32})=0,\\
&\dot{\sigma}_{13}=-\gamma_3\sigma_{13}-g_0\alpha_0^*(\sigma_{33}-\sigma_{11})+g_1\alpha_1^*\sigma_{12}=0,\\
&\dot{\sigma}_{12}=-\gamma_2\sigma_{12}-g_0\alpha_0^*\sigma_{32}-g_1\alpha_1\sigma_{13}=0,\\
&\dot{\sigma}_{23}=-(\gamma_2+\gamma_3)\sigma_{23}-g_1\alpha_1^*(\sigma_{33}-\sigma_{22})+g_0\alpha_0\sigma_{21}=0.
\end{aligned} \tag{1}$$

This system of equations must be augmented by the normalization condition

$$\sigma_{11}+\sigma_{22}+\sigma_{33}=1. \tag{2}$$

Here α_0 and α_1 are the complex amplitudes of the pump and laser fields, which are normalized in such a way that the quantities $n_0=|\alpha_0|^2$ and $n_1=|\alpha_1|^2$ are the numbers of photons in the pump mode and the lasing mode. Other often-used quantities are the Rabi frequencies $\Omega_1=g_1|\alpha_1|$ and $\Omega_0=g_0|\alpha_0|$, where g_0 and g_1 are the constants of dipole interaction of the atom and the pump and laser fields, respectively.

In writing the system of equations we chose the interaction Hamiltonian in the form

$$\begin{aligned}
\hat{V}'=&ig_0(\alpha_0|3\rangle\langle 1|-\alpha_0^*|1\rangle\langle 3|)\\
&+ig_1(\alpha_1|3\rangle\langle 2|-\alpha_1^*|2\rangle\langle 3|),
\end{aligned} \tag{3}$$

where $|i\rangle\langle k|$ $(i,k=1,2,3)$ are the projection operators defined on the atomic states.

Clearly, for both physical situations depicted in Fig. 2 the system of equations is the same.

Under steady-state conditions (all time-derivatives are zero), Eqs. (1) become a system of algebraic equations and can be solved fairly easily. The explicit expressions for the steady-state matrix elements are

$$\sigma_{11}=1,\quad \sigma_{22}=\frac{i_0}{1+i_1},\quad \sigma_{33}=\frac{\tilde{i}_0}{(1+i_1)(1+\tilde{i}_1)}, \tag{4}$$

$$\begin{aligned}
&\sigma_{13}=\frac{g_0\alpha_0^*}{\gamma_3}\frac{1}{1+i_1},\quad \sigma_{12}=-\frac{g_1\alpha_1 g_0\alpha_0^*}{\gamma_2\gamma_3}\frac{1}{1+i_1},\\
&\sigma_{23}=-\frac{g_1\alpha_1^*}{\gamma_3}\frac{\tilde{i}_0}{(1+i_1)(1+\tilde{i}_1)}.
\end{aligned} \tag{5}$$

The dimensionless powers of the pump and laser fields can be expressed in terms of the parameters of the problem:

$$i_0=\frac{\Omega_0^2}{\gamma_3^2},\quad i_1=\frac{\Omega_1^2}{\gamma_3^2},\quad \tilde{i}_0=i_0\frac{\gamma_2}{\gamma_3},\quad \tilde{i}_1=i_1\frac{\gamma_2}{\gamma_3}. \tag{6}$$

In writing the steady-state solutions we assumed that

$$\Omega_0,\gamma_2\ll\gamma_3. \tag{7}$$

We see that under the adopted conditions only a small number of the atoms are excited. Specifically,

$$\sigma_{33},\sigma_{22}\ll\sigma_{11}, \tag{8}$$

with the population of state $|2\rangle$ always higher than that of state $|3\rangle$:

$$\sigma_{33}\ll\sigma_{22}. \tag{9}$$

In what follows we call the condition (7) the inversionless approximation.

3.2. Truncated lasing equations

Let us examine a physical system consisting of an exciting single-mode laser and an excited single-mode laser and consider it a two-model laser with complex-valued lasing amplitudes α_0 (the exciting mode) and α_1 (excited mode). It is well known that

$$\dot{\alpha}_1=-g_1N_1\sigma_{32}-\frac{\kappa_1}{2}\alpha_1, \tag{10}$$

where N_1 is the total number of atoms in the active medium, and κ_1 is the cavity width of the excited mode. Substituting the explicit expression in (5) for the matrix element $\sigma_{32}=\sigma_{23}^*$ in Eq. (10), we arrive at an equation of lasing in the excited mode. Obviously, it depends on the lasing in the exciting mode.

A similar equation can be written for α_0. Here we must bear in mind that the exciting mode interacts with two resonant media: the medium of the exciting laser, and the medium of the excited laser in the process of its excitation.

The sign of σ_{32} is an indication of whether we are dealing with gain or absorption in the laser transition. By separating the components proportional to the two-photon population difference $\sigma_{11}-\sigma_{22}$ we separate the terms (if they exist) responsible for Raman scattering.

From (5) we see that σ_{32} is negative, so that there is gain in this laser transition. Moreover, the coherence of the laser transition depends on the population of the upper laser level and not on the two-photon population difference, the opposite of the case of Raman scattering. Thus, bearing in mind the foregoing conjecture that there is no population inversion in the laser transition, we arrive at the unambiguous conclusion that here we are dealing only with inversionless lasing, which is a characteristic feature of multilevel atomic systems.[3]

We now write Eq. (10) explicitly in the inversionless approximation:

$$\dot{\alpha}_1=\left(\frac{\kappa}{2}\frac{\tilde{i}_1}{1+\tilde{i}_1}\frac{n_0}{n_1}-\frac{\kappa_1}{2}\right)\alpha_1=0. \tag{11}$$

Here n_0 and n_1 are the semiclassical solutions for the numbers of photons inside the first and second cavities, and the quantity

$$\kappa=\frac{\overline{\kappa}}{1+i_1}, \quad \overline{\kappa}=\frac{2g_0^2N_1}{\gamma_3}, \tag{12}$$

has the physical meaning of the rate of absorption of the exciting light in the excited medium (or the rate of excitation of the active medium of the excited laser). This rate depends on the lasing power of the excited laser. In the case of saturation ($\tilde{i}_1 \gg 1$), Eq. (11) yields a useful relationship:

$$\frac{n_0}{n_1}=\frac{\kappa_1}{\kappa}, \tag{13}$$

from which we can easily establish that

$$i_1=\frac{g_1}{g_0}\sqrt{\frac{\overline{\kappa}}{\kappa_1}i_0}, \quad \kappa=\frac{g_0}{g_1}\sqrt{\frac{\kappa_1\overline{\kappa}}{i_0}}. \tag{14}$$

4. THEORY OF LASING OF A SINGLE LASER WITH POISSON AND SUB-POISSON EXCITATION

4.1. The general form of the basic kinetic equation

We construct the quantum statistical theory of the lasing of a cascade scheme of lasers on the basis of the theory of lasing of a single laser. Here and in the sections that follow we assume that there is only one three-level laser whose active medium is excited by the light from another auxiliary laser. We find it convenient not to say anything specific about the auxiliary laser at this point, so that in the final stages of our discussion we can choose between Poisson excitation and sub-Poisson excitation. This problem was partially studied by Ritsch *et al.*[1] Here, however, we discuss not only the intracavity situation but also the aspects of observing such phenomena, and we will show that the results strongly depend on how the laser is excited. Two physical situation will be examined: when the exciting laser is ''oblivious'' of the exciting laser (weak laser coupling), and when reciprocal action is very important (strong coupling).

To describe the statistical properties of lasing we use the kinetic equation for the density matrix ρ of the laser field generated by the laser under investigation and the laser field of the exciting laser, assuming that they are of equal status in the theory. Usually the exciting field is assumed fixed. In our approach, however, we take into account the nonclassical nature of the exciting field and the reciprocal effect of the excited laser on the exciting laser, a characteristic effect in three-level schemes.

To specify the basic kinetic equation, we must assume that the atomic variables vary much more rapidly than the field variables. This standard requirement of kinetic theory can be satisfied under lasing conditions by assuming that the optical cavities of both lasers are high-Q cavities. If this condition is met, we can write the desired kinetic equation in the form

$$\dot{\rho}=\left(r_0\hat{S}_0-\frac{\kappa_0}{2}\hat{R}_0\right)\rho+\left(\hat{S}_1-\frac{\kappa_1}{2}\hat{R}_1\right)\rho. \tag{15}$$

Here the operators $\hat{R}_0$ and $\hat{R}_1$ are responsible for the decay of the intracavity quantum field oscillators of the exciting and excited lasers with rates κ_0 and κ_1, respectively; $\hat{S}_1$ and $\hat{S}_0$ are responsible for the nonlinear interaction of the laser fields and the active media; and r_0 is the average rate of excitation of the upper laser level of the active medium of the exciting laser.

We will attempt to write Eq. (15) explicitly by using the diagonal Glauber representation. Thus, we have $\rho=\rho(\alpha_0,\alpha_1,t)$, where α_0 and α_1, as before, are the complex-valued amplitudes of the pump and laser fields, which are now the eigennumbers of the corresponding annihilation operators:

$$a_j|\alpha_j\rangle=\alpha_j|\alpha_j\rangle, \quad [a_j,a_j^\dagger]=1, \quad j=0,1. \tag{16}$$

For the decay operators of the field oscillators in the Glauber approximation we have the well-known formulas

$$\hat{R}_j=-\frac{\partial}{\partial\alpha_j}\alpha_j+\text{c.c.} \tag{17}$$

Below we will find it convenient to a certain extent not to specify the explicit representation of $\hat{S}_0$, which naturally depends on what laser is chosen as the exciting one. This means, for instance, that in the variant of the theory based on the approximation of small photon fluctuations, which is used in this paper, the two coefficients referring to the exciting laser light, the decay rate Γ_0 of photon fluctuations and the statistical Mandel parameter ξ_0, will not be written explicitly.

Here and in what follows we consider two kinds of exciting lasers. The first uses the common laser, considered, e.g., in Ref. 4. As is known, the light of this laser in the saturation regime is purely Poisson. The explicit expression for $\hat{S}_0$ in the diagonal representation can be found in Ref. 5. We call this the case of Poisson excitation. The second version, sub-Poisson excitation, can easily be studied theoretically if in the first version we replace $\hat{S}_0$ by $\hat{S}_0-\hat{S}_0^2/2$ (see Ref. 2).

Although calculating $\hat{S}_1$ is one of the main problems of this theory, since it is a purely theoretical one and contains cumbersome calculations, we place it in the Appendix. Here we use the formulas from the Appendix and identify only the main steps that must be taken (see the Appendix).

As the first step, we write the fairly obvious equation for the density matrix $\hat{F}(t)$ of two-mode lasing (of the laser field and the pump field) and a single atom, whose energy level structure with allowance for atomic relaxations is depicted in Fig. 2. Then in this equation we pass to the diagonal Glauber representation, i.e., the equation for the density matrix $\hat{F}(\alpha_0,\alpha_1,t)$. This matrix is still an operator in relation to atomic variables. The convenience of using the diagonal representation lies, in particular, in the fact that the equation explicitly acquires terms responsible for the quantization of the electromagnetic field. The terms can be separated from the ''classical'' part (i.e., classical in relation to the field) by a very simple criterion: these are terms with derivatives with respect to the complex-valued amplitudes of the field, and by

discarding them we arrive at an equation that describes the behavior of a single atom (see Eqs. (1)) in external classical fields with the amplitudes α_0 and α_1.

In the next step we factorize the density matrix, i.e., write $\hat{F}=\hat{\sigma}\rho+\hat{\pi}$, where $\hat{\sigma}(t)$ is the density matrix describing the behavior of a single atom in two external ''classical fields'' with the amplitudes α_0 and α_1; $\rho(\alpha_0,\alpha_1,t)$ is the density matrix of the two-mode laser field; and $\hat{\pi}(\alpha_0,\alpha_1,t)$ is the matrix describing the correlations between the atom and the laser fields.

If we combine the initial equation for the density matrix $\hat{F}(\alpha_0,\alpha_1,t)$ of a single atom and two laser fields and the equation for the atomic density matrix $\hat{\sigma}$, we arrive at a system of equations for the matrices ρ and $\hat{\pi}$ (which depend on $\hat{\sigma}$ as a parameter). Solving the equation for $\hat{\pi}$, we find this quantity as a function of ρ. Inserting the result in the equation for ρ, we arrive at a closed equation that describes the behavior of only the field matrix density. A possible approach to this problem is to adopt the so-called kinetic approximation. This approximation assumes that the matter–field correlation time is much shorter than the relaxation time of the field variables. Here the correlation time is the period of revolution of the atom over the levels due to the interaction with the pump and laser fields and due to incoherent processes. On the other hand, the relaxation time is the photon lifetime in the optical cavity (this is common practice in laser problems).

The kinetic approximation makes it possible, in solving the equations for the atomic density matrix $\hat{\sigma}$ and the correlation matrix $\hat{\pi}$, to write only the steady-state solutions. This is quite obvious when we are dealing with purely atomic motion, as in the equation for $\hat{\sigma}$, since this motion is the fastest and the atom need only a short time to reach the steady state. The situation with the equation for $\hat{\pi}$ is somewhat more complicated, since it incorporates the atomic motion and the field motion. Here, however, we can use the adiabatic approximation, when in the first stage we simply ignore the variation of field variables, seek the steady state solution of the problem, and then allow for the temporal variations of the field variables in this steady-state solution.

The transition to steady-state equations transforms the system of differential equations into a system of algebraic equation, which can be solved fairly easily. This scheme makes it possible to obtain the explicit dependence $\hat{\pi}=\hat{\pi}(\rho,\hat{\sigma})$, i.e., the explicit expression for the rate of the variation of the density matrix of the laser field due to the interaction with the single atom.

Finally, we must allow for the presence in the cavity of N_1 statistically independent identical atoms. To do this we multiply the increment of the field density matrix ρ due to the interaction with a single atom by N_1. As a result, we obtain $\dot{\rho}$ as a function of ρ, i.e., the desired development of the field matrix due to the medium–field interaction.

4.2. The approximation of small photon fluctuations

The field characteristics usually discussed in laser physics are the amplitudes (the photon numbers) and phases (phase differences). Therefore, in a theoretical discussion it is convenient to shift from the complex-valued amplitudes α_j (j=0,1) to the polar coordinates u_j and φ_j:

$$\alpha_j=\sqrt{u_j}\exp(i\varphi_j),\quad j=0,1. \tag{18}$$

At this point we make an assumption that is common in problems of statistical laser physics. We assume that photon numbers in the exciting and excited lasers in steady-state lasing conditions fluctuate only slightly near their semiclassical steady-state values:

$$u_j=n_j+\varepsilon_j,\quad n_j\gg\varepsilon_j. \tag{19}$$

Of course, this is possible only under physical conditions such that there is a large buildup of photons in the cavity. The semiclassical solutions for n_j can be found in Eqs. (13) and (14).

Bearing in mind the semiclassical solutions and the results obtained in the Appendix, we can write the kinetic equation for the photon matrix density

$$\rho(\varepsilon_0,\varepsilon_1,t)=\int\int\rho(\varepsilon_0,\varphi_0,\varepsilon_1,\varphi_1,t)\,d\varphi_0\,d\varphi_1, \tag{20}$$

in the form

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\varepsilon_0,\varepsilon_1,t)=&\Gamma_0\frac{\partial}{\partial\varepsilon_0}(\varepsilon_0-\delta_1\varepsilon_1)\rho\\&+\Gamma_1\frac{\partial}{\partial\varepsilon_1}(\varepsilon_1-\delta_0\varepsilon_0)\rho+D_0\frac{\partial^2\rho}{\partial\varepsilon_0^2}\\&+D_1\frac{\partial^2\rho}{\partial\varepsilon_1^2}+D_{01}\frac{\partial^2\rho}{\partial\varepsilon_0\,\partial\varepsilon_1}+\{\cdots\}.\end{aligned} \tag{21}$$

We see that the fluctuations of the exciting field affect the fluctuations of the laser field (due to linear coupling, given by the coefficient δ_0), which was to be expected. But there is also a reciprocal effect (the coefficient δ_1), which stems from the behavior of a three-level scheme, in which the absorption coefficient in the excitation channel depends on the laser field (5). Here by $\{\cdots\}$ we have denoted the entire set of higher-order derivatives with respect to ε.

First we write the coefficients in the inversionless approximation, which do not depend on the type of exciting laser:

$$\begin{aligned}&\Gamma_1=\kappa_1\frac{i_1}{i_1+1}\frac{2\tilde{i}_1+1}{\tilde{i}_1+1},\quad \Gamma_1\delta_0=\kappa,\quad \Gamma_0\delta_1=\kappa_1\frac{i_1+\gamma_3/\gamma_2}{i_1+1},\\&D_1=-\kappa_1 n_1,\quad D_{01}=\kappa_1 n_1\frac{i_1+\gamma_3/\gamma_2}{i_1+1},\quad D_0=\Gamma_0\xi_0 n_0.\end{aligned} \tag{22}$$

These formulas can be used for any type of exciting laser, which is important because in future they will permit us to use this simple basis to study the cascade scheme. The type of exciting laser is determined by the values of the coefficients Γ_0 and ξ_0, where Γ_0 is the decay rate of photon fluctuations in exciting lasing (in the presence of an excited laser), and ξ_0 is the Mandel parameter, which characterizes the intracavity photon fluctuations of the exciting laser (in the absence of an excited laser).

Here we have $\xi_0=0$ for excitation by the light of a Poisson laser in the saturation regime and $\xi_0=-1/2$ for excitation by the light of a sub-Poisson laser in the saturation regime. For both cases $\Gamma_0=\kappa_0+\kappa$.

5. PHOTODETECTION OF LASER LIGHT

5.1. Equations for the average quantities

When detecting the light emitted by a single-mode laser, we can write the photocurrent spectrum as follows:

$$i_\omega^{(2)}=i_{\text{shot}}^{(2)}\left(1+\frac{2\kappa_1}{n_1}\,\text{Re}\int_0^\infty \overline{\varepsilon_1(0)\varepsilon_1(t)}\,\exp(i\omega t)\,dt\right), \tag{23}$$

where i_{shot} is shot noise. Thus, we must find an explicit representation for the average quantity $\overline{\varepsilon_1\varepsilon_1(t)}$. The basic kinetic equation (21) makes it possible, using standard procedures, to write the following system of equations:

$$\begin{aligned}
&\dot{\overline{\varepsilon_1\varepsilon_1(t)}}=-\Gamma_1\,\overline{\varepsilon_1\varepsilon_1(t)}+\Gamma_1\delta_0\,\overline{\varepsilon_1\varepsilon_0(t)}\,,\\
&\dot{\overline{\varepsilon_1\varepsilon_0(t)}}=-\Gamma_0\,\overline{\varepsilon_1\varepsilon_0(t)}+\Gamma_0\delta_1\,\overline{\varepsilon_1\varepsilon_1(t)}\,.
\end{aligned} \tag{24}$$

The solutions of this system depend on the initial conditions, i.e., on the expressions for $\overline{\varepsilon_1\varepsilon_0}$ and $\overline{\varepsilon_1^2}$ in steady-state conditions. To find these steady-state expressions we introduce an additional system of equations, which also follows from the basic kinetic equation:

$$\begin{aligned}
&\dot{\overline{\varepsilon_1^2}}=-2\Gamma_1\,\overline{\varepsilon_1^2}+2\Gamma_1\delta_0\,\overline{\varepsilon_0\varepsilon_1}+2D_1=0,\\
&\dot{\overline{\varepsilon_0\varepsilon_1}}=-(\Gamma_0+\Gamma_1)\overline{\varepsilon_0\varepsilon_1}+\Gamma_0\delta_1\overline{\varepsilon_1^2}+\Gamma_1\delta_0\,\overline{\varepsilon_0^2}+D_{01}=0,\\
&\dot{\overline{\varepsilon_0^2}}=-2\Gamma_0\,\overline{\varepsilon_0^2}+2\Gamma_0\delta_1\,\overline{\varepsilon_0\varepsilon_1}+2D_0=0.
\end{aligned} \tag{25}$$

Note that Eqs. (24) and (25) follow directly from (21), and the terms $\{\cdots\}$ incorporating higher-order derivatives make no actual contribution.

Next we use Eqs. (22) in the saturation regime, i.e., under the condition $\tilde{i}_1\gg 1$, but leave the coefficients Γ_0 and ξ_0 unspecified:

$$\begin{aligned}
&-2\kappa_1\,\overline{\varepsilon_1^2}+\kappa\delta_0\,\overline{\varepsilon_0\varepsilon_1}-\kappa_1 n_1=0,\\
&-(\Gamma_0+2\kappa_1)\overline{\varepsilon_0\varepsilon_1}+\kappa_1\overline{\varepsilon_1^2}+\kappa\overline{\varepsilon_0^2}+\kappa_1 n_1=0,\\
&-\Gamma_0\,\overline{\varepsilon_0^2}+\kappa_1\overline{\varepsilon_0\varepsilon_1}+\Gamma_0\xi_0 n_0=0.
\end{aligned} \tag{26}$$

The solution of this system of algebraic equations can be written in the form

$$\overline{\varepsilon_0\varepsilon_1}=\frac{2\kappa_1\Gamma_0 n_1(\xi_0+1/2)}{(2\Gamma_0-\kappa)(\Gamma_0+2\kappa_1)}, \tag{27}$$

$$\xi_1=-\frac{n_1}{2}+\frac{\kappa\Gamma_0 n_1(\xi_0+1/2)}{(2\Gamma_0-\kappa)(\Gamma_0+2\kappa_1)}, \tag{28}$$

where $\xi_1=\overline{\varepsilon_1^2}/n_1$ is the statistical Mandel parameter.

Clearly at $\xi_0=-1/2$ there can be no one-time cross-correlation, and $\xi_1=-1/2$.

Solving the system of equations (24) and (25), we arrive at an explicit expression for the photocurrent spectrum:

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}}=1-2\kappa_1^2\,\frac{\Gamma_0^2+\omega^2-\Gamma_0\kappa(1+\xi_0)}{[\kappa_1(2\Gamma_0-\kappa)-\omega^2]^2+\omega^2(\Gamma_0+2\kappa_1)^2}. \tag{29}$$

It should be recalled that in the cases of Poisson and sub-Poisson excitation in the saturation regime, $\Gamma_0=\kappa_0+\kappa$, and ξ_0 is equal to 0 and 1/2, respectively.

Below we analyze this expression in the two most interesting limits: weak and strong coupling of the exciting and excited lasers. Since the intracavity volumes of the two lasers intersect in our treatment, the medium of the excited laser introduces additional resonator losses in the exciting laser. Thus, the total field losses in the cavity are determined by $\kappa_0+\kappa$. If

$$\kappa_0\gg\kappa, \tag{30}$$

this additional effect can be ignored, assuming that there is no way in which the lasing of the exciting laser depends on that of the excited laser. We call the inequality (30) the condition for weak laser coupling. At the same time it is clear that the opposite case of strong coupling is also real. Here the main fraction of the light from the exciting laser is used to excite the active medium of the excited laser.

5.2. Weak laser coupling

If the condition (30) holds, formula (29) becomes

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}}=1-2\kappa_1^2\,\frac{\Gamma_0^2+\omega^2}{[2\Gamma_0\kappa_1-\omega^2]^2+\omega^2(\Gamma_0+2\kappa_1)^2}. \tag{31}$$

We see first that in the limit of weak coupling of the exciting and excited lasers, the photocurrent spectrum does not depend on the Mandel parameter ξ_0 in the exciting laser. It is entirely unimportant whether the excitation id Poisson, sub-Poisson, or even super-Poisson. This, in particular, means that we can apply a simplified version of the theory. We need not consider the laser light and the pump light as two subsystems equal in status and write the equation for the density matrix of both laser fields. When the laser coupling is weak, we can always introduce the pump field as a *c*-number into the corresponding interaction Hamiltonian. For any type of exciting laser, the physical picture in the laser wave is always the same and coincides with that in Poisson excitation.

Secondly, maximum suppression of shot noise is attained, as usual, at zero frequency. The extent of this suppression does not depend on the type of excitation, and the relative depth of the dip reaches 1/2.

5.3. Strong laser coupling

As we have just seen, in weak laser coupling the statistical properties of the light from the exciting laser have no effect on the statistical properties of the light from the excited laser. This is quite natural, since only a small fraction of the initial radiation participates in the excitation process,

and this radiation does not retain the properties of all the light. Hence our interest in the opposite case of strong laser coupling,

$$\kappa_0 \ll \kappa, \tag{32}$$

in which the statistical properties of the exciting radiation are assumed to affect lasing. Then instead of (29) we have the following:

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}} = 1 + \frac{2\kappa_1^2}{x_+^2 - x_-^2}\left(\frac{\xi_0^{(0)}\kappa^2 + x_-^2}{\omega^2 + x_-^2} - \frac{\xi_0^{(0)}\kappa^2 + x_+^2}{\omega^2 + x_+^2}\right). \tag{33}$$

Here $x_\pm$ stand for the roots of the biquadratic equation

$$(\kappa_1\kappa + x^2)^2 - x^2(\kappa + 2\kappa_1)^2 = 0. \tag{34}$$

The roots are the simplest in two limits:

$$x_-^2 = \kappa^2, \quad x_+^2 = \kappa_1^2, \quad \kappa \gg \kappa_1, \tag{35}$$

$$x_-^2 = 4\kappa_1^2, \quad x_+^2 = \kappa^2/4, \quad \kappa \ll \kappa_1. \tag{36}$$

Let us write separate formulas for excitation by light from Poisson and sub-Poisson lasers.

We have

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}} = 1, \quad \kappa \gg \kappa_1, \tag{37}$$

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}} = 1 - \frac{1}{2}\frac{4\kappa_1^2}{\omega^2 + 4\kappa_1^2} + \frac{1}{2}\frac{\kappa^2/4}{\omega^2 + \kappa^2/4}, \quad \kappa \ll \kappa_1 \tag{38}$$

in the case of Poisson excitation, and

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}} = 1 - \frac{\kappa_1^2}{\omega^2 + \kappa_1^2}, \quad \kappa \gg \kappa_1, \tag{39}$$

$$\frac{i_\omega^{(2)}}{i_{\text{shot}}^{(2)}} = 1 - \frac{1}{2}\frac{4\kappa_1^2}{\omega^2 + 4\kappa_1^2} - \frac{1}{2}\frac{\kappa^2/4}{\omega^2 + \kappa^2/4}, \quad \kappa \ll \kappa_1 \tag{40}$$

in the case of sub-Poisson excitation. These results can be understood by reasoning in the following manner. In the medium of three-level atoms excited by external radiation there are two mechanisms that control the photon flux in lasing. One is related to the properties of the medium and to the occurrence of negative feedback in the medium, and the other is related to the exciting radiation. Here the first mechanism has indeed a stabilizing effect, while the role of the second differs depends on the ratios of the parameters of the problem. For instance, if the coupling of the exciting and excited lasers is weak, the statistical properties of the exciting radiation are unimportant. They become important for regularizing the photon flux in lasing only when the coupling is strong. The above formulas show that this factor is predominant for very strong coupling, i.e., when κ is much larger not only than κ_0 but also than κ_1. In this case the role of the stabilizing factor of the medium proper becomes unimportant and the statistics of the laser simply coincides with that of the exciting field.

We see that the result of Poisson excitation is Poisson lasing, and the result of sub-Poisson excitation with complete suppression of noise at zero frequency is sub-Poisson lasing with equally complete suppression of noise at zero frequency. If the laser coupling is strong but $\kappa \ll \kappa_1$, both factors act simultaneously, and Eqs. (38) and (40) have two corresponding frequency-dependent terms. The first term is the same in (38) and (40): it is related to the stabilizing properties of the three-level medium and doubly suppresses shot noise at zero frequency. The second term (more precisely, its sign) depends on whether the exciting radiation of Poisson (i.e., destabilizing) or sub-Poisson (i.e., stabilizing). In the first case the stabilizing factor of the medium is balanced to a great extent by the stochasticity of the excitation, so that at zero frequency there is no suppression of shot noise. In the second case both factors act in unison, and shot noise at zero frequency is completely suppressed.

6. A CASCADE SCHEME CONSISTING OF *N* STRONGLY COUPLED LASERS

6.1. The basic kinetic equation for coherent fields

In the preceding sections we established that the quantum properties of the exciting laser may be transferred to the excited laser in the event of strong coupling. Two mechanisms play an important role in specifying the quantum properties of three-level atom lasing: the regularizing role of the laser medium proper, which even under Poisson excitation leads to sub-Poisson lasing, and the regularizing role of the excitation proper, if the excitation is done by sub-Poisson light. In this connection we study a chain of N three-level lasers in which, obviously, both mechanisms act.

The mathematics of this problem is not simple. The problem is that in view of the properties of three-level atomic structures, each laser in the cascade affects not only the next laser in the chain but also the preceding one. This happen only under strong-coupling conditions, but it is precisely this case that we are interested in (as the foregoing reasoning shows) if we want the statistical characteristics to improve. Hence we must, at least in principle, in a certain sense formulate the theory anew and write a kinetic equation for the density matrix of N laser fields. Fortunately, there is no need to go back to the first principles. The actual equation in the approximation of small photon fluctuations in each laser can be written on the basis of very general principles of our understanding of the physics of the problem:

$$\frac{\partial}{\partial t}\rho(\{\varepsilon_i\},t) = \sum_{i=1,2,3,\ldots,N} \Gamma_i \frac{\partial}{\partial \varepsilon_i}(\varepsilon_i - \delta_{ii+1}\varepsilon_{i+1} - \delta_{ii-1}\varepsilon_{i-1})\rho + \sum_{i=1,2,3,\ldots,N}\left[D_i \frac{\partial^2 \rho}{\partial \varepsilon_i^2} + D_{i-1\,i}\frac{\partial^2 \rho}{\partial \varepsilon_{i-1}\,\partial \varepsilon_i}\right]. \tag{41}$$

Here the coefficients $\delta_{i\,i-1}$ allow for the effect of the preceding laser on the next laser (transformation of the photon fluctuations in the lasing of the exciting laser into the lasing of the excited laser in the process of exciting the active medium), the coefficients $\delta_{i\,i+1}$ allow for the effect of the next laser on the preceding laser (transformation of the photon fluctuations in the lasing of the excited laser into the lasing

TABLE I.

	N					
	1	2	3	$\cdots$	$N-1$	N
Γ_1	$2\kappa_1$	$2(\kappa_1+\kappa_{12})$	$2(\kappa_1+\kappa_{12})$	$\cdots$	$2(\kappa_1+\kappa_{12})$	$2(\kappa_1+\kappa_{12})$
Γ_2	–	$2\kappa_2$	$2(\kappa_2+\kappa_{23})$	$\cdots$	$2(\kappa_2+\kappa_{23})$	$2(\kappa_2+\kappa_{23})$
Γ_3	–	–	$2\kappa_3$	$\cdots$	$2(\kappa_3+\kappa_{34})$	$2(\kappa_3+\kappa_{34})$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
Γ_N	–	–	–	$\cdots$	–	$2\kappa_N$
$\Gamma_1\delta_{12}$	–	κ_2	$\kappa_2+\kappa_{23}$	$\cdots$	$\kappa_2+\kappa_{23}$	$\kappa_2+\kappa_{23}$
$\Gamma_2\delta_{23}$	–	–	κ_3	$\cdots$	$\kappa_3+\kappa_{34}$	$\kappa_3+\kappa_{34}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$\Gamma_{N-1}\delta_{N-1\,N}$	–	–	–	$\cdots$	–	κ_N
$\Gamma_2\delta_{21}$	–	κ_{12}	κ_{12}	$\cdots$	κ_{12}	κ_{12}
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$\Gamma_{N-1}\delta_{N-1\,N-2}$	–	–	–	$\cdots$	$\kappa_{N-2\,N-1}$	$\kappa_{N-2\,N-1}$
$\Gamma_N\delta_{N\,N-1}$	–	–	–	$\cdots$	–	$\kappa_{N-1\,N}$

of the exciting laser via the nonlinear absorption coefficient of the exciting light). In the absence of other lasers, the coefficients Γ_i have the physical meaning of the rate of decay of photon fluctuations in the ith laser. Finally, the coefficients D_i and $D_{i\,i+1}$ specify the sources of the corresponding self-correlations and correlations between the neighboring lasers in the cascade.

Recall that earlier we dealt with important parameters in our discussions: the rate κ of absorption of the exciting light in the medium of the excited laser, and the resonant widths κ_0 and κ_1 of the exciting and excited lasers. Now we have many lasers, and the number of parameters increases accordingly. We introduce the following notation: $\kappa_{i-1\,i}$, $i=1,2,\ldots,N$, is the rate of absorption of the light of the $(i-1)$st laser in the active medium of the ith laser, and κ_i is the resonant width of the ith laser.

Let us discuss the structure of the coefficients in Eq. (41) on the basis of Eq. (21) written in the inversionless approximation. We can guess the general structure of the coefficients if we first write them for the case where $N=1$, then for $N=2$, $N=3$ and so on.

To understand the structure of the coefficients $\Gamma_i, \delta_{i-1\,i}$, and $\delta_{i\,i-1}$, we build a table. In Table I the second column consists of these coefficients for the case $N=1$, the third column consists of the coefficients for the case $N=2$, etc., and, finally, the last column consists of the coefficients for the case of an arbitrary number N of lasers. We assume that within each pair in the chain the lasers are strongly coupled, $\kappa_{i\,i+1} \gg \kappa_i$. We can then write the general expressions

$$\delta_{ii+1}=\frac{\kappa_{i+1i+2}}{2\kappa_{ii+1}}, \quad i=1,2,3,\ldots,N-1,$$
$$\delta_{N-1\,N}=\frac{\kappa_N}{2\kappa_{N-1\,N}}, \quad \delta_{N\,N+1}=0. \tag{42}$$

The coefficients δ_{ik} and δ_{ki} are linked by the following relationships:

$$4\delta_{ik}\delta_{ki}=1, \quad i,k=1,2,3,\ldots,N, \tag{43}$$

$$\Gamma_i=2\kappa_{ii+1}, \quad i=1,2,3,\ldots,N-1, \quad \Gamma_N=2\kappa_N. \tag{44}$$

Finally, the expressions for coefficients of the second derivatives can easily be written using Eq. (21):

$$D_i=-\kappa_{ii+1}n_i, \quad D_{i-1i}=\kappa_{ii+1}n_i,$$
$$i=1,2,3,\ldots,N-1, \tag{45}$$
$$D_N=-\kappa_N n_N, \quad D_{N-1N}=\kappa_N n_N.$$

6.2. Suppression of quantum noise in the cascade scheme

Now that the theory at the kinetic-equation level has been formulated, we can do specific calculations. First we use Eq. (41) to write an equation for the one-time averages $\overline{\varepsilon_i\varepsilon_k}$. The simplest way to do this is to begin by writing the equations consecutively for $N=1$, $N=2$, $N=3$, etc., and then attempt to discern a pattern. Analyzing these equations, we arrive at the conclusion that for N arbitrary all mixed averages vanish. This does not mean, however, that there are no correlations between the photon fluctuations of different lasers. Any laser in the chain reacts to the photon fluctuations in neighboring lasers only after a finite time has elapsed. The solutions for the one-time averages are

$$\overline{\varepsilon_i\varepsilon_k}=0, \quad i\neq k, \quad \overline{\varepsilon_i^2}=-n_i/2. \tag{46}$$

Here we must take into account the semiclassical relationships that exist between laser parameters:

$$n_{i-1}\kappa_{i-1i}=n_i\kappa_i, \tag{47}$$

these can also be derived on the basis of our previous discussion.

Thus, the statistical Mandel parameter of each laser in the cascade is the same: $\xi_i=-1/2$. In this sense the situation for each laser does not change when the number of lasers changes and, in particular, in comparison to the case where $N=1$, i.e., when the laser is not in a cascade but is single and is excited by external Poisson light.

We will not attempt to write the formula for the photocurrent spectrum when the light of the entire cascade is being detected. Mathematically the problem is complicated, and for

our analysis it is sufficient to know the size of the dip in the shot noise at zero frequency. We now turn to the calculation of this quantity.

On the basis of this reasoning we can immediately write the final expressions for $N=1$ and $N=2$:

$$\frac{i^{(2)}_{\omega=0}}{i^{(2)}_{\rm shot}}=\frac{1}{2},\quad N=1;\quad \frac{i^{(2)}_{\omega=0}}{i^{(2)}_{\rm shot}}=\frac{1}{3},\quad N=2. \tag{48}$$

Now we analyze the case of $N=3$. Using the basic kinetic equation (41), we can write the following system of equations:

$$\dot{\overline{\varepsilon_3\varepsilon_3(t)}}=-\Gamma_3\overline{\varepsilon_3\varepsilon_3(t)}+\Gamma_3\delta_{32}\overline{\varepsilon_3\varepsilon_2(t)}=0, \tag{49}$$

$$\dot{\overline{\varepsilon_3\varepsilon_2(t)}}=-\Gamma_2\overline{\varepsilon_3\varepsilon_2(t)}+\Gamma_2\delta_{23}\overline{\varepsilon_3\varepsilon_3(t)}+\Gamma_2\delta_{21}\overline{\varepsilon_3\varepsilon_1(t)}=0, \tag{50}$$

$$\dot{\overline{\varepsilon_3\varepsilon_1(t)}}=-\Gamma_1\overline{\varepsilon_3\varepsilon_1(t)}+\Gamma_1\delta_{12}\overline{\varepsilon_3\varepsilon_2(t)}=0. \tag{51}$$

We introduce Laplace transforms,

$$x_{ik}(\omega)=\int_0^\infty\overline{\varepsilon_i(0)\varepsilon_k(t)}\exp(i\omega t)\,dt, \tag{52}$$

and put $\omega=0$ in the formulas:

$$-\overline{\varepsilon_3^2}=-\Gamma_3x_{33}+\Gamma_3\delta_{32}x_{32}, \tag{53}$$

$$0=-x_{32}+\delta_{23}x_{33}+\delta_{21}x_{31}, \tag{54}$$

$$0=-x_{31}+\delta_{12}x_{32}. \tag{55}$$

Clearly, $x_{33}=-(3/8)n_3/\kappa_3$. Thus, for $N=3$ the noise level at zero frequency is

$$\frac{i^{(2)}_{\omega=0}}{i^{(2)}_{\rm shot}}=\frac{1}{4},\quad N=3. \tag{56}$$

Longer cascades can be analyzed in a similar manner. As a result we get

$$\frac{i^{(2)}_{\omega=0}(N)}{i^{(2)}_{\rm shot}}=\frac{1}{N+1}. \tag{57}$$

Thus, the use of a cascade scheme makes it possible to achieve the necessary level of suppression of shot noise.

The research was supported by the Russian Fund for Fundamental Research (Grant No. 98-02-18129), INTAS–RFBR(Grant No. 95-0656), and the State Committee of Higher Learning (Grant No. 95-0-5.4-66).

APPENDIX DERIVATION OF THE BASIC KINETIC EQUATION

1. Deriving the operator $\hat{S}_1$

1.1. The initial equation in the diagonal representation. Let us write the equation for the matrix density of the two-mode lasing field (the laser field and the pump field) and a single three-level atom:

$$\dot{\hat{F}}=-i[\hat{V},\hat{F}]-\hat{R}_{\rm at}\hat{F}. \tag{58}$$

The interaction Hamiltonian consists of two terms, which determine the interaction of the atom with the pump field and with the laser field:

$$\hat{V}=ig_0(a_0|3\rangle\langle1|-a_0^\dagger|1\rangle\langle3|)+ig_1(a_1|3\rangle\langle2|-a_1^\dagger|2\rangle\langle3|). \tag{59}$$

The action of the relaxation operator $\hat{R}_{\rm at}$ agrees with the relaxation processes (or with incoherent excitation processes) in Fig. 2.

A very common approach here is to fix the pump field in the form of an *c*-number field. Then instead of the photon creation and annihilation operators, $a_0^\dagger$ and a_0, in (59) we again write the *c*-numbers α_0^* and α_0, where

$$a|\alpha_0\rangle=\alpha_0|\alpha_0\rangle. \tag{60}$$

We now write Eq. (58) in the diagonal Glauber representation:

$$\dot{\hat{F}}=-i[\hat{V}',\hat{F}]-\hat{R}_{\rm at}\hat{F}+g_1\hat{D}_1\hat{F}+g_0\hat{D}_0\hat{F}. \tag{61}$$

Here the Hamiltonian $\hat{V}'$ coincides with the Hamiltonian $\hat{V}$ if in the latter we replace the photon creation and annihilation operators with the corresponding complex-valued field amplitudes. The action of the operators on derivatives with respect to the complex-valued field amplitudes and in the final analysis form the statistical properties of lasing, is determined by the equations

$$\hat{D}_1\hat{F}=\frac{\partial}{\partial\alpha_1}|2\rangle\langle3|\hat{F}+\frac{\partial}{\partial\alpha_1^*}\hat{F}|3\rangle\langle2|, \tag{62}$$

$$\hat{D}_0\hat{F}=\frac{\partial}{\partial\alpha_0}|1\rangle\langle3|\hat{F}+\frac{\partial}{\partial\alpha_0^*}\hat{F}|3\rangle\langle1|. \tag{63}$$

1.2. Factorization of the density matrix. We now factorize the density matrix $\hat{F}$:

$$\hat{F}=\hat{\sigma}\,\rho+\hat{\pi}. \tag{64}$$

Here $\hat{\sigma}(t)$ is the density matrix of a single atom in two classical external fields with amplitudes α_0 and α_1. The equation for this density matrix is

$$\dot{\hat{\sigma}}=-i[\hat{V}',\hat{\sigma}]-\hat{R}_{\rm at}\hat{\sigma}. \tag{65}$$

The matrix

$$\rho(\alpha_0,\alpha_1,t)={\rm Tr}_{\rm at}\hat{F}(\alpha_0,\alpha_1,t) \tag{66}$$

is the density matrix of the laser field and the pump field in the diagonal representation. It is for this matrix that we have set out to construct an equation. The matrix $\hat{\pi}$ allows for the correlations between the atom and the lasing and pump fields. Clearly, ${\rm Tr}_{\rm at}\hat{\pi}=0$.

Bearing all this in mind, instead of Eq. (61) we derive a system of equations for the matrices $\hat{\pi}$ and ρ with the density matrix $\hat{\sigma}$ entering as a parameter:

$$\dot{\rho}=(g_1\overline{D}_1+g_0\overline{D}_0)\rho+g_1\left(\frac{\partial}{\partial\alpha_1}\pi_{32}+\frac{\partial}{\partial\alpha_1^*}\pi_{23}\right)+g_0\left(\frac{\partial}{\partial\alpha_0}\pi_{31}+\frac{\partial}{\partial\alpha_0^*}\pi_{13}\right), \tag{67}$$

$$\dot{\hat{\pi}}=-i[\hat{V}',\hat{\pi}]-\hat{R}_{\rm at}\hat{\pi}+[g_1(\hat{D}_1-\overline{D}_1)+g_0(\hat{D}_0-\overline{D}_0)]\hat{\sigma}\rho+g_1[\hat{D}_1\hat{\pi}-\hat{\sigma}{\rm Tr}_{\rm at}(\hat{D}_1\hat{\pi})]+g_0[\hat{D}_0\hat{\pi}-\hat{\sigma}{\rm Tr}_{\rm at}(\hat{D}_0\hat{\pi})]. \tag{68}$$

Here

$$\overline{D}_1={\rm Tr}(\hat{D}_1\hat{\sigma})=\frac{\partial}{\partial\alpha_1}\sigma_{32}+\frac{\partial}{\partial\alpha_1^*}\sigma_{23},$$

$$\overline{D}_0={\rm Tr}(\hat{D}_0\hat{\sigma})=\frac{\partial}{\partial\alpha_0}\sigma_{31}+\frac{\partial}{\partial\alpha_0^*}\sigma_{13}. \tag{69}$$

1.3. Steady-state solutions for the matrix $\hat{\sigma}$. Here we use what is known as the kinetic approximation, according to which atomic variables vary much faster than the field variables, with the result that they reach a steady state for any instantaneous values of the field variables. For this to be true, we must use high-Q optical cavities for the exciting and excited lasers. This guarantees that the field variables will vary fairly slowly.

Two things can be done in the kinetic approximation we are using here. First, for the matrix elements σ_{ik} in Eqs. (67) and (68) we can take the steady-state solutions of Eq. (65). Second, we can assume that the main temporal behavior of the correlation matrix $\hat{\pi}$ is determined solely by the temporal behavior of the matrix ρ, so that we are justified in seeking the adiabatic solution of Eq. (68) by setting the time derivative of $\hat{\pi}$ to zero.

Under steady-state conditions, Eq. (65) can be written in the form of an algebraic equation for the matrix elements and can be solved fairly easily. The explicit expressions for the matrix elements are

$$\sigma_{11}=[1-U_1-U_2+V_2+V_1U_2-V_2U_1]D^{-1},$$

$$\sigma_{22}=\frac{1}{D}[-V_1+V_1U_2-V_2U_1]D^{-1},$$

$$\sigma_{33}=\frac{1}{D}[-U_1+V_1U_2-V_2U_1]D^{-1},$$

$$\sigma_{12}=-\frac{g\alpha g_0\,\alpha_0^*}{\gamma_2(\gamma_2+\gamma_3)}\frac{1}{i_1i_0}\times\left[V_2(U_1-V_1)+V_2(1-U_2+V_2)\frac{\gamma_2+\gamma_3}{\gamma_3}\right]D^{-1}, \tag{70}$$

$$\sigma_{13}=-\frac{g_0\alpha_0^*}{\gamma_2+\gamma_3}\frac{1}{i_0}\left[V_2(1-U_1-U_2+V_1+V_2)-\frac{\gamma_3}{\gamma_2}U_1(1-U_2+V_2)\right]D^{-1},$$

$$\sigma_{23}=-\frac{g\alpha}{\gamma_3}\frac{1}{i_1}\left[-V_2(1-U_1-U_2+V_1+V_2)+\frac{\gamma_3}{\gamma_2}U_2(-U_1+V_1)\right]D^{-1},$$

where we have introduced the notation

$$D=(1-U_1-2U_2)(1-V_1+V_2)+(U_2-U_1)(1+V_1+2V_2), \tag{71}$$

$$U_1=\tilde{i}_0\,\frac{i_1-(1+i_0)(\gamma_2+\gamma_3)/\gamma_3}{1+i_0+i_1},\qquad U_2=\tilde{i}_1\,\frac{i_0-(1+i_1)\gamma_3/(\gamma_2+\gamma_3)}{1+i_0+i_1}, \tag{72}$$

$$V_1=-\frac{i_0(1+i_0)(\gamma_2+\gamma_3)/\gamma_3}{1+i_0+i_1},\qquad V_2=-\frac{i_0i_1}{1+i_0+i_1}. \tag{73}$$

The dimensionless powers of the pump and laser fields can be written as follows:

$$i_0=\frac{\Omega_0^2}{\gamma_3^2},\quad i_1=\frac{\Omega_1^2}{\gamma_3(\gamma_3)},\quad \tilde{i}_0=i_0\frac{\gamma_2}{\gamma_3},\quad \tilde{i}_1=i_1\frac{\gamma_2}{\gamma_3}. \tag{74}$$

In the inversionless approximation (7),

$$U_1=\frac{i_1-1}{i_1+1}\tilde{i}_0,\quad U_2=-\frac{i_i\tilde{i}}{1+i_1},$$

$$V_1=-\frac{\tilde{i}_0}{1+i_1},\quad V_2=\frac{i_0i_1}{1+i_1}. \tag{75}$$

The steady-state solutions for the atomic density matrix in the inversionless approximation can be written in the form (4) and (5).

1.4. Adiabatic representation of the operator $\hat{\pi}$. Now, when we have the explicit expressions for the steady-state density matrix of the atom, the two equations for ρ and $\hat{\pi}$, (67) and (68), are fully determined. As mentioned earlier, we will use Eq. (68) to derive from it the explicit $\hat{\pi}=\hat{\pi}\,(\rho)$ dependence in the adiabatic approximation.

Assume that the terms on the right-hand side of Eq. (68) with derivatives with respect to the complex-valued field amplitudes are small. Now it is possible to set up iteration series in powers of these small quantities, i.e., we can write

$$\hat{\pi}=\hat{\pi}_0+\hat{\pi}_1+\hat{\pi}_2+\cdots, \tag{76}$$

where $\hat{\pi}_0$ obeys the equation

$$\dot{\hat{\pi}}_0=-i[\hat{V}',\hat{\pi}_0]-\hat{R}_{\rm at}\hat{\pi}_0. \tag{77}$$

We always have the opportunity to define the initial state of a system consisting of the atom and the field as being statis-

tically independent, i.e., we can put $\hat{\pi}=0$ at $t=0$. Then, since Eq. (77) is homogeneous, $\hat{\pi}_0$ is zero at every subsequent moment and for the first iteration $\hat{\pi}_1$ we have the equation

$$\dot{\hat{\pi}}_1=-i[\hat{V}',\hat{\pi}_1]-\hat{R}_{at}\hat{\pi}_1+g_1(\hat{D}_1-\overline{D}_1)\hat{\sigma}\rho + g_0(\hat{D}_0-\overline{D}_0)\hat{\sigma}\rho. \tag{78}$$

Every subsequent iteration has its own equation.

One must bear in mind the following. Since the iteration series is set up as a series in powers of derivatives with respect to the complex-valued field amplitudes, the truncation of the series reduces to the problem of establishing how important the higher-order derivatives are in the formation of the statistical pattern of lasing. It is well known that if we are dealing with fields that can be called classical, then we can be sure that it is enough to leave only the first iteration, which corresponds to a basic kinetic equation that contains derivatives with respect to the complex-valued field amplitudes of order no higher than the second. On the other hand, there is no reason why for nonclassical fields of the sub-Poisson or squeezed types we must limit ourselves to the lowest-order derivatives. Actually, we must keep all the derivatives, i.e., allow for the compete iteration series (76).

We write Eq. (78) in terms of the matrix elements and set the derivatives equal to zero:

$$\begin{aligned}
&\dot{\pi}_{13}=-\gamma_3\pi_{13}-g_0\alpha_0^*(\pi_{33}-\pi_{11})+g_1\alpha_1^*\pi_{12}+S_{13}=0,\\
&\dot{\pi}_{12}=-\gamma_2\pi_{12}-g_0\alpha_0^*\pi_{32}-g_1\alpha_1\pi_{13}+S_{12}=0,\\
&\dot{\pi}_{23}=-(\gamma_2+\gamma_3)\pi_{23}-g_1\alpha_1^*(\pi_{33}-\pi_{22})\\
&\qquad +g_0\alpha_0\pi_{21}+S_{23}=0,\\
&\dot{\pi}_{22}=2\gamma_2\pi_{22}-g_0\alpha_0(\pi_{31}+\pi_{13})+S_{11}=0,\\
&\dot{\pi}_{33}=-2\gamma_3\pi_{33}+g_0\alpha_0(\pi_{13}+\pi_{31})\\
&\qquad +g_1(\alpha_1\pi_{23}+\alpha_1^*\pi_{32})+S_{33}=0.
\end{aligned} \tag{79}$$

This must be combined with the condition

$$\pi_{11}+\pi_{22}+\pi_{33}=0. \tag{80}$$

The inhomogeneous terms have the form

$$\begin{aligned}
S_{12}&=g_1\left(-\overline{D}_1\sigma_{12}+\frac{\partial}{\partial\alpha_1^*}\sigma_{13}\right)\\
&\quad+g_0\left(-\overline{D}_0\sigma_{12}+\frac{\partial}{\partial\alpha_0}\sigma_{32}\right),\\
S_{13}&=-g_1\overline{D}_1\sigma_{13}+g_0\left(-\overline{D}_0\sigma_{13}+\frac{\partial}{\partial\alpha_0}\sigma_{33}\right),\\
S_{23}&=-g_1\overline{D}_1\sigma_{23}+\frac{\partial}{\partial\alpha_1}\sigma_{33}-g_0\overline{D}_0\sigma_{23},\\
S_{11}&=-g_1\overline{D}_1\sigma_{11}+g_0\overline{D}_0(1-\sigma_{11}),\\
S_{22}&=g_1\overline{D}_1(1-\sigma_{22})-g_0\overline{D}_0\sigma_{22},\\
S_{33}&=-g_1\overline{D}_1\sigma_{33}-g_0\overline{D}_0\sigma_{33}.
\end{aligned} \tag{81}$$

Solving the system of equations (79), we obtain

$$\begin{aligned}
\pi_{11}&=-[U_0(1+V_1+2V_2)+V_0(1-U_1-2U_2)]D^{-1},\\
\pi_{22}&=[U_0(2V_1+V_2)+V_0(1-2U_1-U_2)]D^{-1},\\
\pi_{33}&=[U_0(1-V_1+V_2)-V_0(U_2-U_1)]D^{-1},\\
\pi_{13}&=\lambda_{13}+\frac{g_0\alpha_0}{\gamma_2+\gamma_3}\frac{1}{i_0}\Big[-U_0V_2(1+3V_1+3V_2)\\
&\quad+\frac{\gamma_3}{\gamma_2}U_0U_1(2+3V_2)-V_0V_2(2-3U_1-3U_2)\\
&\quad+\frac{\gamma_3}{\gamma_2}V_0U_1(1-3U_2)\Big]D^{-1},\\
\pi_{23}&=\lambda_{23}+\frac{g\alpha^*}{\gamma_3}\frac{1}{i}\Big[U_0V_2(1+3V_1+3V_2)\\
&\quad+\frac{\gamma_3}{\gamma_2}U_0U_2(1-3V_1)+V_0V_2(2-3U_1-3U_2)\\
&\quad-\frac{\gamma_3}{\gamma_2}V_0U_2(1-3U_1)\Big]D^{-1}.
\end{aligned} \tag{82}$$

Here U_1, U_2, V_1, and V_2 are the same coefficients as before, i.e., they are given by Eqs. (72) and (73). The new coefficients λ_{ik}, U_0, and V_0 are related to the inhomogeneous terms S_{ik} in the initial system of equations and can be expressed in terms of the S_{ik} as follows:

$$\begin{aligned}
\lambda_{12}&=\frac{1}{\gamma_2}\frac{1}{1+i_0+i_1}\\
&\quad\times\left(S_{12}-\frac{g_1\alpha_1}{\gamma_3}S_{13}-\frac{g_0\alpha_0}{\gamma_2+\gamma_3}S_{32}\right),\\
\lambda_{13}&=\frac{g_1\alpha_1^*}{\gamma_3}\lambda_{12}+\frac{1}{\gamma_3}S_{13},\\
\lambda_{23}&=\frac{g_0\alpha_0}{\gamma_2+\gamma_3}\lambda_{21}+\frac{1}{\gamma_2+\gamma_3}S_{23},
\end{aligned} \tag{83}$$

and

$$\begin{aligned}
U_0&=\frac{1}{2\gamma_3g_1}S_{33}+\frac{g_1}{2\gamma_3}(\alpha_1\lambda_{23}+\alpha_1^*\lambda_{32})\\
&\quad+\frac{g_0}{2\gamma_3}(\alpha_0\lambda_{13}+\alpha_0^*\lambda_{31}),\\
V_0&=-\frac{1}{2\gamma_2g_1}S_{11}+\frac{g_0}{2\gamma_2}(\alpha_0\lambda_{13}+\alpha_0^*\lambda_{31}).
\end{aligned} \tag{84}$$

Comparing Eqs. (67) and (15), we can write the desired operator as follows:

$$\hat{S}_1=\frac{2g_1N_1}{\kappa_1}\left(\frac{\partial}{\partial\alpha_1}\pi_{32}+\frac{\partial}{\partial\alpha_1^*}\pi_{23}\right)+\frac{2g_0N_1}{\kappa_1}\left(\frac{\partial}{\partial\alpha_0}\pi_{31}+\frac{\partial}{\partial\alpha_0}\pi_{13}\right). \tag{85}$$

In the inversionless approximation we have

$$\hat{S}_1=\kappa_1\frac{C}{4n_0}\,\frac{\gamma_3/\gamma_2-3\tilde{i}_1}{(1+i_1)^2(1+\tilde{i}_1)^2}\left(\frac{\partial}{\partial\alpha_1}\,\alpha_1+\frac{\partial}{\partial\alpha_1^*}\,\alpha_1^*\right)^2 + \kappa_1 C\,\frac{\gamma_3/\gamma_2+2i_1}{(1+i_1)^2(1+\tilde{i}_1)}\,\frac{\partial^2}{\partial\alpha_1\partial\alpha_1^*}, \tag{86}$$

where

$$C=\frac{2g_1^2\tilde{i}_0}{\gamma_3\kappa_1} \tag{87}$$

is what is known as the cooperative parameter (the ratio of the linear gain of the medium to the linear coefficient of losses in the cavity).

2. Explicit form of the operator $\hat{S}_0$

Here we follow the ideas of Refs. 2 and 5, according to which

$$\hat{S}_0=\frac{\eta_2}{2}\left[2a_0^\dagger\underset{\rightarrow}{a_0}\underset{\leftarrow}{}-a_0\underset{\rightarrow}{a_0^\dagger}-a_0\underset{\leftarrow}{a_0^\dagger}-\frac{\eta_1}{2}(a_0\underset{\rightarrow}{a_0^\dagger}-a_0\underset{\leftarrow}{a_0^\dagger})^2\right] \times\left[1+\frac{\eta_+}{2}(a_0\underset{\rightarrow}{a_0^\dagger}+a_0\underset{\leftarrow}{a_0^\dagger})+\frac{\eta_1\eta_2}{4}(a_0\underset{\rightarrow}{a_0^\dagger}-a_0\underset{\leftarrow}{a_0^\dagger})^2\right]^{-1}. \tag{88}$$

The arrows under the operators indicate the direction in which the operator acts on the expressions to their right, including the density matrix. The explicit expressions for the saturation parameters η_1 and η_2 ($\eta_+=\eta_1+\eta_2$) are unimportant.

After passing to the diagonal representation we obtain

$$\hat{S}_0=-\frac{\eta_2}{2}\left[\left(1+\frac{\eta_1}{2}\right)\left(\frac{\partial}{\partial\alpha_0}\,\alpha_0-\frac{\partial}{\partial\alpha_0^*}\,\alpha_0^*\right)^2+\frac{\eta_1}{2}\left(1+\frac{3\gamma_1}{2\gamma_2}\right) \times\left(\frac{\partial}{\partial\alpha_0}\,\alpha_0+\frac{\partial}{\partial\alpha_0^*}\,\alpha_0^*\right)^2\frac{1}{1+\eta_+|\alpha_0|^2}\right]\frac{1}{1+\eta_+|\alpha_0|^2} +\{\cdots\}. \tag{89}$$

The terms $\{\cdots\}$ contain the higher-order derivatives with respect to the complex-valued amplitudes. In terms of the polar coordinates u_0 and φ_0 we have

$$\hat{S}_0=-\frac{\partial}{\partial u_0}\,\frac{\eta_2 u_0}{1+\eta_+ u_0}+\frac{\partial^2}{\partial u_0^2}\,\frac{\eta_2 u_0}{(1+\eta_+ u_0)^2} + \frac{\eta_2}{4u_0}\,\frac{1+\eta_1 u_0}{1+\eta_+ u_0}\,\frac{\partial^2}{\partial\varphi_0^2}. \tag{90}$$

*)E-mail: yuri_golubev@pop.convey.ru

[1] H. Ritsch, M. A. M. Marte, and P. Zoller, Europhys. Lett. **19**, 7 (1992).
[2] Yu. M. Golubev and I. V. Sokolov, Zh. Eksp. Teor. Fiz. **87**, 408 (1984) [Sov. Phys. JETP **60**, 234 (1984)].
[3] O. A. Kocharovskaya and Ya. I. Khanin, JETP Lett. **48**, 630 (1988); O. A. Kocharovskaya, P. Mandel, and Ya. I. Khanin, Izv. Akad. Nauk SSSR, Ser. Fiz. **54**, 1979 (1990); O. A. Kocharovskaya, P. Mandel, and M. O. Scully, Phys. Rev. Lett. **74**, 2451 (1995).
[4] M. O. Scully and W. E. Lamb, Jr., Phys. Rev. A **159**, 208 (1967).
[5] Yu. M. Golubev, Zh. Eksp. Teor. Fiz. **103**, 832 (1993) [JETP **76**, 408 (1993)].

Translated by Eugene Yankovsky

Indirect (x-ray) irradiation of encapsulated microtargets in the Iskra-5 facility

F. M. Abzaev, S. A. Bel'kov,*) A. V. Bessarab, S. V. Bondarenko, V. A. Gaĭdash, S. G. Garanin, G. V. Dolgoleva, N. V. Zhidkov, V. M. Izgorodin, G. A. Kirillov, G. G. Kochemasov,†) D. N. Litvin, S. P. Martynenko, V. M. Murugov, L. S. Mkhitar'yan, A. V. Pinegin, S. I. Petrov, A. V. Senik, and N. A. Suslov

Russian Federal Nuclear Center, All-Russian Scientific-Research Institute of Experimental Physics, 607190 Sarov, Russia

V. S. Bushuev, V. M. Dorogotovtsev, and Yu. A. Merkul'ev

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 6 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 1993–2003 (December 1998)

Experiments on the indirect (x-ray) irradiation of high-aspect-ratio capsules (with a diameter-to-thickness ratio ≈ 900) filled with DT gas are performed on the Iskra-5 laser facility. It is shown that all the characteristics measured (neutron yield, ion temperature, shell implosion time, etc.) are faithfully reproduced in calculations based on the one-dimensional SNDA (spectral nonequilibrium diffusion of absorption) program for nonequilibrium radiation gas dynamics. The calculations provide an explanation for the experimentally detected generation of a smaller number of neutrons in an experiment with a higher measured value for the ion temperature of DT gas. © *1998 American Institute of Physics.* [S1063-7761(98)00612-X]

1. INTRODUCTION

As the calculations and experiments described in Ref. 1 show, the system for the indirect irradiation of targets with a spherical hohlraum employed in experiments in the Iskra-5 facility permits the achievement of nearly spherically symmetric compression of the DT fuel within glass microshells. It follows from the calculations that the characteristic value of the irradiation nonuniformity with consideration of the angular spread of the laser energy amounts to about 3% when $D/d \simeq 7$, where D and d are the diameter of the hohlraum and the shell, respectively.

However, experiments with shells of relatively large diameter ($D/d \approx 2-3$) would be of interest. Although such shells would be partially illuminated by laser radiation and the degree of symmetrization of x radiation with a comparatively small gap between the hohlraum and the shell would be appreciably smaller than in the case of $D/d \approx 7$, we can attempt to achieve nearly spherically symmetric compression of the DT fuel by employing a shell thickness that is appreciably smaller than the thickness which is heated during an x-ray pulse. As the experiments in Ref. 2 showed, the heated thickness of glass in a hohlraum of diameter $D=2$ mm is $\Delta_{\text{defl}} \simeq 5-7\,\mu$m.

Experiments with thin shells of millimeter diameter are also interesting from the standpoint of developing new methods and improving the accuracy of tools for plasma diagnostics. For example, an increase in the shell implosion time permits improvement of the relative accuracy of the determination of the time of neutron generation, which is important for verifying several fine points in programs for calculating the operation of targets. Two shells of diameter 0.8–0.9 mm with a wall thickness of $\simeq 1\,\mu$m were prepared in the P. N. Lebedev Physics Institute of the Russian Academy of Sciences for performing experiments on the Iskra-5 facility.[3]

This paper presents the results of two experiments with these shells, which were performed according to the indirect drive scheme on the Iskra-5 facility.[3] In one of them the diameter of the gold hohlraum was $D=2$ mm, and in the other it was $D=4$ mm. The characteristic irradiation intensity of the inner surface of the hohlraum by laser rays differs in these experiments by a factor roughly equal to 4. According to the presently prevailing ideas, the inverse bremsstrahlung absorption of laser radiation is the dominant mechanism in the 4-mm hohlraum. In the 2-mm hohlraum the intensity acting on the hohlraum surface is $I \simeq (5-7)\times 10^{14}$ W/cm^2. Therefore, the parameter $I\lambda^2$, which characterizes the nonlinearity of the interaction of the laser radiation with the plasma, is fairly high ($I\lambda^2 \simeq 10^{15}$ W/cm$^2\cdot\mu$m^2). As a result, an appreciable portion of the energy from the laser corona is carried off by fast ions having a velocity of $\simeq 5\times 10^8$ cm/s. As the experiments with a magnesium hohlraum coating showed,[4] these ions can impart to the shell an energy sufficient for the generation of about 10^9 neutrons during a single pulse. Therefore, variation of the hohlraum diameter should make it possible to shed additional light on the role of the fast ions.

This paper also includes a theoretical analysis of the experimental results and provides recommendations for developing this area of research.

2. EXPERIMENTAL SETUP, MEASUREMENT METHODS AND PRINCIPAL RESULTS

2.1. Parameters of the laser radiation

The total energy of the facility at the accelerator exit was $E_\Sigma \approx 9700$ J ($\langle E_{\text{chan}}\rangle = 810$ J) in the first experiment and

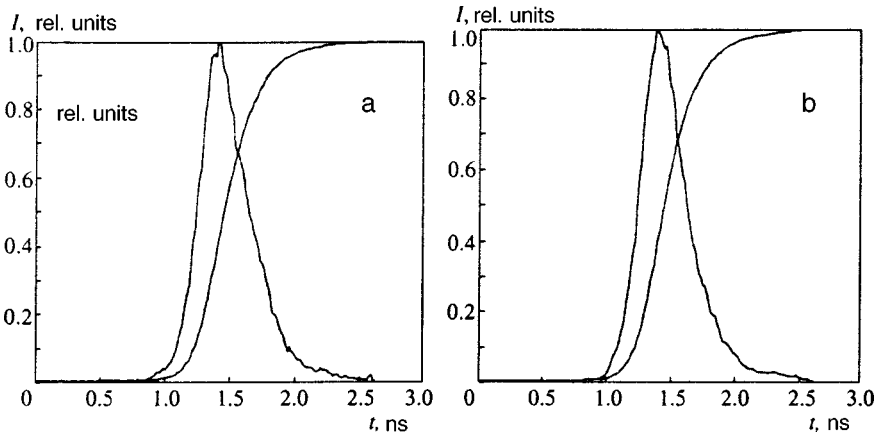

FIG. 1. Form of the laser radiation pulse and integral pulse: a — experiment No. 1, $\tau_{0.5}=0.38$ ns; b — experiment No. 2, $\tau_{0.5}=0.37$ ns.

$E_\Sigma \approx 10\,060$ J ($\langle E_{\text{chan}}\rangle = 840$ J) in the second experiment. The recorded form of the individual pulses is nearly Gaussian.

The divergence of the radiation in the channels is close to the nominal value $\theta_{0.8E} < 10^{-4}$ rad. The contrast of the radiation in all the channels was fairly high: $k_E \geqslant 10^6$, $k_P \geqslant 10^6$. As the experience gained during many years of research indicates, at such a contrast level there is no self-excitation of the accelerators and plasma does not form on the inner surface of the target before the arrival of the main pulse. The accuracy of the transverse aiming at the target is equal to ± 30 μm, and the accuracy of the longitudinal matching of the foci of the alignment and high-power radiation is ± 100 μm. The diameter of the narrow portions of the laser beams is ~ 100 μm.

The form of the total radiation pulse in all the channels obtained by integrating the individual pulses with consideration of the channel energy and the asynchronous arrival of the pulses at the target is shown in Fig. 1. The figure also shows an energy buildup curve obtained by integrating the total laser pulse.

2.2. Features of the target design

The target consisted of a thin-walled spherical copper hohlraum, whose inner surface is coated by a layer of gold with a thickness of ~ 1 μm and whose outer surface is coated by bismuth to a thickness of 0.1–0.3 μm. The diameter of the laser entrance holes was 0.6 mm in the hohlraum with $D=2$ mm and 0.7 mm in the hohlraum with $D=4$ mm. At the center there was a glass microsphere with a diameter of 0.8–0.9 mm and a wall thickness of about 1 μm, which was filled with a gaseous DT mixture to a pressure $P_{\text{DT}} \approx 3.5$ atm. The microspheres were made at the P. N. Lebedev Physics Institute of the Russian Academy of Sciences, and the hohlraums were made in the All-Russian Scientific-Research Institute of Experimental Physics.

2.3. Diagnostics of the parameters of the action of laser radiation on a target

The complex diagnostics instrumentation of the Iskra-5 facility was described in detail in Ref. 5. The main results of the experiments, which are supplemented by features of the measurement methods where necessary, are described below. The main results of the experiments are listed in Table I.

The target images in ordinary x radiation were recorded by a set of pinhole cameras: survey cameras for observing the state of all six laser entrance holes and a pinhole camera for observing the state of the central microtarget. Figure 2 presents the x-ray images of the compressed region of the central capsule observed through a laser entrance hole. It is seen that in both experiments the x-ray luminescence of the compressed region has an annular form with an intensity dip at the center. The diameter of the ring at the luminescence intensity maximum is 250–260 μm in experiment No. 1 and 225–235 μm in experiment No. 2. In experiment No. 1 the distribution of the intensity along the ring perimeter is markedly inhomogeneous. In experiment No. 2 the ring has an appreciably smaller width, and the distribution of the intensity along the perimeter is more homogeneous.

TABLE I. Main results of the experiments.

Experiment No.	D_{hohl}/D_h, mm	D_{sh}, μm	ΔR_{sh}, μm	P_{DT}, atm	E_L, J	$\tau_{\gamma\gamma}$, ns	$\tau_{\gamma n}$, ns	N, 10^9	T_{DT}, keV
1	2/0.6	811	0.9	3.5	7300	1 ± 0.15	0.83 ± 0.05	5.5	1
2	4/0.7	907	1.1	3.5	7500	1.45 ± 0.05	1.10 ± 0.05	0.6	3.2

[a]Notation: D_{hohl}/D_h — diameters of the hohlraum and the laser entrance holes; D_{sh}, ΔR_{sh} — diameter and thickness of the glass microsphere; P_{DT} — pressure of DT gas; E_L — laser radiation energy introduced into the interaction chamber; $\tau_{\gamma\gamma}$ — delay of the onset of the generation of x radiation from the compressed core relative to the onset of the generation of x radiation on the hohlraum wall; $\tau_{\gamma n}$ — delay of the onset of the generation of neutrons relative to the onset of the generation of x radiation on the hohlraum wall; N — integral neutron yield; T_{DT} — temperature of the DT fuel determined by the time-of-flight method.

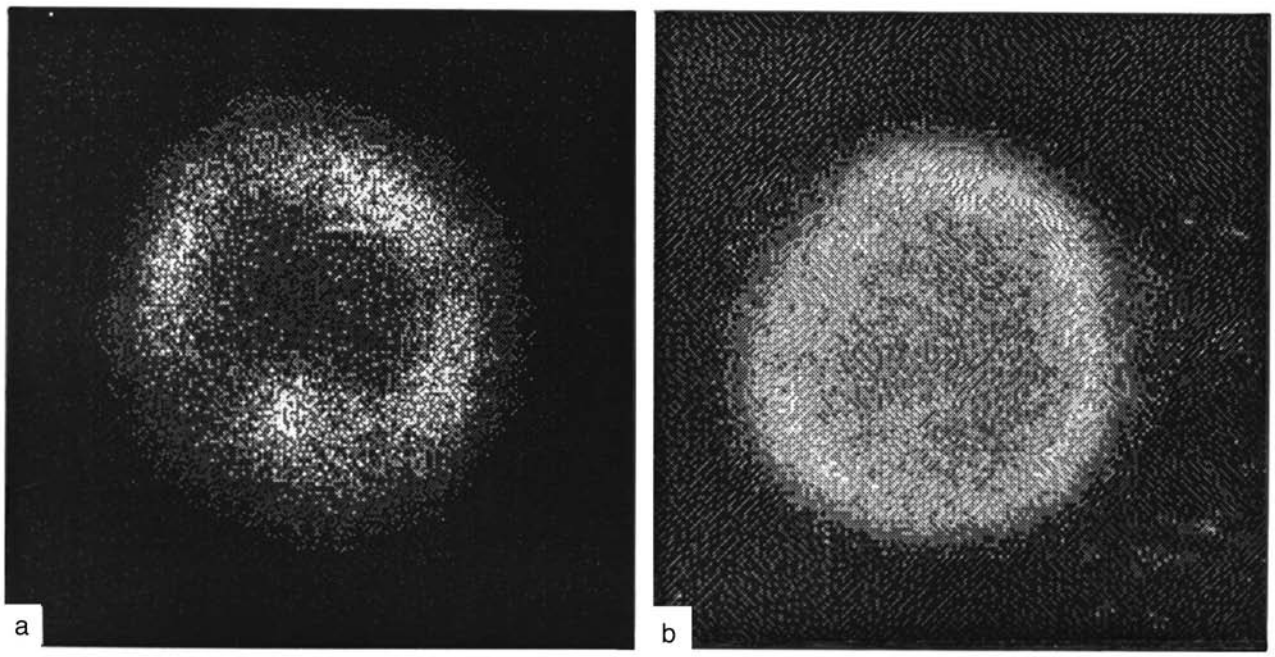

FIG. 2. X-ray images of the compressed region: a — experiment No. 1, filter consisting of 10 μm of PPK+5μm of Ti; b — experiment No. 2, filter consisting of 10 μm of $C_6H_8Cl_4$.

The implosion time of the shells was determined from a time scan of the x-ray image of the region where the central capsule is located in the photon energy range $h\nu \sim 4$ keV on an x-ray chronograph.[6] Figure 3 presents the results of this procedure in the present experiments. The time scan of the capsule luminescence clearly displays two maxima. One corresponds to the x-ray luminescence of the hohlraum, and the other maximum appears as a result of compressing and heating the fuel. The time interval between these pulses characterizes the compression time $\tau_{\gamma\gamma}$, which is listed in Table I. The values of $\tau_{\gamma\gamma}$ were obtained by linearly approximating the leading edges of the pulses to the point of intersection with the time axis and calculating the corresponding delay from these points. The significantly discontinuous form of the leading edge of the pulse accompanying compression of the fuel in the first experiment leads to an appreciable increase in the measurement error for $\tau_{\gamma\gamma}$. In the second experiment the value of $\tau_{\gamma\gamma}$ is somewhat larger, in agreement with the decrease in the energy supplied to the capsule as a result of the increase in the hohlraum diameter.

The integral neutron yield was measured using the battery of methods described in Ref. 5. The measured values of the neutron yield in the experiments performed are also listed in Table I.

Figure 4 presents the result of recording a neutron pulse (time-of-flight method) using a detector positioned at a distance of 16.7 m. The figure also presents calculated signals for an assigned ion temperature. It is seen that the ion tem-

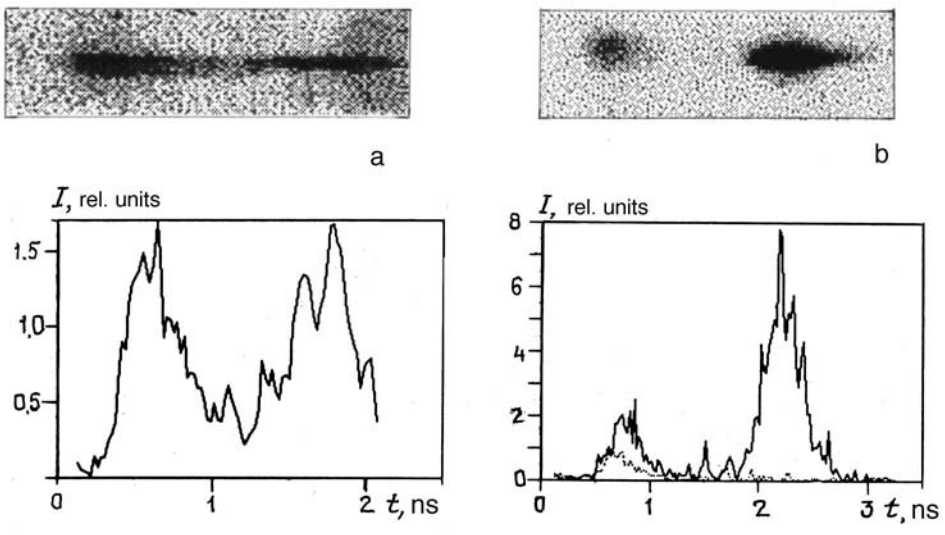

FIG. 3. Results of treatment of chronograms of the x-ray luminescence of the central capsule: a — experiment No. 1; b — experiment No. 2.

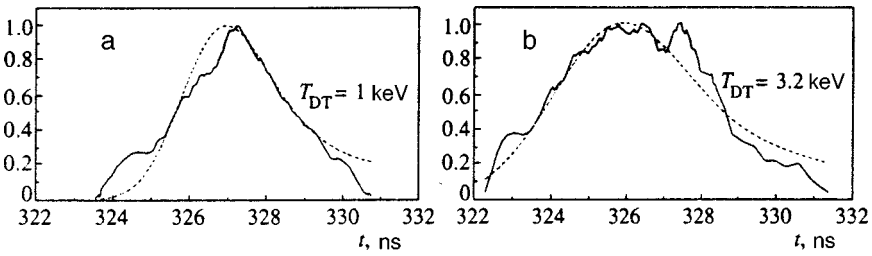

FIG. 4. Numerical plots of oscillograms obtained using the time-of-flight method at a distance of 16.7 m for experiments No.1 (a) and No. 2 (b) (solid curves) and theoretical signal calculated with allowance for the transient function of the detector (dashed curves) for the optimal ion temperature of DT gas.

perature is higher in experiment No. 2 with a lower neutron yield.

The delay between the onset of the generation of x radiation and the time of the production of thermonuclear neutrons $\tau_{\gamma n}$ was determined using a neutron-radiation detector based on an SPPD11-02 p–i–n silicon semiconductor diode.[7] A polyethylene converter was mounted on the front surface of the diode to improve the neutron detection efficiency. The detector was placed in lead shielding and positioned at a distance of about 15 cm from the target. If the detector is positioned closer to the target, the rise time of the transient characteristic of the SPPD11-02 diode does not permit direct measurement of the form of the neutron pulse or its separation from the x-ray pulse. Therefore, the lead shielding of the detector was selected so that any hard x radiation would be essentially completely suppressed, and the output signal of a vacuum-tube x-ray diode which detects x radiation with a photon energy in the vicinity of $h\nu \sim 1.5$ keV served as the time mark. The signals from the vacuum-tube x-ray diode and the p–i–n diode were recorded in one ray of an SRG-7 oscillograph. The value of $\tau_{\gamma n}$ was determined by subtracting the time delay of the recording scheme t_{scheme} and the difference between the times of passage of photons and neutrons over the distance from the target to the detector t_R from the time interval $t_{\gamma n}$ between the x-ray and neutron pulses determined from the oscillogram:

$$\tau_{\gamma n} = t_{\gamma n} - t_{\text{scheme}} - t_R .$$

Figure 5 shows the oscillogram recorded in experiment No. 1. As can be seen in Fig. 5, the semiconductor diode detects not only neutron radiation, but also hard x radiation, which partially passed through the detector shielding. In the second experiment the yield of x radiation was apparently reduced so much that it was below the detection threshold and was not observed on the oscillogram. The amplitude of the neutron pulse was diminished by a factor of 10 in experiment No. 2. The recorded values of the delay time of the generation of neutron radiation relative to x radiation were $\tau_{\gamma n} = 0.83 \pm 0.05$ ns in experiment No. 1 and $\tau_{\gamma n} = 1.10 \pm 0.05$ ns in experiment No. 2.

3. DISCUSSION AND THEORETICAL ANALYSIS OF THE EXPERIMENTAL RESULTS

3.1. X-irradiation uniformity

Let us begin the discussion with estimates of the nonuniformity of the x irradiation of the central capsule containing DT gas. The absorption of the laser radiation was calculated by the Monte Carlo method. The calculation allowed for the real three-dimensional geometry of the laser beams and the laser entrance holes, as well as the absorption and repeated reflection of laser radiation within the hohlraum cavity of the target. It was assumed that the laser radiation is absorbed according to an inverse bremsstrahlung absorption mechanism with an absorption coefficient $k_n = k_0 \cos^3\gamma$ (γ is the angle of incidence of the radiation onto the surface) and $k_0 = 0.5$. The calculations were performed under the assumption of constancy of the shape of the reflecting surface during the action of the laser pulse and for the idealized case of the absence of any imbalance of the energy of the laser radiation in different channels.

Because of the large relative (to the cavity diameter) diameter in experiment No. 1, the surface of the capsule with DT gas was partially in the propagation path of the laser radiation. The distribution of the laser absorption on the capsule surface obtained in the calculation is shown in Fig. 6a.

As can be seen in the figure presented, the bulk of the laser absorption on the capsule surface occurs in regions where the laser beams were tangent to the capsule surface. The percentage of absorption of the laser radiation by the capsule surface in this experiment was equal to 2.9% of the laser energy introduced into the target cavity (the comparatively low percentage of the radiated energy absorbed by the capsule is attributed to the large characteristic values of the angle of incidence γ of the radiation onto the capsule surface and, therefore, to the small values of the inverse bremsstrahlung absorption coefficient in the calculations).

Estimates of the nonuniformity of the x irradiation of the central capsule with DT gas were performed for the geometry of the introduction of laser radiation into the hohlraum corresponding to the experiment. The flux density of the x radiation onto the surface of the hohlraum cavity was deter-

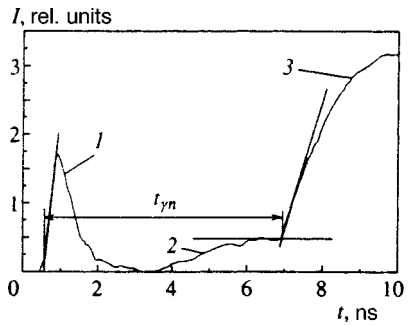

FIG. 5. Oscillogram of the output of the device for recording the time of the generation of neutron radiation (experiment No. 1): *1* — reference; *2* — pulse of hard x radiation, *3* — neutron pulse.

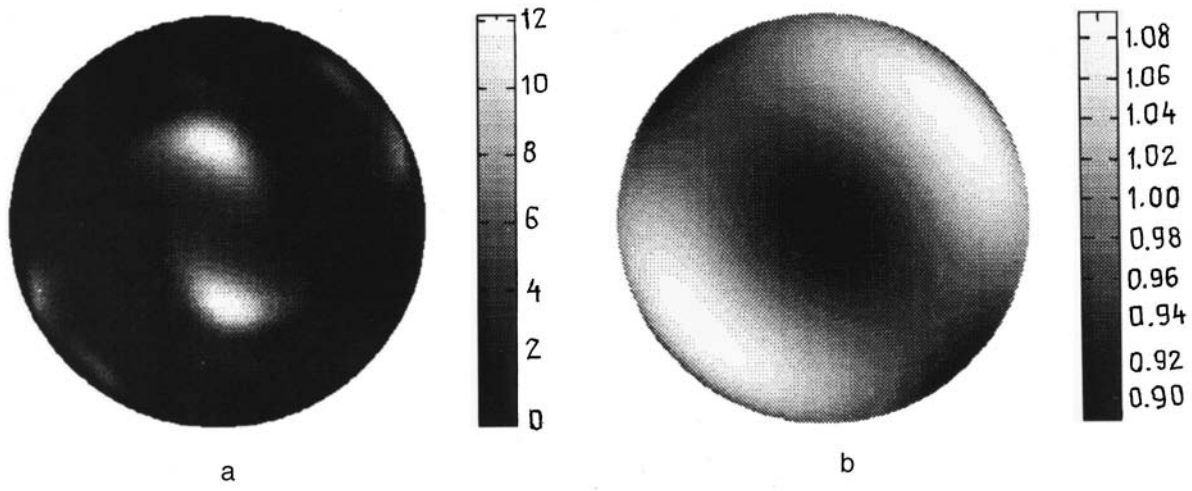

FIG. 6. Normalized distributions of laser absorption (a) and x-ray radiation field (b) on the surface of the central capsule.

mined from the distribution of the absorbed laser energy on the inner surface of the hohlraum on the basis of the energy balance relations and was written in the form

$$S_r(\theta,\varphi)=A_1Q_L(\theta,\varphi)+A_2\overline{Q}_L,$$

where $Q_L(\theta,\varphi)$ is the absorbed laser radiation intensity at a given point on the inner surface of the target cavity, $\overline{Q}_L$ is the mean absorbed laser radiation intensity, and the coefficients A_1 and A_2 are expressed in terms of the albedo α of the x radiation, the conversion coefficient of the laser radiation into primary x radiation η, and the relative area of the laser entrance holes β and equal

$$A_1=0.5(1+\alpha)\,\eta,\quad A_2=\alpha\,\eta(1-\beta)\left[0.5+\frac{\alpha(1-0.5\beta)}{1-\alpha(1-\beta)}\right].$$

The influence of the central capsule on the formation of x radiation on the hohlraum walls was neglected. The degree of this influence can be estimated by the characteristic value of the solid angle within which the central capsule can be seen from the hohlraum surface, i.e. by a quantity $\sim(D_{\rm sh}/D_{\rm hohl})^2$. The results of the calculation of the nonuniformity of the field of x radiation on the surface of the central capsule are presented in Fig. 6b.

The values obtained for the maximum ($\varepsilon_{\max}=|I_{\max}-I_{\min}|/2\overline{I}$) and rms nonuniformity (as well as the losses of laser radiation energy in the holes due to repeated internal reflection) are listed in Table II. This table also presents the amplitudes $\tilde{\gamma}_l=\sqrt{\overline{\alpha}_{lm}\tilde{\alpha}^*_{lm}}$ of the harmonics with a given orbital angular momentum l in the expansion of the normalized distribution $\tilde{I}_x=I/||I||$ of the x radiation impinging on the capsule surface.

3.2. Calculations of target operation

The theoretical analysis was performed using the one-dimensional SNDA (spectral nonequilibrium diffusion of absorption) program for nonequilibrium radiation gas dynamics.[8] The parameters of the targets and the laser pulse assigned in the calculations corresponded to the experimental values.

The following physical processes were taken into account in the calculations: the inverse bremsstrahlung absorption of laser radiation on the inner surface of the hohlraum; the repeated reflection and escape of laser radiation through the entrance holes; the nonequilibrium, nonstationary kinetics of the ionization of the high-Z plasma in the mean-ion approximation;[9] the generation, absorption, and transmission of x radiation (with allowance for the escape of part of it through the laser entrance holes) in the spectral diffusion approximation; electron and ion thermal conductivity; electron-ion relaxation; and neutron generation. All the calculations were performed without allowance for the influence of ''fast'' ions accelerated in the laser corona on the compression dynamics of the central target, i.e., it was as-

TABLE II. Nonuniformity of the x irradiation of the central capsule.

Experiment No.	E_h, %	$E_{\rm sh}$, %	$\varepsilon_{\max}$, %	$\varepsilon_{\rm rms}$, %	$\tilde{\gamma}_1$	$\tilde{\gamma}_2$	$\tilde{\gamma}_3$	$\tilde{\gamma}_4$	$\tilde{\gamma}_5$
1	41	2.9	10	5.0	1.5×10^{-4}	1.2×10^{-2}	3.7×10^{-2}	3.1×10^{-2}	3.4×10^{-3}
2	21	0.028	5.2	2.9	1.4×10^{-3}	1.7×10^{-2}	2.3×10^{-2}	3.6×10^{-3}	1.5×10^{-3}

[a]*Notation:* E_h, $E_{\rm sh}$ — fractions of the laser radiation energy escaping through the entrance holes and absorbed by the shell (as percentages of the radiation introduced into the target); $\varepsilon_{\max}$ — maximum nonuniformity; $\varepsilon_{\rm rms}$ — rms nonuniformity; $\tilde{\gamma}_l$ — amplitudes of the expansion of the normalized distribution of the x radiation field on the surface of the central capsule in spherical harmonics.

TABLE III. Results of the calculations of target operation.

Experiment No.	E_{abs}, kJ	k_{abs}	T_γ, eV	E_γ, kJ	t_N, ns	δ_N	δ_{max}	N, 10^9	T_{DT}, keV
1	4.62	0.63	195	1.79	0.89(0.85)	30	40	3.8(5.0)	1.5(1.0)
2	5.86	0.78	144	0.74	1.05(1.1)	14	26	0.58(0.50)	2.5(3.0)

[a]*Notation:* E_{abs} — laser radiation energy absorbed within the hohlraum, k_{abs} — absorption coefficient, T_γ — maximum effective temperature of the radiation within the hohlraum, E_γ — energy of the x radiation escaping through the laser entrance holes, t_N — time of the generation of the neutron pulse relative to the maximum of the laser pulse, δ_N — compression ratio of DT gas at the time of generation of the neutron pulse, δ_{max} — maximum compression ratio. The experimental values of t_N, N, and T_{DT} are also presented in the table in parentheses.

sumed that compression of the glass microtarget was effected under the action of x radiation generated by the hohlraum walls and the gold plasma moving toward the center of the hohlraum.

The main results of the calculations are presented in Table III. Figure 7 shows R–t diagrams of the motion of the boundary between the gas and glass, as well as the laser and neutron pulses obtained in the calculations. It is noteworthy that neutron generation takes place long before the achievement of maximum compression of DT gas in experiment No. 2. Figure 8 shows the calculated distributions of the ion temperature of the gas at moments in time corresponding to the generation of half of the neutron yield. It can be seen that the gas temperature is essentially uniform throughout the volume in the calculation of experiment No. 1, while the distribution for experiment No. 2 is extremely inhomogeneous. An analysis of the results of the calculation showed that in the case of experiment No. 1 a similar temperature distribution appears at a moment in time close to the moment of focusing of the first shock wave at the capsule center. In this case about 10^8 neutrons are generated. However, the gas then undergoes final compression, during which the temperature at the center decreases. Furthermore, since neutron generation takes place essentially throughout the volume, the total number of neutrons increases by more than 10 fold. It is significant that the gold plasma, which prevents dismemberment of the compressed capsule and increases the retention time of hot DT gas, participates in the final compression. In the calculation for experiment No. 2 the gold plasma does not manage to reach the capsule, the temperature of the gas drops rapidly, and essentially no additional neutron generation occurs at the time of final compression. As a whole, satisfactory agreement is observed between the calculated and experimental results with respect to the absolute neutron yield, its generation time, and the gas temperature.

It can be concluded on the basis of the calculated data presented that, despite the decrease in the neutron yield by an order of magnitude in experiment No. 2 in comparison to experiment No. 1 (see Table I), the gas temperature determined from the neutron spectrum recorded by the time-of-flight method can be significantly higher. This is because neutron generation occurs in a small part of the DT fuel in the second experiment. Thus, the ''contradiction'' observed in Table I, where the lower neutron yield in experiment No. 2 corresponds to a higher ion temperature, can be understood. Such good agreement between the experimental and calculated data is natural for experiment No. 2, in which the mean intensity of the radiation impinging on the hohlraum is $I_0 = E_L/\pi D^2 \tau \simeq 4\times 10^{13}$ W/cm^2 and, therefore, the role of the nonlinear effects in the plasma is negligibly small. For experiment No. 1 we have $I_0 \simeq 1.6\times 10^{14}$ W/cm^2, and the intensity ''acting'' within the cavity $I \simeq 5\times 10^{14}$ W/cm^2. As was noted above, under these conditions the influence of the ''fast'' ions interacting with the central target can be significant. Additional experimental and theoretical studies are needed to clarify the questions that have arisen here.

This work was performed with financial support from the State Committee of the Russian Federation for Science and Technologies in the Iskra-5 Laser Thermonuclear Facility (Registration No. 01-50), as well as the Russian Fund for

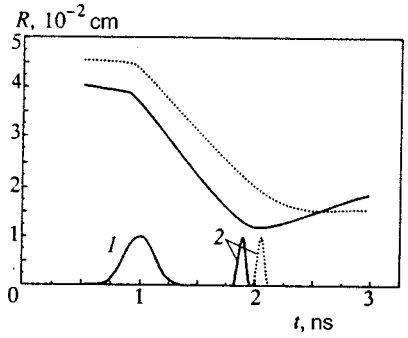

FIG. 7. Calculated R–t diagram of the motion of the boundary between the gas and glass for the laser (*1*) and neutron (*2*) pulses in experiments No. 1 (solid curves) and No. 2 (dotted curves).

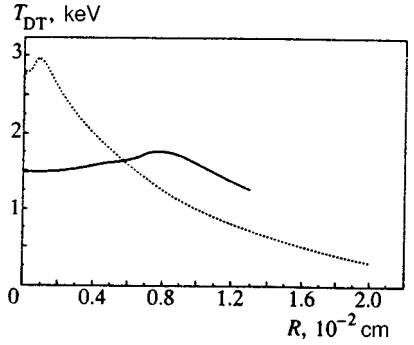

FIG. 8. Calculated profiles of the ion temperature in DT gas at the time of the generation of the maximum neutron flux for experiments No. 1 (solid curve) and No. 2 (dotted curve).

Fundamental Research (Grant No. 96-01-00046).

*)E-mail: belkov@otd13.vniief.ru
†)E-mail: kochemasov@otd13.vniief.ru

[1] G. G. Kochemasov, in *Proceedings of the 23rd European Conference (St. John's College, Oxford, 19–23 September 1994), Institute of Physics Conference Series, No. 140*, (1995), p. 17.
[2] S. A. Bel'kov, A. V. Bessarab, V. A. Gaydash *et al.*, in *Advances in Laser Interaction with Matter and Inertial Fusion (Madrid, Spain)*, G. Velarde and J. M. Martinez-Val (Eds.), World Scientific, Singapore (1997), p. 469.
[3] F. M. Abzaev, V. I. Annenkov, V. G. Bezuglov *et al.*, JETP Lett. **58**, 28 (1993).
[4] S. A. Bel'kov, A. V. Bessarab, V. A. Gaydash *et al.*, in *Advances in Laser Interaction with Matter and Inertial Fusion (Madrid, Spain)*, G. Velarde and J. M. Martinez-Val (Eds.), World Scientific, Singapore (1997), p. 437.
[5] F. M. Abzaev, A. V. Bessarab, G. G. Kirillov *et al.*, Vopr. At. Nauki Tekh., Ser. Mat. Model. Fiz. Protsessov, No. 4, 68 (1992).
[6] V. P. Lazarchuk, V. M. Murugov, S. I. Petrov, and A. V. Senik, Fiz. Plazmy **20**, 101 (1994) [Plasma Phys. Rep. **20**, 94 (1994)].
[7] A. V. Bessarab, S. G. Garanin, G. A. Kirillov *et al.*, in *Abstracts of Reports to the 9th Conference on the Diagnostics of High-Temperature Plasmas (St. Petersburg, Russia, 2–4 June 1997)*.
[8] S. A. Bel'kov and G. V. Dolgoleva, Vopr. At. Nauki Tekh., Ser. Mat. Model. Fiz. Protsessov, No. 1, 59 (1992).
[9] S. A. Bel'kov, P. D. Gasparyan, Yu. K. Kochubeĭ, and E. I. Mitrofanov, Zh. Éksp. Teor. Fiz. **111**, 496 (1997) [JETP **84**, 272 (1997)].

Translated and edited by P. Shelnitz

Dusty plasma induced by solar radiation under microgravitational conditions: an experiment on board the Mir orbiting space station

V. E. Fortov, A. P. Nefedov, O. S. Vaulina,*) A. M. Lipaev, V. I. Molotkov, A. A. Samaryan, V. P. Nikitskiĭ, A. I. Ivanov, S. F. Savin, A. V. Kalmykov, A. Ya. Solov'ev, and P. V. Vinogradov

Scientific-Research Center for the Thermal Physics of Pulsed Effects, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 22 June 1998)
Zh. Éksp. Teor. Fiz. **114**, 2004–2021 (December 1998)

The dynamics of the formation of ordered structures of macroparticles charged by photoemission under the action of solar radiation under microgravitational conditions without the use of electrostatic traps to confine the particles is studied experimentally and theoretically. The working conditions needed for the formation of structures of charged macroparticles are chosen as a result of a numerical solution of the problem posed, the particle charges and the interparticle interaction parameter are determined, and the characteristic times specifying the dynamics of the formation of an ordered system of macroparticles are calculated. The behavior of an ensemble of macroparticles under the effect of solar radiation is observed experimentally on board the Mir space station. An analysis and comparison of the results of the experimental and theoretical investigations permit drawing a conclusion regarding the possibility of the existences of extended ordered formations of macroparticles charged by photoemission under microgravitational conditions. © *1998 American Institute of Physics.* [S1063-7761(98)00712-4]

1. INTRODUCTION

Space research has revealed the importance of dust and dust structures in the formation of stars, planetary systems, and planetary rings, in processes occurring in the upper layers of the atmosphere (the magnetosphere and ionosphere), etc.[1–3] One of the mechanisms for charging dust particles under the conditions of outer space in the presence of intense fluxes of ultraviolet radiation is photoemission. As a result of this process, macroparticles measuring several microns can acquire a positive charge of the order of 10^2-10^5 electron charges.[4] On the other hand, under the conditions of a low-pressure gas discharge, the strong interparticle correlation resulting from the large values of the macroparticle charge (of the order of 10^2-10^5 electron charges) leads to the formation of ordered structures in the arrangement of the macroscopic particles, which are similar to the structures in a liquid or a solid.[5–11] The principal mechanism for charging particles immersed in a radio-frequency (rf) or dc discharge relies on electron and ion fluxes. Because of the higher temperature and mobility of electrons, the particle charge is negative.

One common feature of this group of experiments is the fact that the ordered structures observed do not have a free boundary, since they are confined by the electric field of a striation or the electrodes in the earth's gravitational field and by the potential well formed by the field of the rf discharge or the floating potential of the walls of the gas-discharge vessel in the horizontal direction, respectively. Experiments involving the observation of ordered structures of positively charged cerium oxide particles in a laminar jet of a weakly ionized thermal plasma are exceptions.[11–14]

A plasma with positively charged particles can also form as a result of photoemission when particles are irradiated in a buffer gas by a flux of photons with an energy exceeding the work function of a photoelectron escaping from their surface. Under certain conditions (particle size and concentration, wavelength and intensity of the UV radiation, and photoelectron work function) crystalline structures can appear in such a system.[4,11] The characteristic value of the photoelectron work function for most substances does not exceed 6 eV; therefore, photons with an energy $\leqslant 12$ eV can charge particles without ionizing a buffer gas, such as He or Ar.

Three principal mechanisms can be singled out among the mechanisms by which dust acquires a positive charge. They are the thermionic emission, photoemission, and secondary emission of electrons from the surfaces of dust particles, which, along with charging by electrons, can play a significant role both in the formation of cosmic dust structures and in processes occurring in the upper layers of the atmosphere. The existence of different mechanisms for charging dust in outer space can cause the agglomeration and growth of particles due to the electrostatic attraction of dust particles with opposite charges[15,16] or lead to the formation of plasma-dust structures with a predominant contribution of one of the mechanisms for charging macroparticles under consideration. The investigation of such structures is promising from the standpoint of both basic science and technological applications. It is noteworthy that the possibility of studying plasma-dust crystals with free boundaries can be realized most fully only under the conditions of weightlessness or microgravitation.[17] The study of the formation of ordered structures of charged macroparticles under microgravitational conditions yields new information, which can-

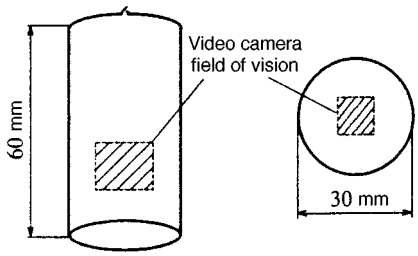

FIG. 1. Geometric dimensions of the working chamber.

not be obtained under laboratory conditions on earth.

The purpose of performing the space experiment was to study the possibility of the existence of plasma-dust structures in the upper layers of the earth's atmosphere when the particles are charged by solar radiation as a result of the photoemission of electrons from their surface. The working conditions needed for the formation of structures of charged macroparticles in the experimental chamber (the type of buffer gas and its pressure, as well as the concentration, material, and size of the particles) were selected as a result of a preliminary numerical analysis of the problem posed; the particle charges induced by solar radiation and the interparticle interaction parameter were determined, and the times specifying the dynamics of the formation of an ordered system of macroparticles (the charging time, braking time, and dispersion times of particles in the working chamber and the times for establishing ordered dust structures) under microgravitational conditions without the use of electrostatic traps to confine the particles were calculated. The calculations were performed for particles of different materials and sizes (1–100 μm) with variation of their concentration and the pressure of the buffer gas.

Experimental investigations of the behavior of an ensemble of macroparticles charged by solar radiation were performed under microgravitational conditions on board the Mir space station.

2. FORMATION OF ORDERED STRUCTURES OF MACROPARTICLES UNDER THE ACTION OF SOLAR RADIATION

2.1. Photoemission charging of particles

Let us consider macroparticles in a neutral gas being irradiated by a source, whose emission has an intensity and a spectral composition corresponding to the characteristics of solar radiation with consideration of the spectral transmission coefficient of the illuminator, the air layer, and the window of the working chamber (Fig. 1) with the particles under investigation.

The conditions for the formation of ordered structures of macroparticles induced by the effects of solar radiation are investigated using particles of cerium oxide (CeO_2), particles of lanthanum boride (LaB_6), and spherical particles of a bronze with a cesium coating. The choice of these particle materials is specified by the efficiency of their photoemission charging and low adhesion, so that the particles under investigation would not adhere to one another and would not precipitate on the walls of the working chamber. The reference data on the quantum efficiency Y of the particle materials in the near-UV and visible regions of the spectrum, the work function W for photoemission, and the densities ρ of the particle materials are listed in Table I, which also indicates the size (radius) r_p and initial concentration n_p of the particles. The initial concentration n_p of the particles is determined by the condition of transparency of the disperse system to the external photoinducing radiation, on the one hand, and the possibility of achieving the maximum values of the interparticle interaction parameter Γ, on the other hand. Taking into account the optical characteristics (the refractive index and sizes) of the particles investigated, we can estimate the optical density τ of the disperse layer as

$$\tau \approx 2\pi r_p^2 n_p H, \tag{1}$$

where r_p is the particle radius and H is the height of the vessel with the particles. The interparticle interaction parameter Γ can be written as a function of the particle concentration n_p in the form

$$\Gamma = (Ze)^2 (4\pi n_p/3)^{1/3}/T_g. \tag{2}$$

Here $l=(4\pi n_p/3)^{-1/3}$ is the mean distance between particles, Z is the particle charge, and T_g is the temperature of the particles, which is equal to the temperature of the buffer gas ($\simeq 0.03$ eV). Thus, an optical density $\tau \approx 1$, which permits achievement of the maximum values of Γ with lowering of the radiation intensity roughly by a factor of $e \approx 2.78$, was selected as a criterion for selecting the value of n_p. The initial particle concentration was determined from (1) as

$$n_p \approx 1/2\pi r_p^2 H. \tag{3}$$

The mass of the particle load in the vessel was calculated from the volume of the vessel (see Fig. 1), the required concentration n_p, and the mass of an individual particle with the mean radius r_p (Table I). The results of experimental investigations of the behavior of CeO_2, LaB_6, and bronze particles under low-pressure conditions (0.01–100 Torr) were also taken into account. More specifically, because of the agglomeration of CeO_2 and LaB_6 particles and their precipitation on the vessel wall, the particle concentration in the working volume decreases by roughly an order of magni-

TABLE I. Values of the radius r_p and the initial concentration n_p of particles with a material of density ρ, the quantum yield Y, the work function W, the limiting particle charge $Z_{\max}$ (4), the interparticle interaction parameter $\Gamma_{\max}$ (2), and the effective photon flux density J (6) of solar radiation.

Particles	r_p, μm	ρ, g/cm^3	W, eV	Y	J, photons/cm^2	n_p, 1/cm^3	$Z_{\max}, e$	$\Gamma_{\max}$
CeO_2	0.5–1.5	7.3	3	10^{-2}	1.33×10^{16}	7.0×10^{6}	7.8×10^{2}	4.2×10^{2}
LaB_6	1–5	2.6	2	10^{-2}	9.16×10^{16}	7.2×10^{6}	4.5×10^{3}	1.4×10^{4}
Bronze	25–50	8.2	1.5	10^{-4}	1.72×10^{17}	5.6×10^{3}	6.9×10^{4}	6.5×10^{5}

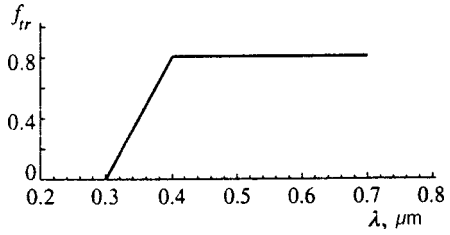

FIG. 2. Spectral transmission function $f_{tr}(\lambda)$ of the device.

tude. Values of n_p ten times smaller than the table values were used to estimate the mean distance between these particles (CeO_2, LaB_6) and Γ_{max} (Table I). The bronze particles did not adhere to one another when the vessel was evacuated, and the table value of the particle concentration corresponds to the basic theoretical parameter n_p.

The limiting (maximum) charges Z_{max} of the particles following the photoemission of electrons from their surface and the value of the interparticle interaction parameter Γ_{max} can be estimated from the condition of equality between the surface potential ϕ_s and the quantity $h\nu_{max}-W$ as

$$Z_{max}=(h\nu_{max}-W)r_p/e, \tag{4}$$

where $h\nu_{max}$ is the maximum photon energy, which corresponds in our case to a wavelength $\lambda_{min}\simeq 0.3\,\mu$m and is determined by the transmission function f_{tr} of the experimental vessel (the transmission of the quartz illuminator, the air layer, and the glass of the working window of the vessel). The dependence of the approximating transmission function f_{tr} of the apparatus on the wavelength λ used in the further calculations is shown in Fig. 2. The values of Z_{max} obtained from Eq. (4) and the values of Γ_{max} obtained from Eq. (2) are listed in Table I. It can easily be seen that the value of Γ and the charge of the particles increase with their size. However, the selection of large particles with $r_p>100\,\mu$m for performing the experiments is undesirable, since it requires lowering the initial particle concentration n_p [see (3)] to values that are unsuitable for observation.

Let us estimate the value of the charge of the particles when they are irradiated by a source with a solar spectrum, taking into account that the plasma-dust system under consideration consists of positively charged macroparticles and photoelectrons emitted by them. The positive potential of the particles is established as a result of the balance between the recombination of electrons on a particle surface and the photoemission electron flux from the particle surface. In the case of a dilute plasma (with a concentration of neutrals $n_n \leqslant 10^{14}$ cm^{-3}), in which the mean free path l_e of photoelectrons before collisions with neutrals greatly exceeds the particle radius r_p ($l_e \gg r_p$), the balance equation can be written in the following form:[4]

$$4n_e\left(\frac{T_e}{2\pi m_e}\right)^{1/2}\left(1+\frac{e\phi_s}{T_e}\right)=YJ\exp\left(-\frac{e\phi_s}{T_{pe}}\right), \tag{5}$$

where m_e is the electron mass, n_e and T_e are the electron concentration in the bulk of the plasma and the electron temperature, J is the photon flux density, Y is the quantum yield of photoelectrons, and T_{pe} is their mean energy. It is assumed here that the efficiency of the absorption of UV radiation is close to unity. The temperature T_{pe} of the electrons leaving a particle surface upon photoemission depends on the particle material and lies in the range from 1 to 2 eV in most cases.[18,19] Assuming that the rate of electron recombination on the particles exceeds the rate of thermal energy loss due to collisions with neutrals, we can set $T_{pe}\simeq T_e$.

In order to determine the equilibrium charge $Z=\phi_s r_p/e$ of the particles under investigation in the working chamber and to estimate the dust charging time, we must calculate the integral number J of solar photons that are capable of causing the photoemission of electrons from a particle surface and estimate the concentration n_e of electrons returning to the particle surface. To calculate the photon flux density J, solar radiation was simulated by a black body with a temperature of 5800 K.[19] The calculation was performed using the formula

$$J=\int_{\lambda_{min}}^{\lambda_{max}}\frac{f_{tr}c_1\lambda^{-4}d\lambda}{\{\exp(c_2/\lambda T_c)-1\}hc}, \tag{6}$$

where f_{tr} is the transmission function of the experimental chamber (Fig. 2), $\lambda_{min}=0.3\,\mu$m, and λ_{max}, i.e., the red edge of the photoeffect, is determined by the work function W for the particle material. The results of the calculations of J for the particles investigated are presented in Table I.

The concentration n_e of electrons returning to the particle surface can be obtained from the solution of the problem of an infinite cylinder uniformly charged throughout its volume. The distributions of the field $E(r)$ and the potential $\phi(r)$ in such a cylinder are specified by the following relations

$$E(r)=2\pi\sigma r, \tag{7a}$$

$$\phi(r)=\pi\sigma(R_{cyl}^2-r^2)+\phi_{wall}, \tag{7b}$$

where $\sigma=e(Zn_p-n_e)=en_e^{wall}$ is the space-charge density and ϕ_{wall} is the potential of the wall of a cylinder of radius R_{cyl}. The floating potential of the surface of the vessel wall ϕ_{wall} is then determined by the photoelectrons escaping from the particles, and the field appearing in the ampul does not allow all the electrons emitted to leave the particle system. The concentration n_e^{wall} of electrons reaching the ampul wall can be estimated from the relation $T_{pe}\simeq e\Delta\phi$, where $e\Delta\phi=e\phi_{av}-e\phi_{wall}$ is the mean energy lost by electrons in the electric field of the vessel. According to the mean-value theorem, from (7b) we have

$$\phi_{av}=\phi_{wall}-\pi\sigma R_{cyl}^2/3.$$

The concentration n_e of electrons remaining in the volume under investigation can be written as

$$n_e=Zn_p-n_e^{wall}\simeq Zn_p-3T_{pe}/\pi e^2R_{cyl}^2. \tag{8}$$

An estimate of the value of n_e^{wall} for photoelectrons with a temperature $T_{pe}=1-2$ eV gives $n_e^{wall}\simeq 5\times 10^6$ cm^{-3} in a vessel with a radius $R_{cyl}=1.5$ cm. The values of Z and Γ calculated on the basis of the values of n_e^{wall} and J obtained are listed in the first row for each type of particle in Table II,

TABLE II. Values of the interparticle interaction parameter Γ and the particle charge Z_p calculated from the balance equation (5) for various values of the macroparticle concentration n_0.

Particles	n_0, cm^{-3}	Z,e	Γ
CeO_2	10^6	3	10^{-2}
	10^5	28	3.0×10^{-2}
	10^4	212	7.5
LaB_6	10^6	14	1.5×10^{-1}
	10^5	131	6.2
	10^4	785	10^2
	10^3	2672	5.5×10^2
Bronze	10^3	5090	2.0×10^3
	10^2	50 070	9.0×10^4

which corresponds to a particle concentration $n_p=n_0$ in the vessel. It should be noted that $n_e\neq0$ only if the initial particle concentration n_0 satisfies the condition

$$n_0 Z_{\max}<n_e^{\text{wall}}.$$

Otherwise, the space-charge density in formulas (7a) and (7b) is defined as $\sigma=eZ_{\max}n_0$.

The results of the calculations presented in this section allow us to postulate the possibility of the formation of crystal- or liquid-type ordered structures of macroparticles for assigned illumination conditions ($h\nu_{\max}$ and J) and macroparticle parameters (W, Y, r_p, and n_0). However, it should be noted that the presence of electrons from the external plasma or photoelectrons returning to the particle surfaces can significantly lower the charge Z of the macroparticles and the value of Γ. In the general case the condition for transparency of the macroparticle cloud to emitted or ''external'' electrons is close to the condition for transparency of the disperse system to the external radiation source causing photoelectron emission from the particle surfaces.

2.2. Characteristic times specifying the dynamics of the formation of dust structures

The possibility of observing ordered structures of macroparticles induced by solar radiation in an experimental chamber with finite dimensions (Fig. 1) is determined by the characteristic times specifying the dynamics of the formation of an ordered system of macroparticles. We shall estimate the charging time, the braking time, the times for establishing ordered macroparticle structures, and the dispersion time of particles in the working chamber with a buffer gas at two different pressures. Neon was selected as the buffer gas for performing the investigations because of its chemical inertness toward the material and cesium coating of the particles, its spectral transparency, and its high ionization potential. The choice of the two different buffer gas pressures (0.01 and 40–70 Torr) was dictated by the possibility of observing the dynamics of the formation of ordered structures for different particle charge values (see Sec. 2.1).

The charging time τ_{ch} of the macroparticles can be determined from the time needed for a particle to acquire the charge $Z=Z_{\max}$ by solving the following differential equation:

TABLE III. Values of the natural frequency ω_0 of charged macroparticles for an initial concentration n_0, the braking time τ_{br}, the drift times of the particles to the vessel wall for lowering the initial particle concentration n_0 by factors of 10 (t_{td1}) 100 and (t_{td2}), and the time for establishing short-range order t_{str} for various values of the pressure P.

Particles	P, Torr	n_0, cm^{-3}	ω_0, s^{-1}	τ_{br}, s	t_{td1}, s	t_{td2}, s	t_{str}, s
CeO_2	0.01	10^6	67.7	1.25	0.023	0.061	0.015
		10^5	21.4		0.074	0.192	0.047
	70	10^6	67.7	1.8×10^{-4}	3.64	40	0.12
		10^5	27.4		36.4	400	1.21
LaB_6	0.01	10^6	124.8	1.32	0.013	0.033	0.008
		10^5	39.5		0.040	0.110	0.025
	70	10^6	124.8	1.9×10^{-4}	1.01	11.14	0.034
		10^5	39.5		10.1	111.4	0.337
Bronze	0.01	10^3	0.778	52	2.00	5.28	1.28
		10^2	0.246		6.41	16.71	4.06
	40	10^3	0.778	0.13	38.11	419.1	1.27
		10^2	0.246		381.1	4191	12.7

$$\frac{dZ}{dt}=\pi r^2\left\{4n_e\left(\frac{T_e}{2\pi m_e}\right)^{1/2}\left(1+\frac{e\phi_s}{T_e}\right)-YJ\exp\left(-\frac{e\phi_s}{T_{\text{pe}}}\right)\right\}. \tag{9}$$

The solution of (9) for both $n_e^{\text{wall}}\simeq5\times10^6$ cm^{-3} and $n_e\simeq0$ when the particles are charged to $Z=Z_{\max}$ gives $\tau_{\text{ch}}<10^{-5}$ s under our conditions.

The braking times τ_{br} of macroparticles in a vessel with two different values of the pressure P are listed in Table III. The braking times τ_{br} for all the particles, except the bronze particles at the ''high'' pressure, were calculated in the free-molecule approximation.[20] The value of τ_{br} for the bronze particles at $P=40$ Torr was determined within the Stokes approximation (the viscosity regime).[21]

The dispersion time of particles in a vessel can be obtained from a numerical molecular-dynamics analysis of the dynamics of the particle system. To investigate the variation of the concentration of charged macroparticles as a function of time, we solved the equation of motion with and without consideration of the thermal (Brownian) motion of the particles:

$$m_d\frac{d^2\mathbf{r}_k}{dt^2}=\sum_j\ \Phi(r)\big|_{r=|\mathbf{r}_k-\mathbf{r}_j|}\frac{\mathbf{r}_k-\mathbf{r}_j}{|\mathbf{r}_k-\mathbf{r}_j|}-m_d\nu\frac{d\mathbf{r}_k}{dt}+\mathbf{F}_B. \tag{10}$$

Here m_d is the mass of a particle, and $\Phi(r)$ is the interparticle interaction parameter. For a Coulomb interaction this quantity can be represented in the form

$$\Phi(r)=\frac{(eZ)^2}{r^2}, \tag{11}$$

where r is the distance between two interacting particles. In the our case Debye screening can be neglected, since the system under consideration is not electroneutral, and the Debye radius is greater than the mean distance between the

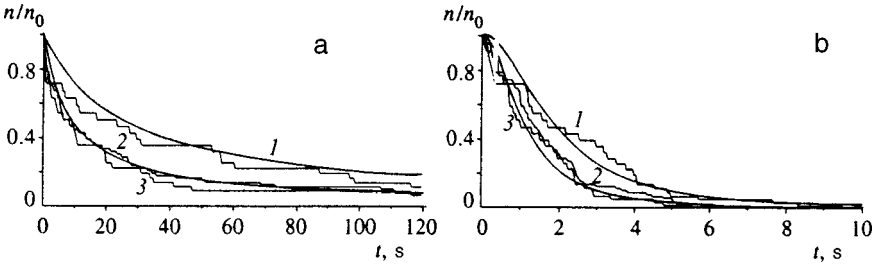

FIG. 3. Molecular-dynamics plots of the concentration $n(t)/n_0$ (solid line) as a function of the time t at the pressures $P=40$ (a) and 0.01 Torr (b) for various values of the initial frequency ω_0: *1* — $\omega_0=0.32$ s^{-1}, $n_0=165$ cm^{-3}, $Z=69\,000$; *2* — $\omega_0=0.53$ s^{-1}, $n_0=465$ cm^{-3}, $Z=69\,000$; *3* — $\omega_0=0.53$ s^{-1}, $n_0=165$ cm^{-3}, $Z=110\,400$. The smooth lines show approximations based on formulas (14a) (a) and (14b) (b).

particles. The system of equations (10) and (11) was solved for a transverse section of a cylindrical vessel with consideration of the absorption of charged macroparticles on the walls under the condition that the particle velocity is equal to zero at the initial moment in time. An analysis of the solution of Eqs. (10) and (11) for the cases of high $[\nu=(\tau_{br})^{-1}\gg 1]$ and low $(\nu\leqslant 1)$ buffer-gas pressures allowed us to conclude that the relative variation of the particle concentration $n(t)/n_0$ as a function of the time t is determined by the natural frequency ω_0 at $t=0$:

$$\omega_0=\sqrt{\frac{(Ze)^2 n_0}{m_d}}, \tag{12}$$

where $n_0=n(t=0)$. The values of ω_0 are given in Table III for various initial concentrations of macroparticles of different materials with the charge $Z=Z_{max}$. In addition, it was found that thermal motion of the particles does not have a significant influence on their dynamic characteristics (dispersion times and correlation functions) at kinetic temperatures K of the particles up to $K=10-50$ eV. Plots of n/n_0 as a function of time are presented in Figs. 3a and 3b for various values of ω_0 and bronze particles with the charge $Z=Z_{max}$ in a gas with high $P=40$ Torr, (Fig. 3a) and low $P=0.01$, (Fig. 3b) pressures.

In order to find an approximation of $n(t)/n_0$ for different values of ω_0, we solved the equation of motion for two charged particles:

$$\frac{d^2r}{dt^2}=-\nu\frac{dr}{dt}+\left(\frac{Ze}{r}\right)^2\frac{1}{m_d}. \tag{13}$$

Here $\nu=(\tau_{br})^{-1}$ is the friction coefficient. Equation (13) was solved for the case of high pressures $(\nu\gg 1)$ in the diffusion approximation with the left-hand side of the equation equal to zero. In the case of low pressures $(\nu\leqslant 1)$ the friction coefficient in (13) was set equal to zero $(\nu=0)$. This allowed us to obtain the following relations for approximating $n(t)/n_0$:

$$n(t)/n_0=(1+3\omega_0^2 t/\nu)^{-1}, \quad \nu\gg 1, \tag{14a}$$

$$n(t)/n_0=8(1+\sqrt{1+4\omega_0^2 t^2})^{-1}, \quad \nu\leqslant 1. \tag{14b}$$

The approximations of $n(t)/n_0$ found are also presented in Figs. 3a and 3b for the corresponding values of ω_0. The functions (14a) and (14b) permit the determination of ω_0 from the results of an experiment in which the macroparticle charge can be obtained for a known concentration n_0 at a certain moment in time $t=0$. The minimum dispersion time of particles of different materials with the charge $Z=Z_{max}$ was estimated on the basis of the dependences obtained. The drift times of the particles to the vessel wall determined from formulas (14a) and (14b) for lowering the initial particle concentration n_0 by factors of 10 (t_{td1}) and 100 (t_{td2}) are listed in Table III for various initial conditions.

The time for establishing ordered macroparticle structures t_{str} in a vessel at low and high buffer-gas pressures can be specified on the basis the solution of (10) and (13) by the following conditions:

$$t_{str}\gg\omega, \quad \nu\leqslant 1, \tag{15a}$$

$$t_{str}\approx\nu/\omega^2, \quad \nu\gg 1, \tag{15b}$$

where $\omega=\sqrt{(Ze)^2 n/m_d}$ is the natural frequency of the charged particles in the structure. The molecular-dynamics calculations show that the times for establishing short-range order correspond to $t_{str}\propto\omega_0$ $(\nu\leqslant 1)$ and $t_{str}\propto 0.1\nu/\omega_0^2$ $(\nu\gg 1)$. These values of t_{str} are listed in Table III for various initial concentrations of particles with the charge $Z=Z_{max}$ and buffer-gas pressures P.

The results of the calculations show that the drift time t_{td} of the particles to the vessel wall is shorter in all the cases considered than the time t_{ch} for photoemission charging of the particles and the time t_{str} for establishing a liquid-type macroparticle structure. Thus, it can be concluded that short-range correlation orders of interparticle distances (liquid-type dust structures) can be observed under microgravitational condition without the use of electrostatic traps to confine the particles. To illustrate the dynamics of the formation of ordered macroparticle structures under these conditions, Figs. 4a and 4b present the results of a molecular-dynamics calculation of pair-correlation functions for bronze particles with a charge $Z_{max}=69\,000$ in a buffer gas (neon) at different pressures. The dashed lines denote correlation functions corresponding to the time t_{str} for the formation of short-range order. The figure also shows fragments of the spatial configuration of the particles in the initial moment in time and at the time $t(n_0/n=10)$ for a decrease in the initial particle concentration $n_0=165$ cm^{-3} by an order of magnitude.

3. EXPERIMENT

3.1. Experimental setup

The experiment was carried out on a setup consisting of the following principal units (Fig. 5):

1. a working chamber;

2. interchangeable glass ampuls containing particles of lanthanum boride LaB_6 (two ampuls), bronze with a cesium

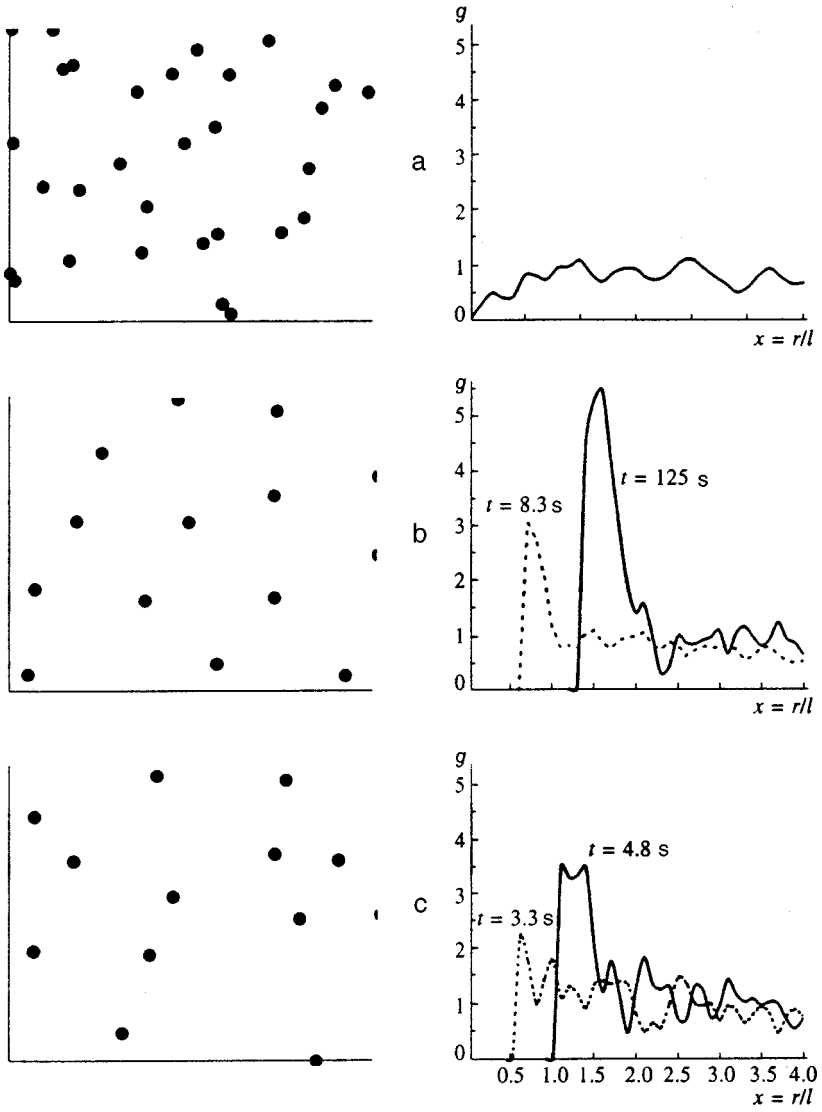

FIG. 4. Fragments of the spatial configuration of the particles and the pair-correlation functions $g(x)$ (where $x=r/l$) for bronze particles with $Z=69\,000$ in a buffer gas with different values of the pressure P at various moments in time t: a — $t=0$ s; b — $P=40$ Torr, $t=125$ s; c — $P=0.01$ Torr, $t=4.8$ s. The dashed lines show the correlation functions corresponding to the time t_{str} for the formation of short-range order.

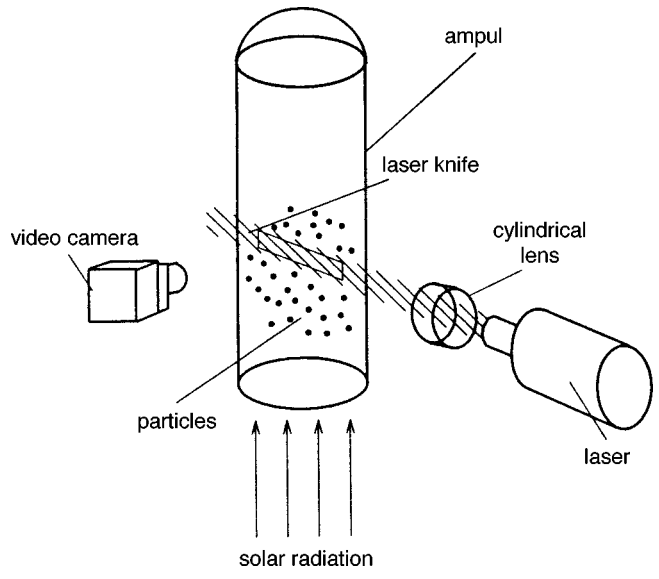

FIG. 5. Schematic representation of the experimental setup.

monolayer (two ampuls), and cerium oxide CeO_2 (two ampuls) in a buffer gas (neon) at different pressures;

3. radiation source — a 30-mW semiconductor laser with a working wavelength of 0.67 μm;

4. a ''Glisser'' television system, including a power supply, an ordinary CCD camera with an objective lens, a tape-recorder module, and a remote-control panel;

5. supports.

The interchangeable glass ampuls had the form of glass cylinders, one of whose end surfaces was a flat uviol glass window and was intended for illuminating the particles with solar radiation (Fig. 1). Immediately before the performance of an experiment, the required ampul was placed in the clamp of the working-chamber holder with its flat end surface toward the illuminator. For diagnostics of the ensemble of particles, the ampul was illuminated by a flat laser beam (a ''laser knife;'' the width of the ''knife'' was no greater than 200 μm), and an image was obtained using the CCD camera, whose signal was recorded on magnetic tape. The field of vision of the video camera had the form of a rectangle measuring roughly 8×10 mm (Fig. 1). The camera was aimed at the center of the ampul (see Fig. 1), and the

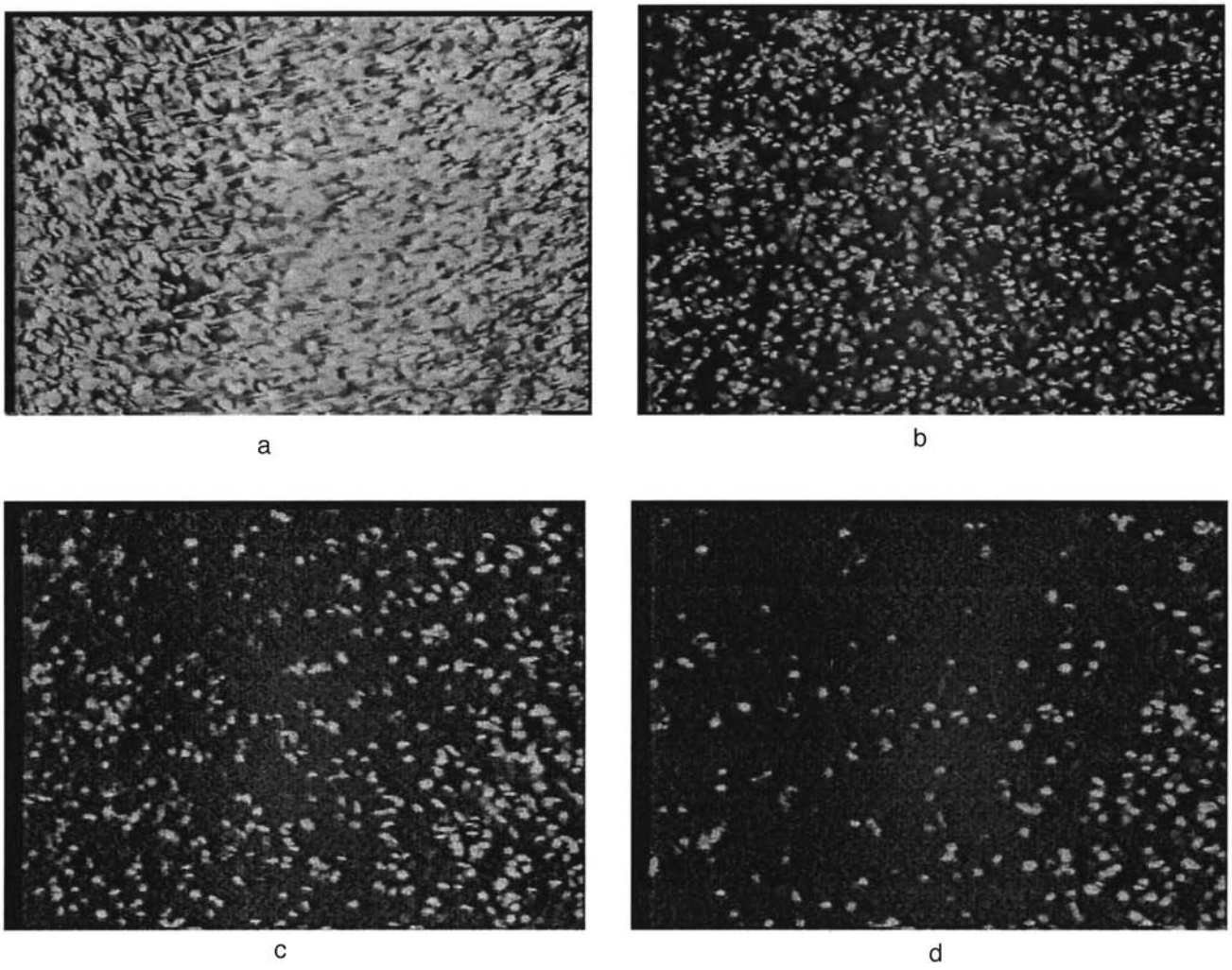

FIG. 6. Successive states of the system of bronze particles in the ampul with $P_1 = 0.01$ Torr following dynamic disturbance of the system.

depth of focus for the *f* stop of 16 chosen was about 10 mm. The experiments were carried out with three values of the working pressure: $P_1 = 0.01$ Torr (for all the types of particles investigated), $P_2 = 40$ Torr (for the bronze particles), and $P_3 = 70$ Torr (for the LaB_6 and CeO_2 particles).

3.2. Behavior of a macroparticle substructure under microgravitational conditions

The first stage of the experiment was confined to observing the behavior of the ensemble of macroparticles placed in the working chamber under the action of solar radiation. In the initial state the particles were on the walls of the ampul; therefore, the experiment was carried out according to the following scheme: a dynamic disturbance (jolt) of the system and relaxation to the initial state, i.e., drift to the wall. The experiments showed that the investigations can be performed only with bronze particles, since the cerium oxide and lanthanum boride particles are incompletely shaken from the walls, rapidly agglomerate, and adhere to the vessel wall, darkening the working region. Therefore, the further analysis was performed only for the bronze particles.

Figures 6a–6d show the successive states of the system of particles in the ampul with $P_1 = 0.01$ Torr following dynamic disturbance of the system, and Figs. 7a–7d show the state of the system in the ampul with $P_2 = 40$ Torr. Observations of the motion of the particles showed that the velocity vectors of the particles are randomly directed in the initial stage and that the particles drift to the walls without a preferential direction. Subsequently, a preferential direction usually appears, but motion along definite trajectories is displayed more strongly in the vessel with the higher pressure (see Fig. 7). Vibration of the particles on a background of the overall translational motion was observed in several experiments, and the treatment of the particle trajectories revealed periodic variations of the magnitude of the particle velocity in all the experiments. These variations of the particle velocity can be associated with fluctuations of the particle charge or with the dynamic action of microscopic accelerating forces arising on board the space station. Variation of the visibility of the particles was observed (one possible cause is rotation). One more interesting finding is the formation of agglomerates, in which the number of particles varies from three or four to several hundred (Figs. 8a and 8b). These agglomerates can depart from the vessel walls in response to

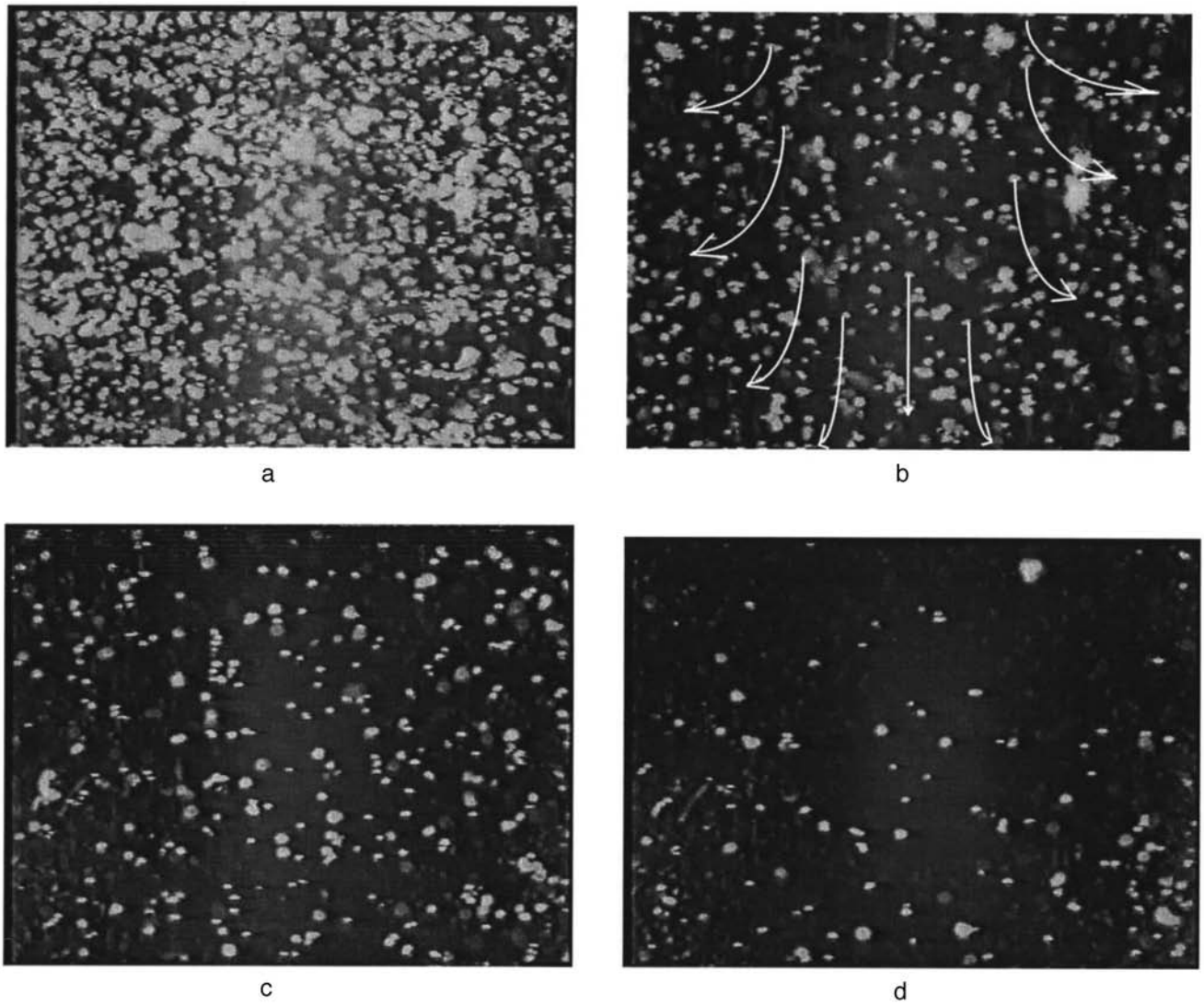

FIG. 7. Successive states of the system of bronze particles in the ampul with $P_1=40$ Torr following dynamic disturbance of the system.

a weak dynamic disturbance. The results of the observations of the behavior of the particles in the vessel with the higher pressure show that the bulk of the agglomerates form in the volume of the vessel during a period of a few seconds following the dynamic disturbance. The agglomeration of particles in the volume of the vessel may occur because the particles acquire opposite charges in the initial moments of illumination: positive charges are acquired as a result of the emission of photoelectrons, and negative charges are imparted by the fluxes of electrons emitted from neighboring particles. A similar effect was observed in Ref. 15.

It was concluded that particles are charged on the basis of observations of the changes in the trajectories of the particles when they come close to one another (collide) or approach the wall. It should also be noted that the particles move very slowly in the vessel with $P_2=40$ Torr when the solar radiation is blocked and that acceleration of the motion occurs when radiation acts on the ensemble of particles. The charge of the macroparticles can be estimated by analyzing their dynamic behavior.

3.3. Determination of the particle charge

According to observations of the behavior of ensembles of particles during illumination, the drift time of the particles to the wall was about 5 s for $P_1=0.01$ Torr and from 3 to 5 min in the vessel with the buffer-gas pressure $P_2=40$ Torr. This suggests (according to the data in Table III) that the bronze particles were charged to a level close to the theoretical value $Z_{\max}=69\,000$. The particle charge can be estimated both from the relative variation of the concentration of the particles and from their trajectories of motion. In the first case approximations of the types (14a) and (14b), which permit the determination of ω_0 from experimental plots of the relative variation of the particle concentration $n(t)/n_0$, can be used to determine the particle charge. Figures 9a and 9b present experimental plots of $n(t)/n_0$ and the approximating functions (14a) and (14b) with $\omega_0=0.315\ \mathrm{s}^{-1}$ which give the best agreement with experiment both for the case of $P_2=40$ Torr and for the vessel with the lower pressure $P_1=0.01$ Torr.

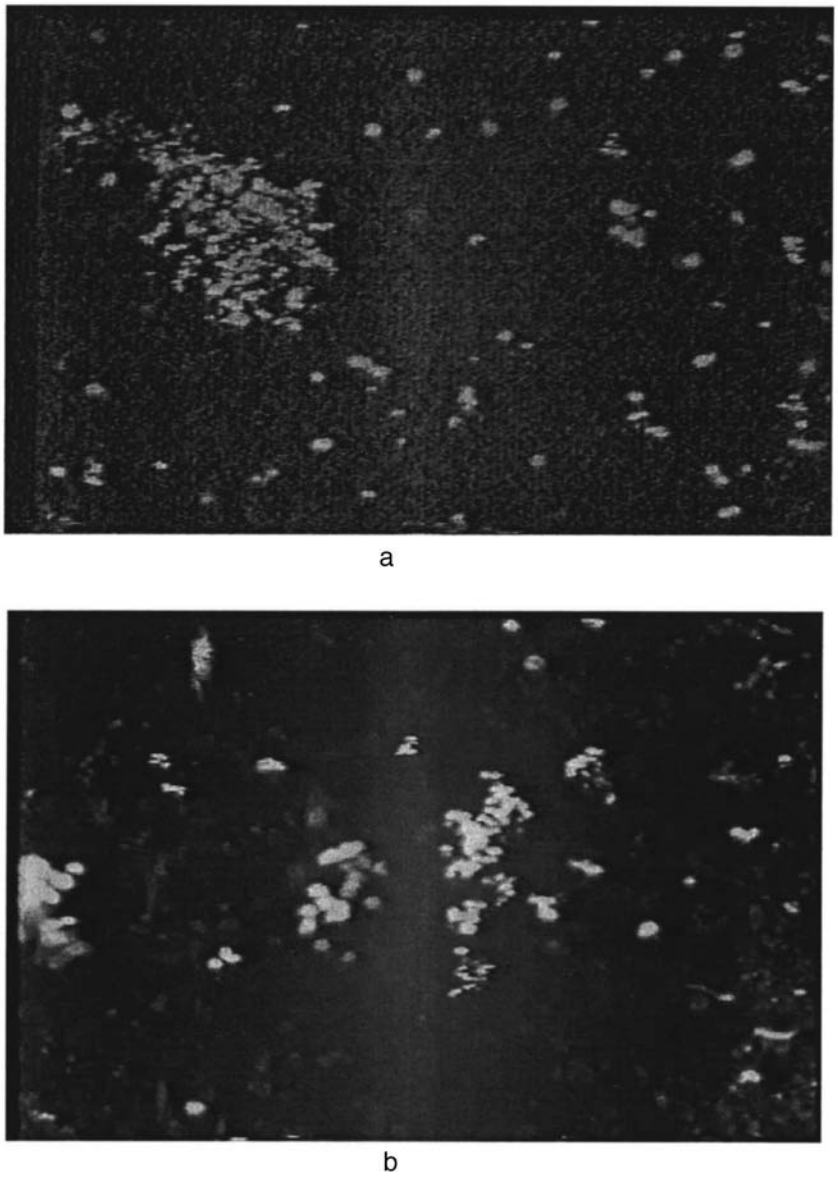

FIG. 8. Formation of macroparticle agglomerates.

The plots of $n(t)/n_0$ were determined during illumination by solar radiation following a period of holding in the dark, during which the particles managed to lower the velocity acquired from the initial jolt (Figs. 9a and 9b). In the case of shaking the vessel while it is exposed to solar radiation, it is very difficult to interpret the experimental data (see Fig. 9), since it is virtually impossible to determine the initial particle concentration n_0 to the required accuracy. The initial concentration n_0 was determined in the absence of solar radiation from the number of particles falling in the plane of the laser knife and ranged from 150 to 250 cm^{-3}. On the basis of the values found $n_0 = 150-250\ \mathrm{cm}^{-3}$ and $\omega_0 = 0.315\ \mathrm{s}^{-1}$, the particle charge can be specified as

$$Z = (\omega_0/e)\sqrt{m_d/n_0} = (6.38 \pm 0.81) \times 10^4.$$

Thus, the charge obtained corresponds to $Z_{\max} = 69\,000$ to within 13%.

The second method for determining the particle charge is

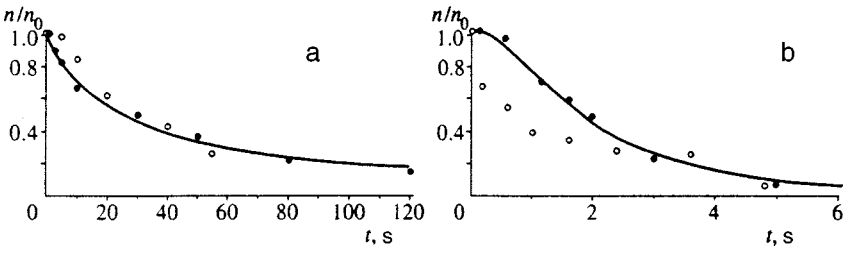

FIG. 9. Experimental plots of the concentration $n(t)/n_0$ as a function of the time t in vessels with pressures equal to 40 (a) and 0.01 Torr (b) obtained at times $t \gg \tau_{\mathrm{br}}$ following dynamic disturbance of the vessel (filled circles) and immediately after agitation of the vessel (unfilled circles). The solid lines show the approximations (14a) and (14b) for the natural frequency $\omega_0 = 0.315\ \mathrm{s}^{-1}$.

TABLE IV. Results of the determination of the charge Z of particles from their trajectories of motion.

No.	P, Torr	n_p, cm^{-3}	a, m/s^2	V, m/s	Z
1		10^1	6×10^{-4}	2.5×10^{-3}	5×10^4
2	0.01	3×10^1	1.2×10^{-3}	6×10^{-3}	10^5
3		5×10^1	1.6×10^{-3}	1.2×10^{-2}	8×10^4
4	40	3×10^1	1.5×10^{-4}	7×10^{-5}	3×10^4

based on solution of the equation of motion of a particle in a known electric field $E(r)$:

$$\frac{d^2r}{dt^2}=-\nu\frac{dr}{dt}+E\frac{Ze}{m_d}. \tag{16}$$

The electric field intensity E can be assigned by Eq. (7) with consideration of averaging over the radius of the cylinder as $E=\pi eZn_pR_{\rm cyl}$. Here we take into account that the space-charge density is $\sigma=eZn_p$ for $Zn_p<n_e^{\rm wall}$ and $\sigma=en_e^{\rm wall}$ for $Zn_p>n_e^{\rm wall}$, since the motion of the particles was analyzed at low concentrations ($n_p<10^2$). After determining the velocity V and the acceleration a of the particles from their trajectories of motion, the particle charge can be found from Eq. (16):

$$Z^2=am_d/(e^2\pi n_pR_{\rm cyl}),\quad P_1=0.01\ \text{Torr}, \tag{17a}$$

$$Z^2=(V+a)m_d/(e^2\tau_{\rm br}\pi n_pR_{\rm cyl}),\quad P_2=40\ \text{Torr}. \tag{17b}$$

Table IV presents the characteristic values of the macroparticle charges, which agree well with the value of the charge obtained from the variation of the relative particle concentration $n(t)/n_0$. It is noteworthy that the slight upward deviation of the charge in the low-pressure case can easily be explained, since lighter particles and, accordingly, smaller particles having a smaller charge leave the vessel volume first under low-viscosity conditions. In the high-viscosity case the drift time of the particles to the wall should not depend strongly on their size.

Despite the high particle charges and the large value of the interaction parameter $\Gamma\simeq10^4$ (see Table IV), no strong correlation between the interparticle distances could be observed. The measured correlation functions exhibit appreciable deviations for the calculated curves (Fig. 4). The typical form of the correlation functions obtained as a result of the treatment of experimental images without illumination (the laser knife) and with solar irradiation is shown in Fig. 10. The difference between the experimental and calculated results can be due to several factors: polydispersity of the particle powder investigated and the formation of agglomerates, in which the number of particles varies from three or four to several hundred. The latter factor makes a highly significant contribution to breakdown of the correlation of the interparticle distances.

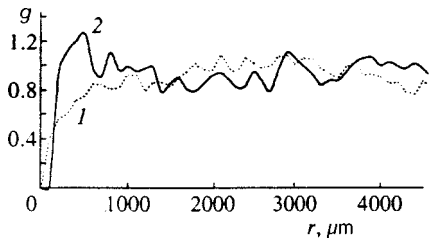

FIG. 10. Experimental correlation functions $g(r)$ obtained as a result of the treatment of images without illumination (*1*) and with solar irradiation (*2*).

4. CONCLUSION

The results of observations of the behavior of an ensemble of macroparticles charged by photoemission under the action of solar radiation under microgravitational conditions have been presented. The particle charges have been estimated, and it has been established that the particles are charged to the maximum possible levels, i.e., several units times 10^4, under the conditions of the experiment. The results of a calculation of the particle charges due to photoemission correspond to the results of the observations. The particle charges have been estimated by two methods, which provide good agreement with the calculation: from the trajectories of motion of the particles and with the use of the approximations obtained for the variation of the relative particle concentration. The values of the charges and the interparticle interaction parameter obtained demonstrate the possibility of the formation of ordered crystal- and liquid-type macroparticle structures under the conditions studied. It is noteworthy that under the conditions of outer space the charging efficiency of particles with a low photoelectron work function is higher due to the additional effect of the UV portion of the solar spectrum and that the interparticle interaction parameter consequently increases.

Although the dynamic behavior of the particles (the variation of the particle concentration in the volume investigated) corresponds well to the behavior determined as a result of a numerical analysis, the form of the experimentally observed correlation functions differs strongly from the theoretical form and attests to the formation of only weakly correlated liquid-type structures. The principal cause of this difference may be the agglomeration of oppositely charged particles during charging. Nevertheless, an analysis and comparison of the results of the experimental and theoretical investigations confirm the conclusion that the existence of extended liquid-type ordered formations of macroparticles charged by solar radiation is possible under microgravitational conditions even if there is substantial agglomeration of the particles.

This work was partially supported by the Russian Fund for for Fundamental Research (Grant No. 98-02-16828).

*)E-mail: ipdustp@redline.ru

[1] S. A. Kaplan, *Interstellar Medium and the Origin of Stars*, [in Russian], Znanie, Moscow (1977).
[2] N. N. Gor'kavyĭ and A. M. Fridman, Usp. Fiz. Nauk **160**, 169 (1990) [Sov. Phys. Usp. **33**(2), 95 (1990)].
[3] F. Melanso and O. Havnes, J. Geophys. Res. **95**, 5837 (1991).
[4] M. Rosenberg and D. A. Mendis, IEEE Trans. Plasma Sci. **23**, 177 (1995).
[5] V. E. Fortov, A. P. Nefedov, V. M. Torchinsky, V. I. Molotkov, O. F. Petrov, A. A. Samarian, A. M. Lipaev, and A. G. Khrapak, Phys. Lett. A **229**, 317 (1997).
[6] J. H. Chu and I. Lin, Phys. Rev. Lett. **72**, 4009 (1994).

[7] H. Thomas, G. E. Morfill, V. Demmel *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
[8] Y. Hayashi and K. Tachibana, J. Phys. Soc. Jpn. **33**, L804 (1994).
[9] A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
[10] V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ, V. I. Molotkov, A. G. Khrapak, O. F. Petrov, and K. F. Volykhin, JETP Lett. **64**, 92 (1996).
[11] A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997) [Phys. Usp. **40**, 1163 (1997)].
[12] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samarian, and A. V. Chernyschev, Phys. Rev. E **54**, 2236 (1996).
[13] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samarian, and A. V. Chernyschev, Phys. Lett. A **219**, 89 (1996).
[14] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samaryan, and A. V. Chernyshev, Zh. Éksp. Teor. Fiz. **111**, 467 (1997) [JETP **84**, 256 (1997)].
[15] T. Yokota and K. Honda, J. Quant. Spectrosc. Radiat. Transf. **56**, 761 (1996).
[16] C. Cui and J. Goree, Trans. Plasma Sci. **22**, 151 (1994).
[17] V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997) [Phys. Usp. **40**, 53 (1997)].
[18] C. K. Goertz, Geophys. Rev. **27**, 271 (1989).
[19] V. A. Grilikhes, P. P. Orlov, and L. B. Popov, *Solar Energy and Space Flights* [in Russian], Nauka, Moscow (1986).
[20] I. T. Yakubov and A. G. Khrapak, Sov. Tech. Rev. B Therm. Phys. **2**, 269 (1989).
[21] D. V. Sivukhin, *General Course in Physics* [in Russian], Nauka, Moscow (1979), Vol. 1, p. 496.

Translated and edited by P. Shelnitz

Director dynamics near the nematic–smectic-*A* phase transition

S. V. Ul'yanov*)

St. Petersburg Institute of Trade and Economics, 194018 St. Petersburg, Russia
(Submitted 25 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 2022–2033 (December 1998)

The Landau–de Gennes model for the free energy of a nematic liquid crystal near the phase transition to the smectic *A*-phase is used to determine the frequency dependence of the fluctuation corrections to the Frank elastic constants. It is shown that the interaction of the fluctuations of the smectic order parameter and the director results in corrections to all the Frank elastic constants. In the low-frequency limit ($\omega \to 0$), the corrections to the Frank elastic constants K_{22} and K_{33} are the largest, and decrease to zero in the infinite-frequency limit. The correction to K_{11} is negative, and vanishes in both limits. The absolute value of the correction to K_{11} is the largest at frequencies in the megahertz range. It is shown that in oriented nematics the interaction of the smectic fluctuations and the director limits deviations of the director from the direction of preferred orientation, as a result of which relaxation of both inhomogeneous and homogeneous distortions of the director field can be observed. It is also shown that this gives rise to a frequency interval in the megahertz range in which shear waves begin to propagate in the nematic. The propagation speed of these waves is roughly a hundred times smaller than that of sound and strongly depends on the direction of propagation. © *1998 American Institute of Physics.* [S1063-7761(98)00812-9]

1. INTRODUCTION

Nematic liquid crystals (NLC) in a state near the phase transition to the smectic *A*-phase (the *N–A*-transition) have been studied for a long time both experimentally and theoretically.[1–5]. De Gennes[6] was the first to phenomenologically describe the properties of NLC near the *N–A*-transition. In fact, using the expression for the free energy derived in Ref. 6, he explained the critical increase of the Frank elastic constants K_{22} and K_{23} as the state of the liquid crystal approaches the *N–A*-transition point by the interaction of the inhomogeneous director field and fluctuations of the smectic order parameter. Later Jahnig and Brochard[7] added the missing factor $1/\sqrt{2}$ to the formulas for K_{22} and K_{33}. These formulas were further refined on the basis of the Landau–de Gennes model[8] and the *NAC*-model.[9]

All the papers just mentioned consider the simplest and most important case, i.e., determining the static corrections to the Frank elastic constants. It is well known, however, that the interaction of hydrodynamic variables (in particular, the director in a NLC is such a variables) and order parameter fluctuations lead to frequency dispersion in the coefficients in the system of equation of motion of the NLC. In the pretransitional region, the viscosity coefficients change significantly, which causes frequency dispersion in the velocity and an anomalous increase in sound absorption.[4]. Jahnig and Brochard[7] and McMillan,[10] who examined the interaction of the director field and the field of the order parameter fluctuations, were the first to predict the anomalous enhancement of some of the viscosity coefficients, an effect manifested in the anisotropy of the acoustic properties of NLC. Later it was found that the interactions of the order parameter and the density[11,12] and the order parameter and the nonuniform velocity field[13–15] result in much larger contributions to the viscosity coefficients. This is reflected in the anomalous increase in isotropic sound absorption[4,16,17] and in the emergence of substantial anisotropy in the velocity dispersion and absorption of sound.[11,18]. As yet, however, there is no satisfactory quantitative description of the experimental data; there are even discrepancies in the data from different experiments. This is especially true of the anisotropy of acoustic properties. The frequency dispersion of the Frank elastic constants may be discerned in experiments in which the NLC is subjected to a periodic external force. This is true of acoustooptic effects observed in a NLC placed in an external field of an acoustic wave, and of Freedericksz optical transitions resulting from periodic variations of the intensity of the incident light.

The aim of the present paper is to use the Landau–de Gennes model for the free energy of a NLC near the *N–A*-transition point to study some of the aspects of the interaction of the director field and the order parameter fluctuations. In Sec. 2, the frequency dispersion of all the Frank elastic constants is established. Section 3 shows that a relaxation mechanism for the attenuation of the distortions of the equilibrium director field arises near t he *N–A*-transitionpoint, which strongly affects the propagation of shear waves. Finally, in Sec. 4, the possibility of experimentally observing the results obtained in this paper is discussed.

2. FREQUENCY DISPERSION OF THE FRANK ELASTIC CONSTANTS

Assuming that the equilibrium orientation of the director vector $\mathbf{n}_0$ is fixed by the conditions at the sample boundaries,

which are assumed to be at an infinite distance from the sample volume of interest, we can write the free energy of a deformed NLC as follows:

$$F = F_0 + F_{\mathrm{Fr}}, \tag{1}$$

where F_0 is the free energy of the undeformed sample, and

$$F_{\mathrm{Fr}} = \frac{1}{2}\int d\mathbf{r}\,[K_{11}(\operatorname{div}\mathbf{n})^2 + K_{22}(\mathbf{n}\cdot\operatorname{curl}\mathbf{n})^2 + K_{33}(\mathbf{n}\times\operatorname{curl}\mathbf{n})^2] \tag{2}$$

is the free energy of Frank deformation, with

$$\mathbf{n}(\mathbf{r},t) = \mathbf{n}_0 + \delta\mathbf{n}(\mathbf{r},t) \tag{3}$$

the local director. Following Refs. 7 and 19, we introduce a molecular field $\mathbf{h}_{\mathrm{Fr}}$ that tends to straighten the distorted director fields. In a system of coordinates in which the z axis coincides with the direction of $\mathbf{n}_0$, the molecular field related to the Frank energy is[7]

$$h_{x\,\mathrm{Fr}}(q,\omega) = -(K_{11}q_x^2 + K_{22}q_y^2 + K_{33}q_z^2)\,\delta n_x(q,\omega) + (K_{22}-K_{11})q_x q_y\,\delta n_y(q,\omega), \tag{4}$$

where we have used the Fourier expansion

$$\mathbf{n}(\mathbf{r},t) = \frac{1}{\sqrt{V}}\sum_{\mathbf{q}}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\mathbf{n}(\mathbf{q},\omega)\exp(i\mathbf{q}\cdot\mathbf{r} - i\omega t). \tag{5}$$

In a NLC that is near the *N–A*-transition point, fluctuations of the smectic order parameter Ψ become important,[1–3,6,7] in view of which to the free energy we must add a term F_Ψ related to the fluctuation-emerging smectic order. In the Landau–de Gennes model,[1–3,6,7]

$$F = F_0 + F_{\mathrm{Fr}} + F_\Psi, \tag{6}$$

where

$$F_\Psi = \int d\mathbf{r}\left[A|\Psi|^2 + \frac{1}{2M_T}|(\nabla_\perp - iq_s\delta\mathbf{n})\Psi|^2 + \frac{1}{2M_V}|\nabla_\parallel\Psi|^2\right], \tag{7}$$

$$\nabla_\perp = \mathbf{i}\frac{\partial}{\partial x} + \mathbf{j}\frac{\partial}{\partial y}, \quad \nabla_\parallel = \frac{\partial}{\partial z}.$$

The factor A vanishes at the transition temperature T_{N-A} and is usually written in the form $A = a_0(T - T_{N-A})^\gamma$, where $\gamma = 3/4$ if we use the helium analogy or $\gamma = 1$ in the mean-field approximation, and $q_s = 2\pi/d$, with d the distance between the smectic layers in the low-temperature phase.

Earlier de Gennes,[6] Jahnig and Brochard,[7] and Andereck and Patton[8] found that the presence in (7) of a term in which $\delta\mathbf{n}$ is coupled with Ψ gives rise to anomalous additional terms in the Frank elastic constants K_{22} and K_{33}, which were found in the static case ($\omega = 0$). When smectic order fluctuations are present, the molecular field consists of two terms:[7]

$$\mathbf{h} = \mathbf{h}_{\mathrm{Fr}} + \langle\mathbf{h}_\Psi\rangle. \tag{8}$$

where

$$\mathbf{h}_\Psi = -\frac{iq_s}{2M_T}(\Psi^*\nabla_\perp\Psi - \Psi\nabla_\perp\Psi^*) - \frac{q_s^2}{M_T}\Psi^*\Psi\,\delta\mathbf{n}, \tag{9}$$

and statistical averaging is over all realizations of the order parameter fluctuations. In the Fourier representation we have

$$\langle\mathbf{h}_\Psi\rangle = \langle\mathbf{h}^{(1)}\rangle + \langle\mathbf{h}^{(2)}\rangle, \tag{10}$$

where

$$\langle\mathbf{h}^{(1)}(\mathbf{q},\omega)\rangle = \frac{q_s}{2M_T V}\sum_{\mathbf{q}'}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}(\mathbf{q}_\perp - 2\mathbf{q}'_\perp) \times\langle\Psi^*(\mathbf{q}',\omega')\Psi(\mathbf{q}-\mathbf{q}',\omega-\omega')\rangle, \tag{11}$$

$$\langle\mathbf{h}^{(2)}(\mathbf{q},\omega)\rangle = -\frac{q_s}{M_T\sqrt{V}}\sum_{\mathbf{q}',\mathbf{q}''}\int_{-\infty}^{\infty}\frac{d\omega'\,d\omega''}{(2\pi)^2} \times\langle\Psi^*(\mathbf{q}',\omega')\Psi(\mathbf{q}'',\omega'') \times\delta\mathbf{n}(\mathbf{q}-\mathbf{q}'-\mathbf{q}'',\omega-\omega'-\omega'')\rangle. \tag{12}$$

To calculate the fluctuation contribution to the molecular field we use Levanyuk's method,[20] which has been repeatedly used in finding the fluctuation corrections to the viscosity coefficients near phase transition points in liquid crystals.[15,21,22]. In this method, the fluctuations of the smectic order parameter Ψ, which develop in a deformed NLC, are described by the Langevin equation

$$\frac{\partial\Psi}{\partial t} = -b\left[A\Psi - \frac{1}{2M_T}(\nabla_\perp - iq_s\delta\mathbf{n})^2\Psi - \frac{1}{2M_V}\nabla_\parallel^2\Psi\right] + f, \tag{13}$$

where f is a random force. The complex conjugate of Eq. (13) is the equation for Ψ^*. When we retain the interaction in (13) to lowest order, only the terms linear in $\delta\mathbf{n}$ survive. In the Fourier representation we have

$$\Psi(\mathbf{q},\omega) = G^0(\mathbf{q},\omega)\left[f(\mathbf{q},\omega) + \frac{bq_s}{2M_T\sqrt{V}} \times\sum_{\mathbf{q}'}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}((2\mathbf{q}_\perp - \mathbf{q}'_\perp)\cdot\delta\mathbf{n}(\mathbf{q}'\omega')) \times\Psi(\mathbf{q}-\mathbf{q}',\omega-\omega')\right], \tag{14}$$

where

$$G^0(\mathbf{q},\omega) = [-i\omega + b\chi^{-1}(\mathbf{q})]^{-1}, \tag{15}$$

$$\chi^{-1}(\mathbf{q}) = A + \frac{q_\perp^2}{2M_T} + \frac{q_Z^2}{2M_V}. \tag{16}$$

The equation for $\Psi^*(\mathbf{q},\omega)$ can be obtained from Eq. (14) by replacing Ψ with Ψ^*, f with f^*, and the plus in front of the integral part on the right-hand side of Eq. (14) with a minus. A formal solution of Eq. (14) can be obtained via iterations in the form of a power series in $\delta\mathbf{n}$. After the second iteration we have

$$\Psi(\mathbf{q},\omega)=G^0(\mathbf{q},\omega)f(\mathbf{q},\omega)+\frac{bq_s}{2M_T\sqrt{V}}\sum_{\mathbf{q}'}\int_{-\infty}^{\infty}\frac{d\omega'}{2\pi}$$
$$\times((2\mathbf{q}_\perp-\mathbf{q}'_\perp)\cdot\delta\mathbf{n}(\mathbf{q}',\omega'))G^0(\mathbf{q},\omega)$$
$$\times G^0(\mathbf{q}-\mathbf{q}',\omega-\omega')f(\mathbf{q}-\mathbf{q}',\omega-\omega'). \quad (17)$$

We then substitute the solution (17) and a similar expression for $\Psi^*(\mathbf{q},\omega)$ into the expressions (11) and (12) for the contributions to the molecular field. Statistical averaging of the resulting expressions with allowance for the fact that the random force f is uncorrelated in both space and time yields

$$\langle f(\mathbf{q},\omega)f^*(\mathbf{q}',\omega')\rangle=\frac{2bk_BT(2\pi)^4}{V}\delta(\mathbf{q}+\mathbf{q}')\,\delta(\omega+\omega'). \quad (18)$$

Integration of the delta functions and integration over the frequencies yields

$$\langle\mathbf{h}^{(1)}(\mathbf{q},\omega)\rangle=\frac{bk_BTq_s^2}{4(2\pi)^3M_T^2}$$
$$\times d\mathbf{q}'\frac{(\mathbf{q}_\perp-2\mathbf{q}'_\perp)((\mathbf{q}_\perp-2\mathbf{q}'_\perp)\cdot\delta\mathbf{n}(\mathbf{q},\omega))}{-i\omega+b(\chi^{-1}(\mathbf{q}')+\chi^{-1}(\mathbf{q}-\mathbf{q}'))}$$
$$\times(\chi(\mathbf{q}')+\chi(\mathbf{q}-\mathbf{q}')), \quad (19)$$

$$\langle\mathbf{h}^{(2)}(\mathbf{q},\omega)\rangle=-\frac{k_BTq_s^2}{(2\pi)^3M_T}\int\frac{d\mathbf{q}'}{\chi^{-1}(\mathbf{q}')}\delta\mathbf{n}(\mathbf{q},\omega). \quad (20)$$

In (19) and (20) we have replaced the summation over the wave vectors by integration:

$$\sum_{\mathbf{q}'}\rightarrow\frac{V}{(2\pi)^3}\int d\mathbf{q}'.$$

To find the corrections to the Frank elastic constants, we need only retain terms proportional to q^2 in the expression for $\langle\mathbf{h}_\Psi\rangle$. Omitting the details of the calculations (which can be found in the Appendix), we arrive at an expression, which in form coincides with (4), for the fluctuation contribution to the molecular field $\langle\mathbf{h}_\Psi(\mathbf{q},\omega)\rangle$. We may assume that the interaction of the director and the smectic order fluctuations gives rise to complex-valued corrections to the Frank elastic constants. Thus,

$$K_{11}(\omega)=K_{11}^0+\frac{k_BTq_s^2\xi_\perp^2}{24\pi\xi_\parallel}[F_1(\tilde{\omega})+iG_1(\tilde{\omega})], \quad (21)$$

$$K_{22}(\omega)=K_{22}^0+\frac{k_BTq_s^2\xi_\perp^2}{24\pi\xi_\parallel}[1+F_2(\tilde{\omega})+iG_2(\tilde{\omega})], \quad (22)$$

$$K_{33}(\omega)=K_{33}^0+\frac{k_BTq_s^2\xi_\parallel^2}{24\pi}[1+F_2(\tilde{\omega})+iG_2(\tilde{\omega})], \quad (23)$$

where K_{ii}^0 $(i=1,2,3)$ are the ''bare'' values of the Frank elastic constants, which ignore the interaction of the director and the order parameter fluctuations. Here we used the ordinary notation for the correlation lengths:

$$\xi_\perp=(2AM_T)^{-1/2}, \quad (24)$$

$$\xi_\parallel=(2AM_V)^{-1/2}. \quad (25)$$

The dependence of the Frank elastic constants on the reduced frequency $\tilde{\omega}$, where

$$\tilde{\omega}=\omega\tau_\Psi=\omega\frac{1}{2bA} \quad (26)$$

is determined by the functions

$$F_1(\tilde{\omega})=\frac{24}{5\tilde{\omega}^3}\left[-\frac{5}{2}\tilde{\omega}+\tilde{\omega}\sqrt{2\sqrt{1+\tilde{\omega}^2}+2}\right.$$
$$\left.+\left(\frac{1}{2}-\frac{3}{16}\tilde{\omega}^2\right)\sqrt{2\sqrt{1+\tilde{\omega}^2}-2}\right], \quad (27)$$

$$G_1(\tilde{\omega})=\frac{24}{5\tilde{\omega}^3}\left[-1+\frac{5}{4}\tilde{\omega}^2+\left(\frac{1}{2}-\frac{3}{16}\tilde{\omega}^2\right)\right.$$
$$\left.\times\sqrt{2\sqrt{1+\tilde{\omega}^2}+2}-\tilde{\omega}\sqrt{2\sqrt{1+\tilde{\omega}^2}-2}\right], \quad (28)$$

$$F_2(\tilde{\omega})=\frac{8}{5\tilde{\omega}^3}\left[-\frac{5}{2}\tilde{\omega}+\tilde{\omega}\sqrt{2\sqrt{1+\tilde{\omega}^2}+2}\right.$$
$$\left.+\left(\frac{1}{2}+\frac{7}{16}\tilde{\omega}^2\right)\sqrt{2\sqrt{1+\tilde{\omega}^2}-2}\right]-1, \quad (29)$$

$$G_2(\tilde{\omega})=\frac{8}{5\tilde{\omega}^3}\left[-1+\left(\frac{1}{2}+\frac{7}{16}\tilde{\omega}^2\right)\sqrt{2\sqrt{1+\tilde{\omega}^2}+2}\right.$$
$$\left.-\tilde{\omega}\sqrt{2\sqrt{1+\tilde{\omega}^2}-2}\right]. \quad (30)$$

From (27)–(30) it follows that in the $\omega\rightarrow0$ limit we have

$$F_1(0)=G_1(0)=F_2(0)=G_2(0)=0, \quad (31)$$

so that in the static case we arrive at the well-known result[6–9]

$$K_{11}(\omega=0)=K_{11}^0, \quad (32)$$

$$K_{22}(\omega=0)=K_{22}^0+\frac{k_BTq_s^2\xi_\perp^2}{24\pi\xi_\parallel}, \quad (33)$$

$$K_{33}(\omega=0)=K_{33}^0+\frac{k_BTq_s^2\xi_\parallel}{24\pi}. \quad (34)$$

At infinite frequency $(\omega\rightarrow\infty)$ the fluctuation corrections disappear:

$$F_1+iG_1\rightarrow0, \quad 1+F_2+iG_2\rightarrow0.$$

Figures 1 and 2 depict F_1, F_2, G_1, and G_2 as functions of the reduced frequency $\tilde{\omega}$, which determine the frequency dispersion of the Frank elastic constants. These constants change significantly as the reduced frequency $\tilde{\omega}$ varies in the 1–100 range. For a NLC near the *N*–*A*-transition point, this corresponds to frequency variations in the megahertz range. At lower frequencies we can use the static corrections to the Frank elastic constants.

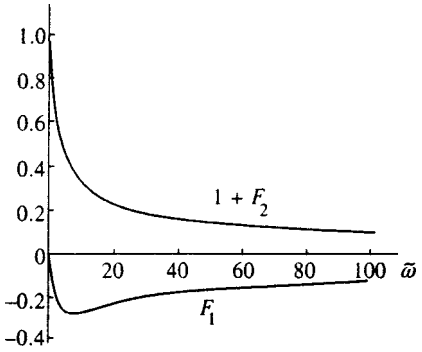

FIG. 1.

The absence of a correction to the Frank elastic constant K_{11} in the static case is a consequence of the invariance of the free energy (7) with respect to equal rotations of the director and the perpendicular smectic layers. The frequency dispersion in K_{11} arises because at frequencies on the order of the reciprocal relaxation time of the smectic order parameter, the smectic structure formed by fluctuations has no time to reorient itself perpendicular to the rapidly changing local director, which results, in the dynamics, in a loss of rotational invariance of the free energy and to the emergence of a fluctuation correction to K_{11}.

3. PROPAGATION OF SHEAR WAVES IN ORIENTED NEMATIC LIQUID CRYSTALS

It can be expected that when the the local director is fairly close to the average direction of preferred orientation, when the direction variation rate is on the order of or exceeds the reciprocal relaxation time of the smectic order parameter, fluctuation-emerging smectic layers are perpendicular to the average direction of the director, since they have no time to reorient themselves in accordance with the changes in the local director. Under such deformations of the director field, the free energy (7) does not remain constant even if the director variations are spatially homogeneous. This generates a restoring force, which opposes deviations (including spatially homogeneous deviations) of the director field from the direction $\mathbf{n}_0$. This relaxation mechanism may strongly affect the natural modes of nematics, particularly the nature of propagation of shear waves at ultrasonic frequencies.

Before estimating the size of this restoring force, we note that we are dealing with phenomena at temperatures for which the fluctuations are relatively weak. More precisely, the fluctuation correction to the specific heat,[11]

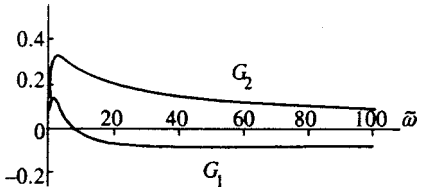

FIG. 2.

$$\Delta C_{\text{fl}}=\frac{k_B T^2 M_T\sqrt{M_V}}{4\sqrt{2}\,\pi\sqrt{A}}\left(\frac{\partial A}{\partial T}\right)^2, \tag{35}$$

must be small compared to the size of the discontinuity in the specific heat at the transition point,[23]

$$\Delta C_{N-A}=\frac{T}{2B}\left(\frac{\partial A}{\partial T}\right)^2, \tag{36}$$

where B is the coefficient of $|\Psi|^4$ in the expansion (7). Thus, in the mean-field approximation, the required temperature range is determined by the condition

$$\frac{k_B^2 T^2 M_T^2 M_V B^2}{8\pi^2 a_0}\ll\frac{T-T_{N-A}}{T_{N-A}}\ll 1. \tag{37}$$

As noted in Ref. 24, such a temperature interval does indeed exist, and it widens as the triple point is approached along the line of N–A-transitions.

Let us estimate the restoring force acting on the director for rapid small spatially homogeneous deviations of the director from $\mathbf{n}_0$. To this end we study the contribution $\langle\mathbf{h}_{\Psi 0}\rangle$, which does not vanish in the $q=0$ limit, to the molecular field averaged over fluctuations:

$$\langle\mathbf{h}_{\Psi 0}(\mathbf{q},\omega)\rangle=-\frac{\gamma}{\tau_0}\,\delta\mathbf{n}(\mathbf{q},\omega), \tag{38}$$

where γ is the rotational viscosity coefficient, and τ_0 has the meaning of the relation time of the director with respect to the direction $\mathbf{n}_0$:

$$\tau_0^{-1}=\frac{k_B T q_s^2}{(2\pi)^3\gamma M_T}\left(\int\frac{d\mathbf{q}'}{\chi^{-1}(\mathbf{q}')}-\frac{2b}{M_T}\int\frac{d\mathbf{q}'\,(\mathbf{q}'_x)^2}{\chi^{-1}(\mathbf{q}')(-i\omega+2b\chi^{-1}(\mathbf{q}'))}\right). \tag{39}$$

First we note that τ_0^{-1} is weakly dependent on frequency and is of the same order at zero and infinite frequencies. For an order-of-magnitude estimate of τ_0^{-1} we set ω to zero (for the sake of simplicity) and ignore the difference between $\xi_\parallel$ and $\xi_\perp$. Moreover, we note that τ_0^{-1} has no critical temperature dependence. Equation (39) shows that the Ornstein–Zernike approximation for the correlation function is insufficient for the convergence of the integrals in (39). Hence, using the fact that the system has a characteristic dimension, the distance between the layers in the smectic-A phase, we can integrate in (39) to the upper limit q_m, assuming that $q_m\sim q_s$. Then we arrive at an estimate for τ_0^{-1}:

$$\tau_0^{-1}=\frac{k_B T q_s^2 q_m}{3\pi^2\gamma}\left(1-\frac{1}{1+q_m^2\xi_\perp^2}\right). \tag{40}$$

We see that near the N–A-transitionpoint the relaxation time of homogeneous distortions of the director field becomes finite. Using the values of parameters typical of liquid crystals,[15,25–27]

$$\xi_\perp \sim 2\times 10^{-8}\left[\frac{T-T_{N-A}}{T_{N-A}}\right]^{-0.5}\ \text{cm},$$

$$\tau_\Psi \sim 10^{-9}\left[\frac{T-T_{N-A}}{T_{N-A}}\right]^{-1}\ \text{s}, \quad q_m \sim q_s \sim 2\times 10^7\ \text{cm},$$

we find that $\tau_0 \sim 10^{-(7+8)}$ s. For the director variation rate the following condition must hold: $\omega \geqslant \tau_\Psi^{-1}$.

Let us find the changes in the spectrum of the shear modes of NLC. If by v_t and n_t we denote, respectively, the velocity and director components that lie in the xy plane and are normal to the wave vector **q**, and by v_l and n_l the components the lie in the xy plane and the plane containing the vectors $\mathbf{n}_0$ and **q**, then the equations for v_t and n_t separate from the other equations in the system of the equation of motion for the NLC.[28]. We write these two equations in the linear approximation with allowance for the term (38):

$$\left[\omega + i\left(\frac{K_t q^2}{\gamma}+\frac{1}{\tau_0}\right)\right] n_t + \frac{\lambda+1}{2} q_z v_t = 0, \tag{41}$$

$$\left(\omega + i\frac{\eta_t q^2}{\rho}\right) v_t + \frac{(\lambda+1)\gamma}{2\rho}\left(\frac{K_t q^2}{\gamma}+\frac{1}{\tau_0}\right) q_z n_t = 0. \tag{42}$$

Here we have used the same notation for the material constants of the NLC as in Refs. 28 and 29, and K_t and η_t are given by the following relationships:

$$K_t = K_{22}\frac{q_\perp^2}{q^2} + K_{33}\frac{q_z^2}{q^2}, \tag{43}$$

$$\eta_t = \eta_1 \frac{q_\perp^2}{q^2} + \frac{\eta_3}{2}\frac{q_z^2}{q^2}. \tag{44}$$

We examine the case where the molecular field is mainly determined by the relaxation term, i.e.,

$$\frac{1}{\tau_0} \gg \frac{K_t q^2}{\gamma}. \tag{45}$$

Using the values of parameters typical of nematic liquid crystals, $K_t \sim 10^{-6}$ dyn, $\gamma \sim (0.1-1)$ P, and $\tau_0 \sim (10^{-8} - 10^{-7})$ s, we arrive at an estimate for the inhomogeneity wavelength: $\lambda \gg 2\times 10^{-5}$ cm.

Equating to zero the determinant of the system of Eqs. (41) and (42), we arrive at an equation for the natural frequencies:

$$\omega^2 + i\left(\frac{1}{\tau_0}+\frac{\eta_t q^2}{\rho}\right)\omega - \frac{1}{\tau_0}\left(\frac{\eta_t q^2}{\rho}+\frac{(\lambda+1)^2\gamma q_z^2}{4\rho}\right) = 0. \tag{46}$$

We see that if

$$\frac{1}{\tau_0} \approx \frac{\eta_t q^2}{\rho}, \tag{47}$$

i.e., $l \sim 10^{-4}$ cm, the solution of the dispersion equation (46) takes the form

$$\omega_{1,2} = \pm c_t q - i\frac{1}{\tau_0}, \tag{48}$$

where

$$c_t = \sqrt{\frac{\gamma}{\rho\tau_0}}\,\frac{\lambda+1}{2}\,|\cos\theta|, \tag{49}$$

with θ the angle between the direction of propagation of the shear wave, **q**, and the vector $\mathbf{n}_0$. Combining (48) and (49), we arrive at an estimate for speed and attenuation constant of the first pair of shear waves: $c_t \sim 3\times 10^3$ cm/s and $1/\tau_0 \sim 10^7\ \text{s}^{-1}$ For a wavelength $l \sim 2\pi\sqrt{\eta_t\tau_0/\rho} \sim 10^{-4}$ cm, a shear wave can propagate in a NLC with moderate rotational viscosity ($\gamma \sim 0.1$ P), which corresponds to $1/\tau_0 \sim 10^7\ \text{s}^{-1}$, since the wave's amplitude changes after one period by a factor of $\exp(-l/c_t\tau_0) \approx 0.7$.

From the equations for the components v_z, v_l, and n_l we can easily obtain a dispersion equation for another pair of shear modes if the vibrations are transverse, or $q_z v_z + q_\perp v_l = 0$. The equation is

$$\omega^2 + i\left(\frac{1}{\tau_0}+\frac{\eta_l q^2}{\rho}\right)\omega - \frac{1}{\tau_0}\left(\frac{\eta_l q^2}{\rho}+\frac{\Lambda_l^2\gamma q^2}{\rho}\right) = 0, \tag{50}$$

where

$$\eta_l = (\eta_1+\eta_2-2\eta_4+\eta_5)\frac{q_z^2 q_\perp^2}{q^4} + \frac{\eta_3}{2}\frac{(q_z^2-q_\perp^2)^2}{q^4}, \tag{51}$$

$$\Lambda_l = \frac{\lambda+1}{2}\frac{q_z^2}{q^2} + \frac{1-\lambda}{2}\frac{q_\perp^2}{q^2}. \tag{52}$$

This implies that if

$$\frac{1}{\tau_0} \approx \frac{\eta_l q^2}{\rho}, \tag{53}$$

we can solve the dispersion equation corresponding to a propagating shear wave:

$$\omega_{3,4} = \pm c_l q - i\frac{1}{\tau_0}, \tag{54}$$

where

$$c_l = \sqrt{\frac{\gamma}{\rho\tau_0}}\,\frac{\lambda\cos 2\theta + 1}{2}. \tag{55}$$

The propagation speeds c_t and c_l of the shear waves strongly depend on the direction of propagation, and for $\lambda = 1$, i.e., for nematic liquid crystals consisting of rod-shaped molecules, they are almost the same. The shear modes that become propagating modes near the *N–A*-transition point are precursors, so to speak, of second sound existing in smectics *A*.

4. CONCLUSION

In conclusion, let us discuss the possibility of experimentally observing the phenomena we have just discussed. Usually the geometry of experiments conducted to detect the acoustooptic effect is such that optical transmission depends on the Frank elastic constant K_{33} (see Ref. 4). Equation (23) implies that near the *N–A*-transition point, K_{33} strongly depends on the ultrasonic frequency, and by varying this frequency one can detect the related variations in optical transmission and in the dimensions of the dark and bright bands

or rings. As follows from Eq. (21), the Frank elastic constant K_{11} has the most interesting frequency dependence. At low and high frequencies, K_{11} does not depend on how close the system is to the phase transition point, but at ultrasonic frequencies an appreciable decrease in K_{11} should be observed. Unfortunately, no experimental data sensitive to variations in K_{11} at ultrasonic frequencies are available.

As for the shear waves near the *N*–*A*-transition point, their speed is almost ten times smaller than that of second sound in the smectic *A*-phase. Here transverse waves with deflections perpendicular to the plane with $\mathbf{n}_0$ and $\mathbf{q}$ can also propagate in the nematic phase. Note that according to (47) and (53) these waves propagate only when their wavelengths lie within a fairly narrow range. In the smectic *A*-phase such waves are purely damped waves. The frequencies of these waves fall in the megahertz range, so that, because in Eq. (39) we ignore all frequency dependence, formulas (49) and (55) for the speeds of the shear waves can be used only for making estimates. Exact allowance for the frequency dependence in (39), which results in the same qualitative conclusions, gives rise to extremely complicated expressions, which are omitted from this paper.

ACKNOWLEDGMENTS

The author would like to express his gratitude to A. Yu. Val'kov and V. P. Romanov for their interest in the work and for fruitful discussions.

APPENDIX

In Eq. (19) for $\langle \mathbf{h}^{(1)}(\mathbf{q},\omega)\rangle$ we leave only the terms that contain factors $\sim q^2$. If we denote their contribution to the averaged molecular field by $\langle \delta\mathbf{h}_\Psi\rangle$, then

$$\langle \delta\mathbf{h}_\Psi\rangle = \frac{bk_BTq_s^2}{M_T^2(2\pi)^3}\int d\mathbf{q}'\frac{\mathbf{q}'_\perp(\mathbf{q}'_\perp\cdot\delta\mathbf{n}(\mathbf{q},\omega))}{[-i\omega+2b\chi^{-1}(\mathbf{q}')]\chi^{-1}(\mathbf{q}')}$$
$$\times\left\{-\left(\frac{\chi(\mathbf{q}')}{2}+b[-i\omega+2b\chi^{-1}(\mathbf{q}')]^{-1}\right)\right.$$
$$\left.\times\left(\frac{q_\perp^2}{2M_T}+\frac{q_z^2}{2M_V}\right)+2\chi^2(\mathbf{q}')\left[\frac{\mathbf{q}_\perp\cdot\mathbf{q}'_\perp}{2M_T}+\frac{q_zq'_z}{2M_V}\right]^2\right\}. \tag{A1}$$

Changing the integration variables,

$$\frac{\mathbf{q}'_\perp}{\sqrt{2M_TA}}=\mathbf{k}_\perp,\quad \frac{q'_z}{\sqrt{2M_VA}}=k_z, \tag{A2}$$

we obtain

$$\langle \delta\mathbf{h}_\Psi\rangle = \frac{k_BTq_s^2}{(2\pi)^3}\sqrt{\frac{2M_V}{A}}\left\{-\left(\frac{q_\perp^2}{2M_T}+\frac{q_z^2}{2M_V}\right)\right.$$
$$\times\int d\mathbf{k}\frac{\mathbf{k}_\perp(\mathbf{l}_\perp\cdot\delta\mathbf{n}(\mathbf{q},\omega))}{(k^2+1-i\tilde\omega)(k^2+1)}$$
$$\times\left(\frac{1}{k^2+1-i\tilde\omega}+\frac{1}{k^2+1}\right)$$
$$+\int d\mathbf{k}\frac{\mathbf{k}_\perp(\mathbf{k}_\perp\cdot\delta\mathbf{n}(\mathbf{q},\omega))}{(k^2+1-i\tilde\omega)(k^2+1)^3}$$
$$\left.\times\left[\frac{2(\mathbf{q}_\perp\cdot\mathbf{k}_\perp)^2}{M_T}+\frac{2(q_zk_z)^2}{M_V}\right]\right\}. \tag{A3}$$

This yields the following expression for the *x*-component of the fluctuation contribution to the molecular field:

$$\langle \delta h_{\Psi x}(\mathbf{q},\omega)\rangle = -\frac{k_BTq_s^2}{\pi^2}\sqrt{\frac{2M_V}{A}}$$
$$\times\left[\left(\frac{I_1(\tilde\omega)+I_2(\tilde\omega)}{12M_T}-\frac{I_3(\tilde\omega)}{5M_T}\right)(q_x^2+q_y^2)\right.$$
$$\left.+\left(\frac{I_1(\tilde\omega)+I_2(\tilde\omega)}{12M_T}-\frac{I_3(\tilde\omega)}{15M_V}\right)q_z^2\right]\delta n_x(\mathbf{q},\omega)$$
$$+\frac{k_BTq_s^2}{\pi^2}\sqrt{\frac{2M_V}{A}}\frac{2I_3(\tilde\omega)}{15M_T}q_xq_y\,\delta n_y(\mathbf{q},\omega), \tag{A4}$$

where the following notation has been introduced:

$$I_1(\tilde\omega)=\int_0^\infty\frac{k^4\,dk}{(k^2+1-i\tilde\omega)(k^2+1)^2}$$
$$=\frac{\pi}{2\tilde\omega^2}[1+\tilde\omega H_1(\tilde\omega)-H_2(\tilde\omega)]$$
$$+i\frac{\pi}{2\tilde\omega^2}\left[H_1(\tilde\omega)+\tilde\omega H_2(\tilde\omega)-\frac{3}{2}\tilde\omega\right], \tag{A5}$$

$$I_2(\tilde\omega)=\int_0^\infty\frac{k^4\,dk}{(k^2+1-i\tilde\omega)^2(k^2+1)}$$
$$=\frac{\pi}{2\tilde\omega^2}\left[-1+\frac{1}{2}\tilde\omega H_1(\tilde\omega)+H_2(\tilde\omega)\right]$$
$$+i\frac{\pi}{2\tilde\omega^2}\left[-H_1(\tilde\omega)+\frac{1}{2}\tilde\omega H_2(\tilde\omega)\right], \tag{A6}$$

$$I_3(\tilde\omega)=\int_0^\infty\frac{k^6\,dk}{(k^2+1-i\tilde\omega)(k^2+1)^3}$$
$$=\frac{\pi}{2\tilde\omega^3}\left[\frac{5}{2}\tilde\omega-(1-\tilde\omega^2)H_1(\tilde\omega)-2\tilde\omega H_2(\tilde\omega)\right]$$
$$+i\frac{\pi}{2\tilde\omega^3}\left[1-\frac{15}{8}\tilde\omega^2+2\tilde\omega H_1(\tilde\omega)\right.$$
$$\left.-(1-\tilde\omega^2)H_2(\tilde\omega)\right], \tag{A7}$$

$$H_1(\tilde\omega)=\frac{1}{\sqrt{2}}\sqrt{\sqrt{1+\tilde\omega^2}-1}, \tag{A8}$$

$$H_2(\tilde{\omega}) = \frac{1}{\sqrt{2}}\sqrt{\sqrt{1+\tilde{\omega}^2}+1}\,, \tag{A9}$$

and the reduced frequency frequency $\tilde{\omega}$ is determined by Eq. (26). Comparing (A4) with (4), we find the contributions (21)–(23).

*)E-mail: mathem@list.spb.su

[1] P. G. de Gennes, *The Physics of Liquid Crystals*, Oxford Univ. Press, London (1974).
[2] S. Chandrasekhar, *Liquid Crystals*, Cambridge Univ. Press, London (1977).
[3] T. C. Lubensky, J. Chim. Phys. **80**, 31 (1983).
[4] A. P. Kapustin and O. A. Kapustina, *The Acoustics of Liquid Crystals* [in Russian], Nauka, Moscow (1986).
[5] A. Yu. Val'kov, V. P. Romanov, and A. N. Shalaginov, Usp. Fiz. Nauk **164**, 149 (1994) [Phys. Usp. **36**, 139 (1994].
[6] P. G. de Gennes, Solid State Commun. **10**, 753 (1972).
[7] F. Jahnig and F. Brochard, J. Phys. (Paris) **35**, 301 (1974).
[8] B. S. Andereck and B. R. Patton, Phys. Rev. E **49**, 1393 (1994).
[9] Jing-huei Chen and T. C. Lubensky, Phys. Rev. A **14**, 1202 (1976).
[10] W. L. McMillan, Phys. Rev. A **9**, 1720 (1974).
[11] F. Kiry and P. Martinoty, J. Phys. (Paris) **39**, 1019 (1978).
[12] J. Swift and B. J. Mulvaney, J. Phys. (Paris) Lett. **40**, L287 (1979).
[13] Kh. A. Hossain, J. Swift, J.-H. Chen, and T. C. Lubensky, Phys. Rev. B **19**, 432 (1972).
[14] J. Swift and B. J. Mulvaney, Phys. Rev. B **22**, 4523 (1980).
[15] V. P. Romanov and S. V. Ul'yanov, Phys. Rev. E **55**, 5623 (1977).
[16] S. Bhattcharaya, B. K. Sarma, and J. B. Ketterson, Phys. Rev. Lett. **40**, 1582 (1978).
[17] S. Bhattcharaya, B. K. Sarma, and J. B. Ketterson, Phys. Rev. B **23**, 2397 (1981).
[18] S. V. Pasechnik and V. A. Balandin, Zh. Éksp. Teor. Fiz. **83**, 195 (1982) [Sov. Phys. JETP **56**, 106 (1982)].
[19] Orsay Liquid Crystal Group, J. Chem. Phys. **51**, 816 (1969).
[20] A. P. Levanyuk, Zh. Éksp. Teor. Fiz. **49**, 1304 (1965) [Sov. Phys. JETP **22**, 901 (1966)].
[21] B. S. Andereck and J. Swift, Phys. Rev. A **25**, 1084 (1982).
[22] V. P. Romanov and S. V. Ul'yanov, Akust. Zh. **37**, 386 (1991) [Sov. Phys. JETP **37**, 195 (1991)].
[23] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Part 1, 3rd ed., Pergamon Press, Oxford (1980).
[24] M. A. Anisimov, *Critical Phenomena in Liquids and Liquid Crystals*, Gordon & Breach, Philadelphia (1991).
[25] I. Haller, H. Higgins, H. R. Liliental, and T. R. McGuire, J. Chem. Phys. **77**, 950 (1973).
[26] E. Kuss, Mol. Cryst. Liq. Cryst. **47**, 71 (1978).
[27] N. I. Alekseev, V. P. Romanov, and S. V. Ul'yanov, Akust. Zh. **34**, 398 (1988) [Sov. Phys. Acoust. **34**, 232 (1988)].
[28] E. I. Kats and V. V. Lebedev, *Fluctuational Effects in the Dynamics of Liquid Crystals*, Springer, New York (1993).
[29] L. D. Landau and E. M. Lifshitz, *Theory of Elasticity*, Pergamon Press, Oxford (1986).

Translated by Eugene Yankovsky

Temperature dependence of surface tension and capillary waves at liquid metal surfaces

V. Kolevzon*)

Institute of Physical Chemistry, University of Karlsruhe, 76128 Karlsruhe, Germany
(Submitted 21 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2034–2042 (December 1998)

The temperature dependence of the surface tension $\gamma(T)$ is treated theoretically and experimentally. The theoretical model based on the Gibbs thermodynamics of a one-component fluid relates $\partial\gamma/\partial T$ to the surface excess entropy density $-\Delta S$. All specific surface effects, namely ordering, capillary waves, and double layer influence the surface entropy, which in turn governs the sign and the magnitude of $\partial\gamma/\partial T$. Experimental data collected at a free Hg surface in the temperature range from 0 °C to 30 °C show that $\partial\gamma/\partial T$ is negative.
 [S1063-7761(98)00912-3]

1. INTRODUCTION

The temperature dependence of the surface tension of liquids is very important in applications such as Marangoni convection and crystal growth. However, little is known about the specific surface forces influencing this dependence in liquid metals.

Our previous work[1] was concerned with the surface tension of mercury as a function of temperature, which was studied using quasielastic light scattering from capillary waves.[2] Although the experimental data show a decrease in surface tension with T, the theoretical explanation of this fact seems to be non trivial. The recent discovery of surface layering in liquid gallium and mercury indicates highly ordered metal structure perpendicular to the surface which is about a few atomic diameters thick.[3,4] This ordering can drastically influence the entropy density profile in the surface zone, reducing the entropy density at the surface compared to the entropy density of the bulk. If this were the case then the surface excess entropy ΔS defined by

$$\Delta S=\int_{-\infty}^{\infty} dz[S(z)-S_b\theta(z)] \tag{1}$$

would be negative.[1] Here $\theta(z)=0$ for $z>0$ and $\theta(z)=1$ for $z\leqslant 0$; $z=0$ denotes the surface position, and the integration is performed from the liquid bulk ($z=-\infty$) to a vapor phase ($z=\infty$). Insofar as the surface tension derivative of a one-component liquid is related to ΔS by[5,6]

$$\Delta S=-\partial\gamma/\partial T, \tag{2}$$

the temperature dependence is expected to be an increasing function of T, as demonstrated at surfaces of normal alkanes over a certain T range.[7] However, the experimental data collected at a Hg-vapor interface contradict this expectation. In order to reconcile experiment and theory, we suggested that a liquid metal surface is a two-component system comprising quasi-free surface electrons and positive ions. In this case $\partial\gamma/\partial T$ is not given by Eq. (2) but depends on the chemical potential of electrons as well.[1] However, the ionic and electronic profiles of any complicated shape should be approximated by the same right-angled profile in order to preserve the electrical neutrality of metal as a whole. Thus, the two-component model is inapplicable to a free metal surface, whose excess entropy should be evaluated in the framework of a one-component model. It will be shown that the effect of surface ordering is not a single contribution on the surface entropy. Other specific surface effects opposing layering tend to increase the entropy of the interface.

2. THEORETICAL ANALYSIS

As noted in Ref. 6, the interface thickness of van-der-Waals liquids is determined by thermal fluctuations, which take the form of thermally excited capillary waves. The r.m.s. amplitude of these waves is usually slightly greater than an atomic diameter.[2] One would thus expect the interface thickness at zero temperature to vanish, and the density profile to be a discontinuous step function dropping from the bulk density ρ_b to zero. In such a situation it is natural to suppose that the location of the Gibbs surface coincides with the step profile at 0 K. The real surface profile varies with T in such a way that the surface excess density given by

$$\Gamma_s=\int_{-\infty}^{\infty} dz[\rho(z)-\theta(z)\rho_b] \tag{3}$$

remains zero for all T.

Consider now a liquid metal comprising two components: free electrons and positive ions. Free electrons behave as a quantum medium: even at zero temperature their energy is nonzero and is usually written in terms of the Fermi energy E_f: $E=E_fN$, where N is the number of electrons. It is instructive to consider the profiles of both components, electronic and ionic, at $T=0$. Due to nonzero wavelength the electronic density profile has a nonzero width, which leads to some redistribution of electrons between bulk and the surface. In other words, some electrons are ejected from the bulk and concentrate on the vapor side of the interface. This charge separation leads to a surface double layer with an electric field $E=Q_s/\epsilon_0$, where Q_s is the surface charge density and ϵ_0 is the dielectric constant. The existence of such a

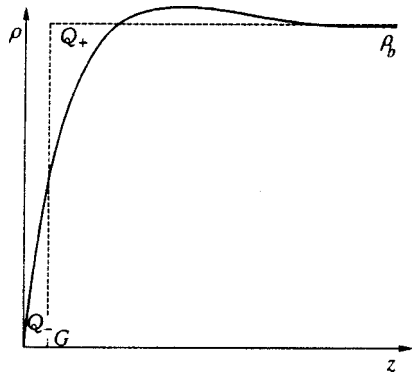

FIG. 1. Schematic variations of the electron density (solid line) and positive ionic jellium represented as a step function which coincides with the Gibbs surface, denoted by *G*. Friedel oscillations of small amplitude are shown in the electron density profile. Surface double layer comprises a negative electrode formed by the area to the left to the step function (labeled with Q_-) and a positive electrode formed by the area to the right of *G* labeled with Q_+.

double layer, suggested by Frenkel,[8] leads to significant changes in the surface interactions. We consider first the conditions that must be satisfied even in the presence of charge separation. The first is electrical neutrality: the total electric charge of a metal must be zero,

$$\int_{-\infty}^{\infty} dz[\rho_i(z)-\rho_e(z)]=0, \tag{4}$$

where ρ_i and ρ_e are the ionic and electronic densities.

The second quantity that must be taken into account is the surface excess density. In general, this quantity is a relative one; it depends on the choice of the Gibbs surface. If this surface coincides with the step ionic profile, the surface excess mass density of electrons can be written via Eq. (3):

$$\Gamma_e=\int_{-\infty}^{\infty} dz[\rho_e(z)-\theta(z)\rho_b].$$

From a comparison with Eq. (4) it is clear that $\Gamma_e=0$ in our choice of Gibbs surface.

We concentrate first on changes in the electronic density with *T* dictated by the surface concentration of electrons in the double layer. This concentration is temperature dependent, although this dependence can be different from the temperature dependence of the bulk electron concentration. This means that the area under the tail on the vapor side (to the left of the Gibbs surface) would vary with *T* (see Fig. 1). If we anticipate that the positive ions do not follow these changes, then the ionic profile will have a shorter tail but a higher amplitude in order to satisfy Eq. (4). However, an ionic density profile of any complicated shape is approximated by a step function that drops from ρ_b to zero at some *z*. If this step function deviates from the step function that approximates the electronic distribution, electrical neutrality of the metal as a whole will be violated. Although the electronic and ionic profiles do not necessarily coincide, they can be approximated by the same right-angled profile as the Gibbs surface in Fig. 1. We have shown that the surface excess density of a metal surface should be zero according to Eqs. (4) and (3).

Unfortunately, no theory describing the surface tension of liquid metals is well established. Numerical simulations of the electronic and ionic density profiles[9] do not provide clear insight into the different forces acting in the surface zone. To the author's knowledge, only one paper treats this problem analytically.[10] Due to its importance for the present context we repeat the main results of this paper.

The electronic density ρ_e at a metal surface varies upon an ionic jellium that approximates the real profile of the ionic density. Following the ideas of Kirkwood and Buff-Bakker (see Ref. 6 for example), the surface tension of any liquid can be expressed via the tensor of anisotropic stresses ($\Pi_{\parallel} \neq \Pi_{\perp}$) and results from the density gradient in the surface zone:

$$\gamma=\int_{-\infty}^{\infty} dz(\Pi_{\perp}-\Pi_{\parallel}), \tag{5}$$

where integration is taken from $-\infty$ to ∞ in order to take into account the effect of free electrons on the vapor side of the interface. The forces acting at the surface at 0 K are nothing more than the quantum pressure of the electron gas and electrostatic force stemming from the potential drop in the double layer. A detailed analysis done by Samojlovich ensures that the long-range electrostatic force $-\rho_e\nabla\phi$ leads to Maxwellian elastic stresses $\Pi_{zz}=-\Pi_{xx}=\epsilon_0E^2/2$ that deform the ionic fluid. This deformation is impossible for the rigid ionic continuum suggested by the jellium model. Therefore, some external pressure $P=\epsilon_0E^2/2$ must be introduced in order to provide the equilibrium of the whole system. This isotropic pressure acting on the topmost layer makes the ions there over-compressed compared to the ions in the bulk. These considerations confirm some increase of the ionic density in the surface layer, which was demonstrated recently on the surface of liquid gallium and mercury. Unfortunately, this conclusion was not drawn by Samojlovich who used a simple monotonic profile to calculate the surface tension. However, his analysis helps to understand surface layering, which is closely related to electrostatic pressure due to a surface double layer.

The temperature behavior of the surface tension of a liquid metal is a most intriguing question. While the tension changes only by a few percent over 100 K, the temperature derivative $\partial\gamma/\partial T$ is extremely sensitive to the density profile. Here Eq. (2) is considered to be a basic thermodynamic equation relating the temperature derivative $\partial\gamma/\partial T$ to the surface excess entropy. It is plausible to suppose that the surface excess entropy comprises different parts:[6] the first is responsible for surface ordering, and the second is pertinent to all fluids, due to capillary waves. The third part is determined by the electric field in the double layer.

We first concentrate on the orientational part. The recent discovery of surface layering in liquid gallium and mercury shows that surface atoms are more ordered than bulk atoms (along the surface normal).[3,4] Hence the orientational part of the entropy in the surface zone might be less than that part of the bulk entropy if the surface atoms exhibit long-range order in the surface plane. Only order along the surface normal was reported at the Hg surface; in-plane surface order has not

been confirmed experimentally.[3,4] Therefore we cannot make any definitive predictions about the orientational part of the entropy in the surface zone.

As we already noted, the surface excess entropy should split into different parts; the orientational part describing the effect of surface layering remains beyond our simplified approach. The other surface effect that can be treated quantitatively is the effect of capillary waves. Following Frenkel's ideas, the surface free energy due to thermal motion in the form of capillary waves is given by[8]

$$F_s = F_{s0} - n_s k_B T \ln \frac{k_B T}{\hbar \tilde{\omega}_c}, \tag{6}$$

where n_s is the number of atoms per unit area, F_{s0} is the surface free energy at 0 K, and $\hbar$ is Planck's constant; $\overline{\omega}$ is the mean frequency of capillary waves, defined as

$$\tilde{\omega}_c = \frac{1}{n_s} \int_0^{q_{\max}} 2\pi q dq \omega(q).$$

Using the usual thermodynamic relation $S = -\partial F/\partial T$, we find the part of the surface entropy associated with capillary waves[8]

$$S_{\text{cap}} = n_s k_B \ln \frac{k_B T}{\hbar \tilde{\omega}_c}. \tag{7}$$

The surface excess entropy density is given by the difference between the densities at the surface S_s and in the bulk S_b: $\Delta S \simeq [S(z) - S_b]\delta z$. The part of the bulk entropy associated with thermal motion is defined by the number of modes of sound waves (or phonons) propagating in the bulk at any $T \neq 0$:

$$S_b = n_b k_B \ln \frac{k_B T}{\hbar \overline{\omega}_p},$$

where $\overline{\omega}_p$ is the mean frequency of bulk phonons.[5] Combining the last two equations, we obtain for the surface excess entropy

$$\Delta S = n_s k_B \ln \frac{\overline{\omega}_p}{\overline{\omega}_c}. \tag{8}$$

We see that ΔS is governed by the ratio of two mean frequencies surface and bulk.

This theory can be applied only to relatively long surface waves whose frequencies are much less than the inverse relaxation time τ_0 typical of each liquid.[8] If the frequency of capillary waves is such that $\omega \geqslant 1/\tau_0$, then vibrations in the liquid propagate as in a solid body. Hence, capillary waves should be replaced by Rayleigh surface waves which propagate at the surface of a liquid or solid body with a velocity $u_s = 0.9\sqrt{G/\rho}$, where G is the shear modulus and ρ is the density. The high-frequency part (with $\omega \geqslant 1/\tau$) makes the main contribution to the spectrum of surface waves.[8] Therefore, the question of the number of modes at the liquid surface reduces to a calculation of these modes at a solid surface. The situation is complicated by the influence of surface waves on the bulk modes.[8] A thorough analysis by Frenkel shows that the surface excess entropy is

$$\Delta S_s = \frac{3}{4}\left(\frac{3\pi}{2}\right)^{1/3} \frac{k_B n_s}{(u_l/u_t)^2[(2+\sqrt{G/Y})^{2/3}]^{2/3}}, \tag{9}$$

where Y is Young's modulus and u_l and u_t are the velocities of longitudinal and transverse sound waves, respectively. The surface tension derivative evaluated using this model is negative and is in good agreement with the tabulated $\partial\gamma/\partial T$ for mercury.

The second mechanism contributing to the surface excess entropy is the electric field in the surface double layer. The impact of this field on the surface free energy is given by $F_s \propto \epsilon_0 E^2 \delta z$. However, this term appears in the free surface energy with a negative sign due to the specific distribution of electric stresses in the double layer. As we discussed above, the effect of electric fields is expressed via the Maxwell stress tensor S_{ij} with components $-S_{zz} = S_{xx} = \epsilon_0 E^2/2$. The change in the surface free energy due to the electric field is[8]

$$F = \int_0^\infty dz (S_{zz} - S_{xx}) = -\int_0^\infty dz \epsilon_0 E^2. \tag{10}$$

It is clear that the surface excess entropy should have a term stemming from electrostatic energy that is specific solely to the surface zone and is zero in the bulk metal. The standard expression $S = -\partial F/\partial T$ yields for the entropy

$$\Delta S_{\text{el}} = \epsilon_0 E^2 \frac{\partial \delta z}{\partial T}, \tag{11}$$

where the derivative is taken at constant surface charge Q_s. This electric part of the surface excess entropy is positive as the thickness of the surface zone increases with T. To the author's knowledge, no analytical predictions exist about the temperature dependence of the intrinsic length scale for the surface electrostatic interactions in liquid metals. Therefore the amplitude of capillary waves will be used as an approximate width of the interface. The r.m.s. value of the interface thickness can be found from the theory of capillary waves:[2]

$$\delta z = \sqrt{\frac{k_B T}{2\pi\gamma} \ln \frac{q_{\min}}{q_{\max}}}. \tag{12}$$

The upper cutoff of capillary waves can be estimated in terms of a molecular size a: $q_{\max} = 1/a$. The minimum wavenumber is usually related to the capillary length: $l_c = \sqrt{\gamma/(\rho g)}$. Bearing this in mind, the part of the surface excess entropy due to the double layer can be written

$$\Delta S = \frac{1}{2}\epsilon_0 \left(\frac{\phi}{\delta z}\right)^2 \sqrt{\frac{k_B}{2\pi\gamma T} \ln \frac{l_c}{1/a}}. \tag{13}$$

Simple estimates based on the values $\delta z \sim 1$ Å and $\phi \sim 1$ V, $\epsilon_0 = 10^{-11} Q^2/(\text{Nm}^2)$, $l_c \sim 1$ mm yield

$$\Delta S = 10^{-4}\sqrt{\ln(10^{-3}/10^{-10})} = 0.4\ \text{mN/(m·K)}.$$

Note that Eq. (13) describes a nonlinear dependence of $\Delta S(T)$ and consequently of $\partial\gamma/\partial T(T)$. It should be stressed that near 0 K the interface thickness in Eq. (11) cannot be given by the r.m.s. amplitude of capillary waves but is equal to the wavelength of quasi-free surface electrons possessing the Fermi energy. Our Eq. (13) does not contain a term $\propto (k_B T/E_f)^2$ typical of the T-dependence of the energy of a

free electron fluid. This is possibly due to our phenomenological approach to the electric field of the double layer.

These estimates are based on the assumption that capillary waves survive up to the upper cutoff of q, i.e., on an atomic scale. From the discussion above we know that at higher frequencies liquid surfaces are similar to solid ones, so capillary waves must be replaced by Rayleigh waves propagating along the surface without dispersion. The surface elastic energy per unit area associated with this mode is

$$F_s = \int_0^\infty dz \frac{1}{2} G(\nabla_s u)^2, \tag{14}$$

where G is the elastic (or shear modulus), u is the displacement, and ∇_s is the differential operator in the surface plane. Then the squared wave amplitude in the q-domain can be written analogously to that of capillary waves,[2]

$$\langle \chi_q \rangle^2 = \frac{k_B T}{A G_s q^2}, \tag{15}$$

where A is the surface area. In order to calculate the r.m.s. displacement of the surface using Eq. (12) one should replace γ with the shear modulus G_s whose magnitude can be estimated using data on the bulk modulus:

$$G_s \approx G_b \delta z = 10^{11} \times 10^{-8} = 10^3 \text{ dyn/cm}.$$

The surface excess entropy density of the Hg surface, calculated using Eqs. (15) and (11), is 0.7 times lower than ΔS estimated from Eq. (12).

3. EXPERIMENTAL METHODS

A liquid mercury surface was prepared by distillation in vacuum (about 10^{-4} Torr) from a batch of 10 ml Hg (99.998%, Merck). The quartz glass distillation apparatus consisted of a compartment filled with Hg in the open air and a condenser (water cooled) connected by a U-tube to a stainless steel capillary mounted in the wall of the working chamber. The chamber had a vacuum flange and an optical window for laser access. Mercury dropped through the capillary to the working chamber and formed a layer about 2 mm deep. The working trough was rectangular with dimensions of 60×60 mm^2, and was machined from stainless steel. Prior to the experiment, the working trough and capillary were cleaned with chromic acid and carefully rinsed in double-distilled water. Distillation was carried out at 200 °C, and about 2 hours were necessary to complete the continuous layer. While no efforts were made to achieve wetting, the Hg surface was relatively flat, as confirmed by the minimal divergence of a reflected beam. Heating or cooling was carried out from below through a copper plate. This plate was heated with a resistive heater or cooled using liquid nitrogen. The temperature was measured with a thermocouple glued to the thin bottom of the working chamber.

Our light scattering technique is described in detail elsewhere.[2,11,12] In brief, a beam from a 5 mW He–Ne laser (TEM_{00}, $\lambda = 632$ nm) fell on the liquid surface. Small-angle scattered light was optically mixed (on a photodetector) with a portion of the original beam, providing all the necessary conditions for optical heterodyning. The output of an avalanche photodiode was modulated at the propagation frequency of a capillary wave with the selected wavenumber q. The spectral representation of the signal was recorded in the frequency domain with a spectrum analyzer. The whole apparatus was placed on an optical table, vibration isolation being provided by four pressurized air cylinders in the legs.

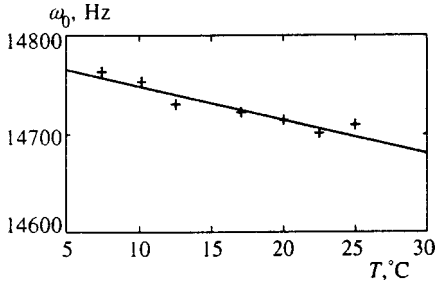

FIG. 2. Temperature variations of peak frequencies of capillary waves of wave number $q = 619$ cm^{-1} at the free surface of mercury. The errors in ω_0 are less than the size of the data points. The solid line is the best-fit solution in the form of Eq. (20).

Capillary waves, present on all liquid surfaces up to the critical point, scatter light mainly at small angles about the reflected beam. The spectrum of the scattered light is the power spectrum of capillary waves, which is approximately Lorentzian.[2] The data were fitted with a theoretical function that incorporates the effects of instrumental broadening.[13,14] The latter arises from illumination of the detector by light scattered by more than one wave-vector $\mathbf{q}$ on the surface. The spread δq in the wave numbers gives a corresponding broadening $\Delta\omega$ in the spectrum. For the Gaussian beam the instrumental function is also a Gaussian.[13,14] The convolution of an ideal Lorentzian and the Gaussian instrumental function of width β yields[14]

$$P(\omega) = \int_{-\infty}^{\infty} \frac{(\Gamma/\beta)\exp[-(\omega-\omega')^2/\beta^2]}{\Gamma^2 + (\omega' - \omega_0)^2} d\omega'. \tag{16}$$

This integral can be evaluated in terms of the complementary error function of the complex argument:[13,15]

$$S(\omega) = A \,\mathrm{Re}\left\{ \exp\left[-\left(\frac{i\Gamma}{\beta} + \frac{\omega - \omega_0}{\beta} \right)^2 \right] \times \mathrm{erfc}\left[-i\left(\frac{i\Gamma}{\beta} + \frac{\omega - \omega_0}{\beta} \right) \right] \right\} + B, \tag{17}$$

where A is the scaling amplitude and B the background. Thus, five properties were extracted from the fit of experimental spectra: frequency ω_0, damping constant Γ, instrumental width β, amplitude A, and background B. In the present context we concentrate only on peak frequencies ω_0.

4. RESULTS

The temperature behavior of peak frequencies ω_0 of capillary waves at the free Hg surface is shown in Fig. 2. To first order, the roots of the dispersion relation describing the propagation of a capillary wave with a particular wavenumber q are[2]

$$\omega_0 = \sqrt{\gamma_0 q^3/\rho}, \tag{18}$$

$$\Gamma = 2\eta q^2/\rho. \tag{19}$$

where γ_0, η, and ρ are the surface tension, bulk viscosity, and density, respectively. Equation (18) serves as a good basis for evaluation of the tension. The data in Fig. 2 are fit by assuming a linear temperature dependence of the surface tension. $\gamma_0 = C_1(T - T_0) + C_2$, where T_0 is the melting point:

$$\omega_0(T) = \sqrt{[C_1(T - T_0) + C_2]q^3/\rho}. \tag{20}$$

The best-fit estimate of the slope C_1 of the temperature dependence is: $d\gamma/dT = C_1 = -0.27 \pm 0.07$ mN/(m·K). The best estimate of C_2 corresponding to the tension at the melting point (-39 °C) is $C_2 = 526 \pm 7$ mN/m, which is slightly greater than the tension (497 mN/m) extrapolated to the melting point using the tabulated slope $d\gamma/dT = -0.2$ mN/(m·K).[16,17]

The light scattering experiment does not directly yield the surface tension; only the peak frequency is directly measurable. In order to obtain additional information on the sign of $\partial\gamma/\partial T$, the following experiment was carried out. A temperature gradient was imposed along the bottom of our cell. In this case a liquid metal flow (usually called thermocapillary convection) driven by the gradient of the surface tension is expected. The surface tension gradient is indirectly related to the liquid velocity v_x via the boundary condition at the free surface, which for the tangential stress is:[18]

$$\frac{\partial\gamma}{\partial T}\nabla T = \eta\frac{\partial v_x}{\partial z}. \tag{21}$$

Thermocapillary convection usually couples to buoyancy-driven convection. However, thermocapillary flow dominates buoyancy convection for a 2-mm deep Hg layer (see Ref. 1). The direction of flow on the liquid surface is expected to be from hot to cold if the temperature derivative of the tension is negative. In the opposite case, the surface flow should be directed toward the hot end. The flow was made visible by means of light scattered from the beam footprint on the surface. We observed the surface flow always to be directed from hot to cold in the middle of the trough, with a characteristic velocity of a few mm/s, which ensures a decreasing dependence $\gamma(T)$.

5. CONCLUSIONS

Surface light scattering from a free mercury surface exhibits the temperature-dependent behavior $\omega(T)$ that corresponds to a decrease in surface tension upon heating. This phenomenon clearly indicates that the surface excess entropy is positive, as it should be in a one-component fluid. The present theoretical treatment ensures that a liquid metal surface is a one-component substance, despite its comprising two components: positive ions and delocalized free electrons. The surface excess entropy ΔS splits into three parts, describing surface layering, capillary waves, and a surface double layer. Since the capillary waves contribution is proportional to $k_B T$, it is expected that this effect is smallest near the melting point. Since surface layering is most pronounced there, ΔS might become negative in some T-range. Unfortunately, it is not possible to estimate the contribution of surface layering to ΔS in the framework of our simple model. However, we hope that surface light scattering is sensitive enough to detect a possible change in the temperature dependence of the surface tension.

*)E-mail: vladi@ipclibml.chemie.uni-karlsruhe.de

[1] V. Kolevzon and G. Pozdniakov, J. Phys.: Condens. Matter **9**, 6815 (1997).
[2] D. Langevin (Ed.), *Light Scattering by Liquid Surfaces and Complementary Techniques*, Marcel Dekker, New York (1992).
[3] M. J. Regan, E. H. Kawamoto, S. Lee, P. S. Pershan, N. Maskil, M. Deutsch, O. M. Magnussen, B. M. Ocko, and L. E. Berman, Phys. Rev. Lett. **75**, 2498 (1995).
[4] O. M. Magnussen, B. M. Ocko, M. J. Regan, K. Penanen, P. S. Pershan, and M. Deutsch, Phys. Rev. Lett. **74**, 4444 (1995).
[5] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pt. J, Pergamon Press, Oxford (1980).
[6] Clive A. Croxton, *Statistical Mechanics of the Liquid Surface*, John Wiley & Sons, New York (1980).
[7] C. J. Hughes and J. C. Earnshaw, Phys. Rev. E **47**, 3485 (1993).
[8] J. Frenkel, *Kinetic Theory of Liquids*, Dover, New York (1955).
[9] M. P. D'Evelin and S. A. Rice, J. Chem. Phys. **78**, 5225 (1982); J. G. Harris, J. Griko, and Ṣ. A. Rice, J. Chem. Phys. **87**, 3069 (1987).
[10] A. Samojlovich, Zh. Éksp. Teor. Fiz., **16**, 135 (1946).
[11] J. C. Earnshaw and R. C. McGivern, J. Phys. D **20**, 82 (1987).
[12] V. Kolevzon and G. Gerbeth, J. Phys. D **29**, 2071 (1996).
[13] S. Hard, Y. Hamnerus, and O. Nilsson, J. Appl. Phys. **47**, 2433 (1976).
[14] D. Byrne and J. C. Earnshaw, J. Phys. D **12**, 1133 (1979).
[15] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions*, Dover, New York (1972).
[16] TAPP, *A Database of Thermochemical and Physical Properties*, E. S. Microware, Inc. (1991).
[17] M. Shimoji, *Liquid Metals: Introduction to the Physics and Chemistry of Metals in the Liquid State*, Academic Press, London (1977).
[18] V. G. Levich, *Physicochemical Hydrodynamics*, Prentice-Hall, Englewood Cliffs, New Jersey (1962).

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

Formation of singularities on the surface of a liquid metal in a strong electric field

N. M. Zubarev*)

Electrophysics Institute, Ural Branch of the Russian Academy of Sciences, 620049 Ekaterinburg, Russia
(Submitted 13 June 1998)
Zh. Éksp. Teor. Fiz. **114**, 2043–2054 (December 1998)

The nonlinear dynamics of the free surface of an ideal conducting liquid in a strong external electric field is studied. It is established that the equations of motion for such a liquid can be solved in the approximation in which the surface deviates from a plane by small angles. This makes it possible to show that on an initially smooth surface for almost any initial conditions points with an infinite curvature corresponding to branch points of the root type can form in a finite time. © *1998 American Institute of Physics.* [S1063-7761(98)01012-9]

1. INTRODUCTION

A flat surface of a conducting liquid placed in a strong external electric field is known[1–3] to become unstable if the field strength E exceeds a critical value $E_c^2=8\pi\sqrt{g\alpha\rho}$, where g is the acceleration of free fall, α is the surface tension, and ρ is the density of the medium. The interaction of the electric field and the charges induced by this field on the surface of the liquid causes surface perturbations to grow rapidly and regions of substantial curvature to form in a finite time.[4,5]. The result may be an increase in the field energy density at the surface, initialization of emission processes, and, finally, vacuum breakdown.[6]. Moreover, there are indications that the liquid phase plays an important role in the initial stages of explosive electron emission.[7]. All this means that one must build a meaningful theoretical model of the nonlinear stages in the development of an instability, a model that can be used to describe the formation of a singular profile of the surface of the medium (a liquid metal in applications).

The present paper studies the nonlinear dynamics of an electrohydrodynamic instability in the limit of a strong electric field, $E\gg E_c$, when both surface tension and gravity can be ignored. The interest in this limit is due, in particular, to the recent discovery of systems with anomalously low critical fields, $E_c\sim 1\ \mathrm{kV\,cm^{-1}}$ (Ref. 8). The nonlinear stages in the development of an instability are studied by perturbation techniques that use series expansions in a small parameter, the angle of the slope of the surface. Of course, the introduction of such a small parameter makes it impossible to describe the formation (observed in experiments) of strong singularities, with corresponding slope angles of order unity. Nevertheless, using the model adopted in this paper, we can show that for almost any initial conditions at the surface of the conducting liquid, it takes only a finite time for points with infinite curvature to form on the surface. Thus, even in the weakly nonlinear stages in the development of a nonlinearity there is the tendency for a singular profile of the liquid surface to form.

The plan of the paper is as follows. In Sec. 2 we derive the main equations of motion, which describe the evolution of the free surface of an ideal conducting liquid in a strong external electric field. In Sec. 3 we use the approximation of small angles characterizing the slope of the surface to build a nonlinear model of the development of an electrohydrodynamic instability. Section 4 is devoted to a study of the dynamics of one-dimensional surface perturbations. Integration of the model equations shows that it takes only a finite time for weak singularities of the root type to form in the system, i.e., singular points at which the curvature of the surface is infinite.

Mathematically, the formation of singularities can be explained by the violation of the analyticity of the complex velocity potential due to the movement of singularities, or branch points, to the boundary. On the whole, such behavior is similar to that of the velocity potential of an ideal liquid in the absence of external forces.[9–11]. In Sec. 5 we use the example of the evolution of single perturbations to show that the formation of singularities occurs before the small-angle condition is violated because of the development of a linear instability (the branch point of the root type agrees with the small-angle approximation). In Sec. 6 we study the behavior of the boundary of a liquid metal under the assumption that self-similarity is retained in a small neighborhood of a singularity in the crossover from one-dimensional perturbations of the surface to arbitrary perturbations. Finally, in Sec. 7 we discuss the role that branch points of the root type play in the evolution of the system.

2. INITIAL EQUATIONS

Consider the motion of a conducting liquid that occupies the region $-\infty<z\leqslant\eta(x,y,t)$ and is subjected to a strong external electric field E. We assume that this liquid is ideal and its motion is vortex-free. Then the potential Φ of the liquid velocity is determined by the time-dependent Bernoulli equation

$$\Phi_t+\frac{(\nabla\Phi)^2}{2}+\frac{p}{\rho}=F(t),$$

where p is the pressure and F is a function of time. Moreover, for potential flow of an incompressible fluid we have $\Delta\Phi=0$. The equations of motion must be augmented by the kinematic condition at the free surface,

$$\eta_t=[\Phi_z-\nabla\eta\cdot\nabla\Phi]|_{z=\eta},$$

by the condition at infinity,

$$\nabla\Phi|_{z\to-\infty}\to 0,$$

and the condition imposed on the pressure at the conductor–vacuum boundary,[3]

$$[8\pi p+(\nabla\varphi)^2]|_{z=\eta}=0,$$

where φ is the potential of the electric field.

The electric potential in the absence of space charges is described by the Laplace equation $\Delta\varphi=0$ together with the conditions that everywhere on the surface of the conductor the potential be the same, $\varphi|_{z=\eta}=0$, and that the field be uniform at infinity, $\varphi|_{z\to\infty}\to -Ez$.

Note that these equations of motion have a Hamiltonian structure and the functions $\eta(x,y,t)$ and $\psi(x,y,t)=\Phi|_{z=\eta}$ are canonically conjugate:[12,13]

$$\frac{\partial\psi}{\partial t}=-\frac{\delta H}{\delta\eta},\quad \frac{\partial\eta}{\partial t}=\frac{\delta H}{\delta\psi},$$

where the Hamiltonian

$$H=\int_{z\leqslant\eta}\frac{(\nabla\Phi)^2}{2}d^3r-\int_{z\geqslant\eta}\frac{(\nabla\varphi)^2}{8\pi\rho}d^3r$$

coincides, to within a constant, with the total energy of the system.

3. THE SMALL-ANGLE APPROXIMATION

Using Green's formulas, we can write the Hamiltonian in the form of a surface integral:

$$H=\int_s\left[\frac{\psi}{2}\frac{\partial\Phi}{\partial n}+\frac{E^2\eta}{8\pi\rho}\frac{\partial\tilde{\varphi}}{\partial n}\right]ds,$$

where $\tilde{\varphi}=(\varphi+Ez)/E$ is the perturbation of the scalar potential, ds is the surface area element, and $\partial/\partial n$ is the normal derivative at the surface s.

From now on we assume $|\nabla\eta|\ll 1$, which corresponds to small surface-slope angles. This allows expanding the normal derivatives in powers series of the canonical variables. Then for the Hamiltonian we have

$$H=\int\frac{\psi}{2}(\hat{T}_+\hat{k}\hat{T}_+^{-1}\psi-\nabla\eta\cdot\hat{T}_+\nabla\hat{T}_+^{-1}\psi)\,d^2r$$
$$-\int\frac{E^2\eta}{8\pi\rho}(\hat{T}_-\hat{k}\hat{T}_-^{-1}\eta+\nabla\eta\cdot\hat{T}_-\nabla\hat{T}_-^{-1}\eta)\,d^2r.$$

Here $\hat{k}$ is the two-dimensional integral operator with a kernel whose Fourier transform is equal to the absolute value of the wave vector:

$$\hat{k}f=-\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{f(x',y')\,dx'\,dy'}{[(x'-x)^2+(y'-y)^2]^{3/2}}.$$

The nonlinear operators $\hat{T}_\pm$ defined as

$$\hat{T}_\pm=\sum_{n=0}^{\infty}\frac{(\pm\eta)^n\hat{k}^n}{n!}$$

act as shift operators (i.e., $f|_{z=\eta}=\hat{T}f|_{z=0}$) for harmonic functions that decay as $z\to\mp\infty$.

If we limit ourselves to second- and third-order terms and introduce scaling

$$t\to\frac{t(4\pi\rho)^{1/2}}{E},\quad \psi\to\frac{\psi E}{(4\pi\rho)^{1/2}},\quad H\to\frac{HE^2}{4\pi\rho},$$

we arrive at an expression for the Hamiltonian:

$$H=\frac{1}{2}\int[\psi\hat{k}\psi-\eta\hat{k}\eta+\eta((\nabla\psi)^2$$
$$-(\hat{k}\psi)^2+(\nabla\eta)^2-(\hat{k}\eta)^2)]\,d^2r. \tag{1}$$

The equations of motion corresponding to this Hamiltonian are

$$g_t+\hat{k}g=\frac{1}{2}[(\hat{k}f)^2-(\nabla f)^2+(\hat{k}g)^2-(\nabla g)^2]$$
$$+\hat{k}[(f-g)\hat{k}f]+\nabla\cdot[(f-g)\nabla f], \tag{2}$$

$$f_t-\hat{k}f=\frac{1}{2}[(\hat{k}f)^2-(\nabla f)^2+(\hat{k}g)^2-(\nabla\hat{k}g)^2]$$
$$+\hat{k}[(g-f)\hat{k}g]+\nabla\cdot[(g-f)\nabla g], \tag{3}$$

where we have changed from the variables η and ψ to the normal variables f and g:

$$f=\frac{\psi+\eta}{2},\quad g=\frac{\psi-\eta}{2}.$$

In the linear approximation, Eq. (2) describes the relaxation of g to zero with a characteristic times $1/|k|$. In this case in the right-hand sides of Eqs. (2) and (3) we can put $g=0$, which means we are examining the perturbation-buildup branch with allowance for a quadratic nonlinearity. This leads us to the following system of equations:

$$g_t+\hat{k}g=\tfrac{1}{2}(\hat{k}f)^2-\tfrac{1}{2}(\nabla f)^2+\hat{k}(f\hat{k}f)+\nabla\cdot(f\nabla f), \tag{4}$$

$$f_t-\hat{k}f=\tfrac{1}{2}(\hat{k}f)^2-\tfrac{1}{2}(\nabla f)^2. \tag{5}$$

Thus, we have shown that studying the dynamics of perturbations of the surface of a conducting medium in a strong electric field in the small-angle approximation amounts to studying the system of equations (4) and (5). What is important about this system is that the nonlinear equation (5) does not contain the function g and that Eq. (4) is linear in g and can easily be solved:

$$g=\frac{1}{2\pi}\int_0^t\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{G(x',y',t')(t-t')\,dx'\,dy'\,dt'}{[(x'-x)^2+(y'-y)^2+(t'-t)^2]^{3/2}}, \tag{6}$$

$$G(x,y,t)=\frac{1}{2}(\hat{k}f)^2-\frac{1}{2}(\nabla f)^2+\hat{k}(f\hat{k}f)+\nabla\cdot(f\nabla f), \tag{7}$$

where we assumed that $g|_{t=0}=0$.

4. FORMATION OF A BRANCH POINT OF THE ROOT TYPE

In the case of one-dimensional perturbations of the surface (we ignore the dependence of all the quantities on y), the integral operator $\hat{k}$ can be expressed in terms of the Hilbert operator $\hat{H}$:

$$\hat{k}=-\frac{\partial}{\partial x}\hat{H}, \quad \hat{H}f=\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{f(x')}{x'-x}dx'.$$

Then the model equations (4) and (5) can written

$$g_t-\hat{H}g_x=\tfrac{1}{2}(\hat{H}f_x)^2-\tfrac{1}{2}(f_x)^2+\hat{H}(f\hat{H}f_x)_x+(ff_x)_x, \tag{8}$$

$$f_t+\hat{H}f_x=\tfrac{1}{2}(\hat{H}f_x)^2-\tfrac{1}{2}(f_x)^2. \tag{9}$$

For further discussions it is convenient to introduce a function that is analytic in the upper half-plane of the complex variable x:

$$\phi=\hat{P}f, \quad \chi=\hat{P}g,$$

where $\hat{P}=(1-i\hat{H})/2$. Since applying the Hilbert operator to a function that is analytic in the upper half-plane amounts to multiplying that function by the unit imaginary number, Eqs. (8) and (9) take the form

$$\mathrm{Re}(\phi_t+i\phi_x+\phi_x^2)=0,$$

$$\mathrm{Re}(\chi_t-i\chi_x+\phi_x^2-2\hat{P}(\phi\phi_x^*)_x)=0.$$

Thus, the integro-differential equations (8) and (9) can be studied simply by analyzing the inhomogeneous linear equation

$$\chi_t-i\chi_x=-\phi_x^2+2\hat{P}(\phi\phi_x^*)_x \tag{10}$$

and (separately) the nonlinear partial differential equation

$$\phi_t+i\phi_x=-\phi_x^2. \tag{11}$$

For the sake of convenience we introduce a new function, $v=\phi_x$. In terms of this function, Eq. (11) becomes

$$v_t+iv_x=-2vv_x.$$

Note that this equation coincides with the one proposed by Zhdanov and Trubnikov,[14,15] who used it to describe the nonlinear stages in the development of tangential discontinuities in hydrodynamics. More than that, if we replace x by $x+it$, we arrive at the equation derived in Refs. 9–11 as a result of a discussion of the nonlinear dynamics of a free surface of ideal liquid in the absence of external forces, where it describes the temporal evolution of the complex-valued velocity. The solution of this first-order partial differential equation can be found by using the method of characteristics:

$$v=Q(x'), \tag{12}$$

$$x=x'+it+2Q(x')t, \tag{13}$$

where the function Q is determined by the initial conditions $Q(x)=v|_{t=0}$.

Let us show, by analogy with Refs. 9–11, that these relations describe (if we require that v be analytic) the formation of a singularity in a finite time. The problem of finding the explicit solution reduces to analyzing the map $x\to x'$ specified by Eq. (13). Generally, this map ceases to be one-to-one at points where

$$\frac{\partial x}{\partial x'}=1+2Q_{x'}t=0. \tag{14}$$

This relationship (14) specifies a path $x'=x'(t)$ in the complex x' plane. Then the motion of the branch point of the function v is given by

$$x(t)=x'(t)+it+2Q(x'(t))t.$$

At the time t_0 when the branch point reaches the real axis the analyticity of v is violated and the solutions of Eq. (9) become singular.

Let us examine the behavior of the solutions near a singularity. Expanding (12) and (13) in a small neighborhood of the point $t=t_0$, $x=x_0=x(t_0)$, $x'=x_0'=x'(t_0)$, in the leading order we get

$$v=q_0-\frac{\delta x'}{2t_0}, \quad \delta x=i\,\delta t+2q_0\delta t+q''t_0(\delta x')^2,$$

where $q_0=Q(x_0')$, $q''=Q_{x'x'}(x_0')$, $\delta x=x-x_0$, $\delta x'=x'-x_0'$, and $\delta t=t-t_0$.

Excluding $\delta x'$ from these expressions, we obtain

$$v=q_0-\left[\frac{\delta x-i\,\delta t-2q_0\delta t}{4q''t_0^3}\right]^2. \tag{15}$$

The derivative of this expression with respect to x is

$$\phi_{xx}\equiv v_x=-[16q''t_0^3 2(\delta x-i\,\delta t-2q_0\delta t)]^{-1/2}, \tag{16}$$

which shows that ϕ_{xx} behaves in a self-similar manner ($\delta x\sim\delta t$) and becomes infinite as $\delta t\to 0$.

As for the complex-valued function χ, the equation that determines its temporal dynamics (Eq. (10)) can be integrated by the method of characteristics (see Eqs. (6) and (7)). Taking the initial condition in the form $\chi|_{t=0}=0$ yields

$$\chi=\int_0^t D(x+it-it',t')\,dt', \quad D(x,t)=-\phi_x^2+2\hat{P}(\phi\phi_x^*)_x.$$

Inserting (15) into this expression, we see than near the singularity the derivative χ_{xx} can be expressed in terms of ϕ_{xx}:

$$\chi_{xx}=\frac{q_0^*-q_0}{q_0+i}\phi_{xx}.$$

This means that the analyticity of χ_{xx} is violated at time $t=t_0$.

How does the surface of the liquid metal behave at the time when the singularities develop in the solutions of Eqs. (10) and (11)? Allowing for the fact that $\eta=f-g$, we find that the surface curvature

$$K=\eta_{xx}(1+\eta_x^2)^{-3/2}$$

is specified, to within a quadratic nonlinearity, by the expression

$$K\approx\eta_{xx}=2\,\mathrm{Re}(\phi_{xx}-\chi_{xx}).$$

Substituting the expressions for ϕ_{xx} and χ_{xx} found earlier, we find that in a small neighborhood of the singular point

$$K \approx 2\,\mathrm{Re}\left[1-\frac{q_0^*-q_0}{q_0+i}\right]\phi_{xx}. \tag{17}$$

Since ϕ_{xx} is given by (16), we have

$$K|_{x=x_0} \propto |\delta t|^{-1/2}, \quad K|_{t=t_0} \propto |\delta x|^{-1/2},$$

i.e., it takes a finite time for a singularity of root type (branch point) to form at the surface, and the curvature of the surface of the liquid at this point is infinite.

To conclude this section we note that since we have $\psi = f+g$, near the singularitya relationship holds for the complex-valued potential of the liquid flow, $\Psi \equiv 2\hat{P}\psi$:

$$\Psi_{xx} = 2(\phi_{xx}+\chi_{xx}) \approx 2\,\frac{q_0^*+i}{q_0+i}\,\phi_{xx},$$

i.e., the first derivative of the complex-valued velocity also exhibits singular behavior as $\delta t \to 0$. This means that, as in Refs. 9–11, the formation of singularities can be interpreted as the result of violation of the analyticity of the complex-valued potential due to the movement of the singularities of the potential to the boundary.

5. EVOLUTION OF A SINGLE PERTURBATION

We use a simple example to show that at the time when a singularity in the solutions of Eqs. (8) and (9) develops the applicability conditions for our model are met.

We take the initial perturbation in the form

$$f|_{t=0} = -\varepsilon a^m \hat{k}^{m-1} \ln(x^2+a^2), \tag{18}$$

where m is a positive integer, and the parameters a and ε take positive values $a>0$ and $\varepsilon>0$. This situation corresponds to a one-dimensional single perturbation of the surface symmetric with respect to point $x=0$, at which the surface curvature is negative. The characteristic slope angles of the surface are determined by the parameter ε, which we assume small.

Note that in the linear approximation Eq. (8) becomes

$$f_t + \hat{H} f_x = 0.$$

Its solution with the initial condition (18) is

$$f(x,t) = -\varepsilon a^m \hat{k}^{m-1} \ln(x^2+(a-t)^2),$$

i.e., within the linear model the perturbation grows without limit and becomes infinite at the time $t=a$, which of course violates the applicability conditions for this model.

Will introducing nonlinearity into model permit a singularity to develop in the solution before the condition $|\eta_x| \approx |f_x| \ll 1$ breaks down? (The branch-point nature of this singularity agrees with the condition that the angles be small.) To answer this question, we will examine the evolution of the perturbation (18) according to the nonlinear equation (9).

The symmetry of (18) implies that the singularity forms at point $x=0$. Then from (13) and (14) it follows that the time t_0 at which the singularity develops can be found by solving the following equations simultaneously:

$$x_0' + i t_0 + 2Q(x_0')t_0 = 0, \quad 1 + 2Q_{x'}(x_0')t_0 = 0,$$

where the function Q corresponding to (18) has a pole of order m at the point $x' = -ia$:

$$Q(x') = i\varepsilon(m-1)! \left(\frac{ia}{x'+ia}\right)^m.$$

Expanding in a power series in the small parameter ε, we obtain to leading order the following:

$$t_0 \approx a\left[1-\frac{m+1}{m}(2\varepsilon m!)^{1/(m+1)}\right],$$

$$x_0' \approx -ia[1-(2\varepsilon m!)^{1/(m+1)}].$$

Since in the linear approximation the singularity is formed at time $t=a$, the above expression for t_0 implies that the nonlinearity accelerates the formation of the singularity (but if $\varepsilon<0$ holds, the nonlinearity delays the onset of the instability).

Plugging the above expression for x_0' into the expression for Q and its second derivative $Q_{x'x'}$, we obtain

$$q'' \approx -\frac{i(m+1)}{2a^2}(2\varepsilon m!)^{-1/(m+1)},$$

$$q_0 \approx \frac{i}{2m}(2\varepsilon m!)^{1/(m+1)}.$$

Thus, for perturbations of the form (18) the parameter q'' is finite. This means that the dynamics of surface perturbations near a singular point is described fairly accurately by Eqs. (15)–(17). As for the parameter q_0, the smallness of ε implies $|q_0| \ll 1$. This is an important result. The important point is that this parameter, as (12) and (15) imply, determines the characteristic angles of slope of the surface by the moment of singularity formation. Then for the derivative η_x at the time of collapse the following estimate holds:

$$|\eta_x| \sim \varepsilon^{1/(m+1)} \ll 1,$$

i.e., the characteristic angles remain small, even though they increased by a factor of $\varepsilon^{-m/(m+1)}$. In this case there is not enough time for the small-angle condition to be violated as a result of the development of a linear instability, and the proposed model (Eqs. (8) and (9) is closed in the sense that if the initial perturbation meets all the conditions needed for the model to be valid, this property is retained throughout the entire evolution until the time of collapse, t_0.

We now discuss the behavior of a perturbation of the electric field at the conducting surface,

$$\delta E(x,t) \equiv -E - \left.\frac{\partial \varphi}{\partial n}\right|_{z=\eta}$$

near the singularity. Clearly, in the linear approximation the field perturbation is specified by the formula

$$\delta E \approx -E\hat{H} f_x = 2E\,\mathrm{Im}(v).$$

Substituting v of Eq. (15) in this expression, we find that at the singular point,

$$\delta E|_{\delta x=\delta t=0} \approx 2E\,\mathrm{Im}(q_0).$$

Since the parameter q_0 is small, the perturbation of the electric field at the time of singularity formation remain much weaker than the external field (both δE_x and δE_t are singular).

6. SELF-SIMILAR SOLUTIONS IN THE GENERAL CASE

In Sec. 5 we found that $|q_0| \ll 1$ holds in the small-angle approximation. This allows us to ignore the dependence on this parameter in Eqs. (16) and (17). Then, if $q'' \neq 0$, the dynamics of the surface curvature in the one-dimensional case is given by the expression

$$K \approx -\mathrm{Re}[4q'' t_0^3(\delta x - i\,\delta t)]^{-1/2}. \tag{19}$$

In particular, for a perturbation that is symmetric with about the point $x = x_0$ we have

$$K \approx -\frac{1}{\sqrt{8t_0^3|q''|}}\left[\frac{-\delta t + \sqrt{\delta x^2 + \delta t^2}}{\delta x^2 + \delta t^2}\right]^{1/2}.$$

Here the initial conditions determine only the constant factor. Thus, the behavior of the system near a singular point is of a universal nature.

Let us take the particular case $q'' = 0$. Suppose, for instance, that

$$\left.\frac{\partial^2 Q}{\partial x'^2}\right|_{x'=x_0'} = \ldots = \left.\frac{\partial^{n-1} Q}{\partial x'^{n-1}}\right|_{x'=x_0'} = 0,$$

$$q_n \equiv \left.\frac{\partial^n Q}{\partial x'^n}\right|_{x'=x_0'} \neq 0,$$

where $n > 2$. Then, expanding (13) in a power series in $\delta x'$ up to the nth power, in the leading order we get

$$K \approx -\frac{1}{nt_0}\mathrm{Re}\left(\frac{n!}{2t_0 q_n}\right)^{1/n}(\delta x - i\,\delta t)^{1/n-1}. \tag{20}$$

The formulas (19) and (20) show that for an arbitrary one-dimensional perturbation of the surface satisfying the condition $|\eta_x| \ll 1$ the curvature near near the singularity behaves self-similarly:

$$K \approx \frac{1}{|\delta t|^p} h\left(\frac{\delta x}{|\delta t|}\right), \tag{21}$$

where h is an unknown function, and the exponent is given by the expression

$$p = (n-1)/n, \tag{22}$$

with n a positive integer.

Note that (19) and (20) are the exact solutions of the linear integro-differential equation

$$K_t + \hat{H} K_x = 0,$$

which describes the evolution of the surface curvature in the linear approximation, as follows from (9) with allowance for the fact that $K = f_{xx}$ holds in the leading order. For an arbitrary exponent p, Eq. (21) specifies the class of self-similar solutions of this equation. This means that, on the one hand, the dynamics of the surface near a singularity is described by the self-similar solutions of the linearized equations of the model and, on the other, that the presence of a nonlinearity leads to a situation in which of all the possible self-similar solutions only those with rational values of p specified by the condition (22) are realized (from general considerations it follows that $p = 1/2$).

It is therefore natural to assume that in the two-dimensional case, as in the one-dimensional, the solutions in a small neighborhood of the singularity are self-similar:

$$K \approx \frac{1}{|\delta t|^p} h\left(\frac{\delta x}{|\delta t|}, \frac{\delta y}{|\delta t|}\right), \tag{23}$$

where p satisfies the condition (22). A characteristic feature of the weak-nonlinearity approximation in our problem is that the specific form of the dependence of all quantities on the self-similar variables can be treated using the equation

$$K_t = \hat{k} K, \tag{24}$$

whose linearity makes it possible to effectively study the formation of two-dimensional singularities at the surface of a conducting medium. Note that this statement is valid if $p < 1$, which, as condition (22) shows, is met in our case in a natural manner. The point is that at $p = 1$ an expression of the form (23) corresponds to the symmetries of the initial nonlinear equations of motion. This means that near a singularity the contribution of a nonlinearity is comparable to that of the linear terms, and the analysis of the behavior of the surface lies outside the scope of this paper.

Substituting (23) in (24), we arrive at the following integro-differential equation for the unknown function h:

$$\xi h_\xi + \zeta h_\zeta + p h = \hat{k}(\xi, \zeta) h,$$

where $\xi = \delta x/|\delta t|$, and $\zeta = \delta y/|\delta t|$. Since the profile of the surface begin to form at the periphery and only then is propagated to the point $\delta x = \delta y = 0$, at the time of collapse the curvature of the surface in a small neighborhood of the singular point is determined by the asymptotic solutions of this equation as $\xi^2 + \zeta^2 \to \infty$. As can easily be shown, these solutions are described by the partial differential equation

$$\xi h_\xi + \zeta h_\zeta + p h = 0,$$

whose general solution is

$$h = [\xi^2 + \zeta^2]^{-p/2} F(\zeta/\xi),$$

where F an unknown function. Plugging this expression into (23) and introducing polar coordinates,

$$\delta x = r\cos\beta, \quad \delta y = r\sin\beta,$$

we arrive at the following formula for the curvature of the surface near the singular point:

$$K|_{t=t_0} \approx \frac{F(\tan\beta)}{r^p}.$$

We see that we are again dealing with a branch point of the root type.

7. CONCLUSION

Our analysis of the evolution of the boundary of a conducting liquid in a strong electric field within the small-angle approximation has shown that for almost any initial conditions on an initially smooth surface the presence of a nonlinearity gives rise to points at which the curvature of the surface becomes infinite. These points correspond to branch points of the root type. However, the presence of such singularities does not ensure a significant concentration of the electric field near the surface of the conductor and, hence, cannot by itself lead to vacuum breakdown. In this case we may assume that the main role of these branch points in the general evolution of the system is, in time, to generate stronger singularities capable of substantially influencing the emission from liquid metals; in particular, capable of ensuring the conditions needed for the initiation of explosive electron emission. Among such singularities are, for instance, discontinuities in the first derivative of the surface profile, which were observed in experiments.[5,7]. Note that a theoretical study of such singularities lies outside the scope of the small-angle approximation and requires allowing for surface tension. Indeed, the applicability of the model adopted in the present paper is limited to the scales $|\delta x| \gg l$, where the parameter l is the characteristic length on which capillary effects become important. This parameter, on dimensional grounds, is determined by the ratio of surface tension to the electric-field energy density: $l \propto \alpha/E^2$. The stabilizing effect of the surface pressure means that at time t_0 the curvature of the surface is still finite ($K \sim 1/l$) and the profile of the surface is smooth, and the formation of a singular profile begins in the later stages in the development of an instability.

The author would like to express his gratitude to A. M. Iskodol'skiĭ and N. B. Volkov for stimulating discussions and to E. A. Kuznetsov who kindly pointed out Refs. 9–11. This work was supported by the Russian Fund for Fundamental Research (Project 97-02-16177).

*)E-mail: nick@ami.uran.ru

[1] L. Tonks, Phys. Rev. **48**, 562 (1935).
[2] Ya. I. Frenkel, Zh. Éksp. Teor. Fiz. **6**, 347 (1936).
[3] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media*, Pergamon Press, Oxford (1984).
[4] M. D. Gabovich and V. Ya. Poritskiĭ, JETP Lett. **33**, 304 (1981).
[5] A. V. Batrakov, S. A. Popov, and D. I. Proskurovskiĭ, Pis'ma Zh. Tekh. Fiz. **19**(10), 66, 71 (1993) [Tech. Phys. Lett. **19**, 627, 629 (1993)].
[6] L. I. Pranevichyus, I. Yu. Bartashyus, and V. I. Ilgunas, Izv. Vyssh. Uchebn. Zaved. Fiz. No. 4 , 44 (1969).
[7] A. V. Batrakov, Ph. D. thesis in Physics and Mathematics, ISÉ, Siberian Branch of the Russian Academy of Sciences, Tomsk (1997) [in Russian].
[8] A. V. Batrakov, S. A. Popov, and D. I. Proskurovskiĭ, JETP Lett. **63**, 615 (1996).
[9] E. A. Kuznetsov, M. D. Spector, and V. E. Zakharov, Phys. Lett. A **182**, 387 (1993).
[10] E. A. Kuznetsov, M. D. Spector, and V. E. Zakharov, Phys. Rev. E **49**, 1283 (1994).
[11] A. I. D'yachenko, V. E. Zakharov, and E. A. Kuznetsov, Fiz. Plazmy **22**, 916 (1996) [Plasma Phys. Rep. **22**, 829 (1996)].
[12] V. E. Zakharov, Zh. Prikl. Mekh. Tekh. Fiz. No. 2, 86 (1968).
[13] E. A. Kuznetsov and M. D. Spektor, Zh. Éksp. Teor. Fiz. **71**, 262 (1976) [Sov. Phys. JETP **44**, 136 (1976)].
[14] S. K. Zhdanov and B. A. Trubnikov, Zh. Éksp. Teor. Fiz. **94**(8), 104 (1988) [Sov. Phys. JETP **67**, 1575 (1988)].
[15] S. K. Zhdanov and B. A. Trubnikov, *Quasigas Unstable Media* [in Russian], Nauka, Moscow (1991).

Translated by Eugene Yankovsky

Mechanism of structural changes of Si(111) surfaces subjected to low-energy ion pulses during molecular-beam epitaxy

A. V. Dvurechenskiĭ,*) , V. A. Zinov'ev, and V. A. Markov

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

(Submitted 30 December 1997)
Zh. Éksp. Teor. Fiz. **114**, 2055–2064 (December 1998)

Reflected high-energy electron diffraction (RHEED) and detection of the intensity oscillations of the specular reflection have been used to investigate morphological changes in Si(111) associated with the two-dimensional layer-by-layer mechanism of silicon growth from a molecular beam under conditions of pulsed (0.25−1 s) bombardment with low-energy (80−150 eV) Kr ions in the interval of small total radiative fluxes (10^{11}–10^{12} cm^{-2}), for which the density of radiation-generated defects is small in comparison with the surface density of the atoms. After pulsed ion bombardment an increase in the intensity of the specular reflection is observed if the degree of filling of the monolayer satisfies $0.5<\theta<1$. No increase in the intensity occurs during the initial stages of filling of the monolayer. The maximum amplitude increment of the oscillations is reached at $\theta\approx 0.8$. The magnitude of the amplitude increment of the RHEED oscillations increases with temperature up to 400 °C and then falls. At temperatures above 500 °C amplification of the reflection intensity essentially vanishes. Experiments on multiple ion bombardment of each growing layer showed that the magnitude of the amplitude increment of the oscillations decreased as a function of the number of deposited layers (the order of the RHEED oscillation). A mechanism for the observed phenomena is proposed, based on the concept of surface reconstruction by pulsed ion bombardment accompanied by formation of a (7×7) superstructure, which corresponds to a decrease of the activation energy of surface diffusion of the adatoms. Within the framework of the proposed mechanism the results of Monte Carlo modeling agree with the main experimental data.
[S1063-7761(98)01112-3]

1. INTRODUCTION

Low-energy (~100 eV) ion bombardment during molecular-beam epitaxy leads to drastic changes in the surface morphology and microstructure of growing films.[1–3] The mechanisms of such changes have so far not been established. A substantial advance in our understanding of these processes became possible when the continuous beam of accelerated particles was replaced by a pulsed beam during growth of the layers. Thus, the use of a pulsed beam of Ar^+ ions with an energy of 600 eV during continuous deposition of Ag on an Ag(111) surface altered the growth mechanism: growth of three-dimensional islands during Ag molecular-beam epitaxy gave way to two-dimensional layer-by-layer growth if the pulsed ion bombardment occurred with a period corresponding to deposition of one monolayer.[1] In an earlier study of silicon molecular-beam homoepitaxy it was established that pulsed bombardment with Kr^+ ions with an energy of 145 eV reduces the roughness of the surface on which growth takes place preferentially by the two-dimensional layer-by-layer mechanism of layer formation.[4] The present paper reports results of a study of the effect of pulsed low-energy ion bombardment during silicon molecular-beam homoepitaxy.

2. EXPERIMENTAL TECHNIQUE

The experiments were carried out in a superhigh-vacuum chamber with residual gas pressure $<10^{-8}$ Pa. The silicon wafers had orientation (111) within 0.1° limits according to the x-ray diffraction data. The samples were chemically cleaned and heated in a high-vacuum chamber at a temperature of 900 °C. A silicon buffer layer of thickness ~200 nm was then grown on the wafers at a temperature of 610 °C. The wafers were then heated to 770 °C until distinct (7×7) superstructure reflections appeared, which testified to the formation of an atomically clean Si(111) surface. An electron-beam Si evaporation block served as the source of the silicon molecular flux. The flux density was controlled by varying the silicon melt region and amounted to 10^{14} -10^{15} atoms/($cm^2\cdot s$). A flow regulator for SNA-2 gas was built into the superhigh-vacuum chamber and was controlled by an applied external electric field. A system for ionizing the admitted gas and accelerating the ionized particles by applying an accelerating voltage up to 200 V was located inside the chamber. The pulsed gas feed block allowed the ion current pulse to vary from 0.25 s (lower limit, determined by the response time of the system for opening the intake slit) to 1 s (upper limit, defining the conditions of maintaining a high vacuum in the system after pulsed bombardment). The angle of incidence of the ion beam was 54.5°. The ex-

periments were carried out at temperatures of 200–600 °C. A cylinder containing spectrally pure argon was connected to the feed block. The purity of the gas was monitored with the help of a mass analyzer built into the growth chamber. The ion current density was varied within the range $0.1-0.6\ \mu A/cm^2$. For the chosen conditions of pulsed bombardment the total flux varied within the limits $10^{11}-10^{12}\ cm^{-2}$.

To record the surface states *in situ*, we employed reflected high-energy electron diffraction (RHEED) and measured the intensity oscillations of the specular reflection associated with the two-dimensional layer-by-layer mechanism of silicon growth.[5] The period of the intensity oscillations under such conditions is equal to the deposition time of a single monolayer. For the case of a Si(111) surface, a complete monolayer is biatomic in thickness (1 monolayer $=1.57\times10^{15}$ atoms/cm^2).[6] Pulses of Kr ions weer applied at different phases of the intensity oscillations of the probing electron beam. The intensity was measured under Bragg diffraction conditions, which are very sensitive to changes in the surface morphology by virtue of electron scattering from the boundaries of steps, islands, and other surface defects. Under these conditions, a one-to-one correspondence is observed between minimum surface roughness and the maximum intensity of the reflected electron beam.[7]

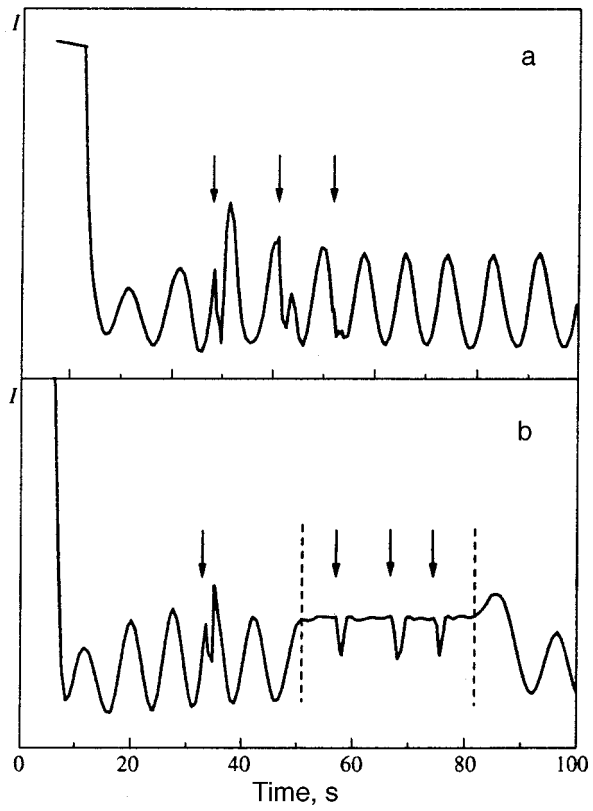

FIG. 1. Variation of the intensity of the specular reflection I under pulsed ion bombardment of the Si(111) surface: a) during molecular-beam epitaxy at $T=400$°C; b) after interruption of the silicon molecular beam at the degree of filling of the monolayer $\theta\sim0.8$ and $T=400$°C. The arrows indicate the onset times of the ion current pulse having current density $0.12\ \mu A/cm^2$ and duration 0.5 s with he energy of the Kr^+ ions in the beam equal to 145 eV. The dashed lines bound the region of interruption of the silicon flux.

3. EXPERIMENTAL RESULTS

Pulsed bombardment with Kr atoms of an initially atomically-smooth surface at temperatures of 200–600 °C reduced the intensity of the specular reflection over a time corresponding to the duration of one pulse, with subsequent recovery of the intensity to a level near its original value. The difference between the initial value of the reflection intensity and its final value after ion bombardment increased as a function of the beam ion energy and decreased as the substrate temperature was raised. This accords with the idea of the introduction of defects into a silicon surface by the ion beam and their subsequent annealing, where the efficiency of annealing, of course, increases with temperature. Experiments on pulsed bombardment of a silicon surface with the gas ionization system switched off showed that the intensity of the specular reflection is maintained at its initial level after pulsed application of the gas flux both for a single pulse and for multiple pulses.

During silicon molecular-beam epitaxy, oscillations of the RHEED specular intensity were observed in the temperature interval 200–550 °C, in line with the known data in the literature.[8] Under these conditions, after deposition of one monolayer the initial (7×7) superstructure changed into a mixture of (5×5) and (7×7) superstructures. Variations in the intensity of the specular reflection during pulsed ion bombardment are shown in Fig. 1(a) at various phases of the RHEED oscillations for a temperature of 400 °C and a deposition rate of about 0.1 monolayer per second. The arrows indicate the switching-on times of the ion current pulse having duration 0.5 s, current density $0.12\ \mu A/cm^2$, and ion energy 145 eV. An abrupt falloff of the intensity of the recorded signal was observed during each ion bombardment pulse. After termination of ion bombardment the variation of the RHEED signal intensity depended strongly on the degree of filling of the growing layer. When the degree of filling θ was in the range $0.5<\theta<1$ the RHEED signal intensified and exceeded the level characteristic of ordinary conditions of growth. In the initial stages of filling of the layer ($\theta<0.5$) amplification of the intensity was not observed. The largest growth of the intensity of the specular reflection due to ion pulse occurred during bombardment before the maximum of the RHEED oscillations, when θ was near 0.8.

Experiments on pulsed ion bombardment of a Si(111) surface were also carried out directly after termination of the molecular flux at various stages of filling of the surface monolayer. It was established that under these conditions the intensity of the specular reflection does not increase for any degree of filling of the monolayer. During ion bombardment the intensity fell and then returned to its initial value during a time comparable to the duration of the pulse [Fig. 1b]. This implies that the Bragg diffraction conditions remain unchanged after ion bombardment and the observed effect is not the result of a change in the conditions of observation.

A comparison of the diffraction pattern under conditions of ordinary growth and growth with ion bombardment made it possible to determine the increase in the fraction of the

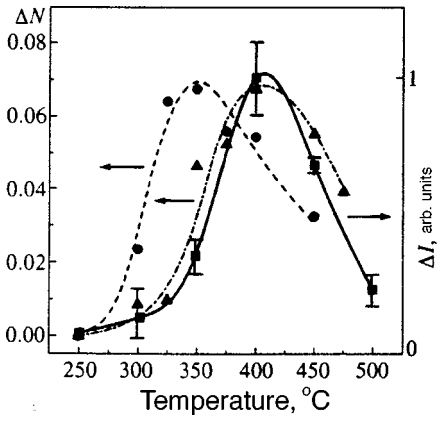

FIG. 2. Temperature dependence of the amplitude increment of growth oscillations of the intensity of the specular reflection ΔI and of the calculated oscillations of the density of steps ΔN (see Fig. 4) due to pulsed ion bombardment for $\theta \sim 0.8$: experimental data (■); results of modeling for $E_1 = 1.1$ eV (●) and for $E_1 = 1.2$ eV (▲).

(7×7) superstructure after pulsed ion bombardment. To establish the connection between this phenomenon and bombardment by a beam of accelerated particles, we carried out analogous experiments with the ionization system switched off and the accelerating voltage absent. No changes in the diffraction pattern were observed in this case.

Studies of the temperature dependence of the amplitude increment of the RHEED oscillations ΔI due to ion bombardment were performed for a fixed degree of filling of the monolayer ($\theta \sim 0.8$) and number of grown-on layers (before the maximum of the third growth oscillation for each substrate temperature). The experiments showed that the value of ΔI increases with temperature up to 400 °C, and then decreases; at temperatures greater than 500 °C amplification of the reflection intensity essentially vanishes (Fig. 2, solid curve).

Experiments on multiple ion bombardment of each growing layer for a fixed degree of filling of that layer (~ 0.8) showed that the magnitude of the amplitude increment of the oscillations depends on the amount of deposited material (Fig. 3). The amplitude increment of the oscillations decreased with increase of the number of deposited layers (the number of the RHEED oscillation).

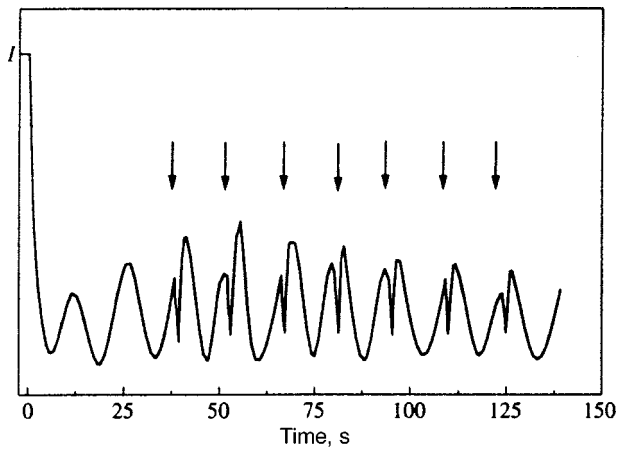

FIG. 3. Oscillations of the electron beam intensity for multiple pulsed ion bombardment of each growing monolayer for $\theta \sim 0.8$ at $T = 400$°C.

4. DESCRIPTION OF MODEL

To describe layer-by-layer growth of silicon from a molecular beam, we utilized a model developed by Vvedenskiĭ and colleagues[7,8] to numerically model epitaxy of group-IV semiconductors. This model assumes that growth is completely controlled by two kinetic processes: deposition of atoms from the molecular beam and surface diffusion of adatoms. Surface migration takes place as a result of diffusion hops of atoms to neighboring unoccupied sites. The frequency of such hops has an activation-like temperature dependence:

$$\nu(E,T) = \nu_0 \exp(-E/kT), \tag{1}$$

where ν_0 is the frequency of vibrations of the surface atoms and E is the activation energy of surface diffusion, which is determined by the binding energy of an atom with its nearest neighbors.

An important property of the initial stage of epitaxial growth of Si(111) is modification of the superstructure of the growing surface during deposition from the molecular beam. For $T \leqslant 550$°C the thermodynamically stable (7×7) superstructure is not able to form completely. As a result, along with the (7×7) superstructure a series of metastable surface phases is formed: (3×3), (5×5), (9×9), and others, which possess excess surface energy relative to (7×7) (Refs. 6 and 9). As a result, the mobility of the atoms on the surface of the epitaxial layer is substantially lower than the mobility on the original surface. To take this aspect of the growth process into account, the activation energy of surface diffusion is modified as follows:

$$E = nE_1 + mE_2 + lE_3. \tag{2}$$

Here nE_1 is the binding energy of the atom with the n first-tier nearest neighbors ($n \leqslant 3$), mE_2 is the binding energy with the m second-tier nearest neighbors ($m \leqslant 6$) in the plane parallel to the surface, E_3 is a correction to the binding energy that takes account of the surface phase, l is the number of atoms in the underlying layer ($1 \leqslant l \leqslant 7$) (Ref. 10). For the initial, atomically clean Si(111) surface it is assumed that $E_3 = 0$. As the first atomic layer is deposited, lE_3 grows due to the transition to the new surface phase, which is characterized by a higher activation energy of surface diffusion.

In the Monte Carlo simulation the analog of the oscillations of the RHEED signal during growth is the variation of the surface density of steps N. In a sense, this quantity characterizes the number of atoms along the perimeter of the islands and vacancy clusters; therefore it is very sensitive to changes in the morphology of the growing surface. It has been shown that under Bragg diffraction conditions the quantity $1-N$ is directly proportional to the intensity of the reflected electron beam.[7,8] In the present paper N is defined as follows:

$$N = \frac{1}{4M} \sum_{i=1}^{\sqrt{M}} \sum_{j=1}^{\sqrt{M}} [|h_{i,j} - h_{i+1,j}| + |h_{i,j} - h_{i,j+1}|], \tag{3}$$

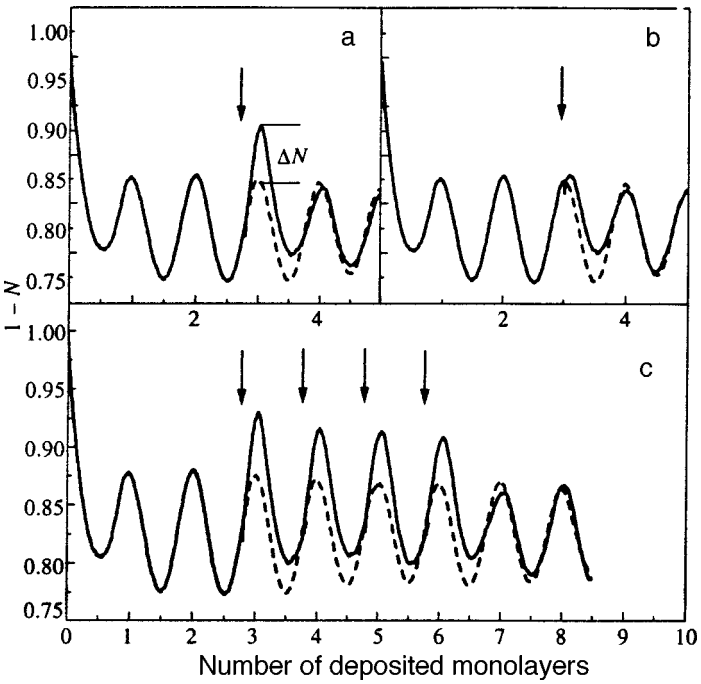

FIG. 4. Calculated oscillations of the density of steps N during growth of silicon from a molecular beam for a single-pulse ion bombardment (a, b): $\theta=0.8$ (*a*), $\theta\sim1$ (b); for multiple pulsed bombardment (c) for fixed degree of filling of the monolayer $\theta=0.8$. Deposition rate 0.1 monolayer/s, $T=400°C$.

where M is the number of surface sites at which the atom can be found and $h_{i,j}$ is the height of the atom above the initial surface. This definition of N allows one to treat differences in the height of the surface steps and therefore differences in their scattering power during RHEED.

To describe the morphological changes on the Si(111) surface resulting from the interaction with low-energy ions we have made use of modeling results based on the molecular dynamics method.[10] According to these results, the incidence of one atom with energy ~200 eV leads on average to the sputtering of one target atom. At the ion impact site a surface vacancy cluster is formed with primarily monoatomic depth (16 vacancies). The target atoms go into adatom sites and are arranged about the vacancy cluster at some distance from its edges (15 adatoms). The molecular dynamics modeling results were used in our model as initial conditions in the Monte Carlo modeling of pulsed ion bombardment during epitaxial growth of a Si(111) surface.

Our model assumed that ion bombardment leads to a modification of the surface phase or to a restructuring of the superstructure of the growing surface. A transition takes place from the metastable surface phase to the thermodynamically more stable (7×7) phase. As a consequence of this, the diffusion coefficient of the adatoms on the Si(111) surface grows after an ion pulse. In the model calculations we made the simplifying assumption that the ion flux density in the experiment was sufficient for reconstruction of the entire surface. Within the framework of the proposed model, this corresponds to a lowering of the activation energy of surface diffusion since the addition to the binding energy E_3 vanishes due to the change in the surface phase during epitaxy. The calculations assumed that pulsed bombardment results in an instantaneous modification of the surface morphology as a result of a combination of the results of the interaction of each ion with the surface. For the total fluxes we used, the probability of the incidence of two ions at one site is essentially zero. We assumed that the reconstruction of the (7×7) superstructure was transformed into another surface phase after the accretion of one monolayer.

5. MODELING RESULTS

The epitaxial growth of Si on Si(111) was modeled using cyclical boundary conditions on a 147×147 two-dimensional grid, whose nodes corresponded to regular sites of the atoms in the crystalline lattice. The numerical parameters were as follows: $E_1=1.1$ (1.2) eV; $E_2=0.2$ eV; $E_3=0.02$ eV. Upon incorporating our proposed mechanism of structural changes into the model of pulsed ion bombardment, we found the results of the numerical experiments to be in qualitative agreement with the experimental data. The density of steps on the surface decreased noticeably some time after pulsed ion bombardment. This effect was most noticeable when the degree of filling of the monolayer was chosen in the range 0.6–0.9. The minimal step density and, consequently, the maximum value of $1-N$ were reached at $\theta=0.8$ [Fig. 4a].

We found that the amplitude increment of the calculated oscillations ΔN (see Fig. 4) during pulsed ion bombardment for $\theta=0.8$ behaves as a function of temperature in a way similar to the amplitude increment of the RHEED oscillations ΔI in the experiment (Fig. 2). The position of the maximum of the calculated dependence (Fig. 2, dashed line) is determined by just one model parameter, E_1, which corresponds to the activation energy of adatom diffusion on the initial, atomically clean surface. For $E_1=1.2$ eV the calculated results are in quite good agreement with experiment. As E_1 is decreased, the maximum of the calculated dependence is shifted toward lower temperatures without any substantial change in the functional dependence of the curve.

For a fixed degree of filling of the monolayer $\theta=0.8$ and fixed substrate temperature, we calculated the changes in the

step density after the action of a single ion pulse on layers at different distances from the initial surface. Comparison of the calculated results implies that pulsed ion bombardment is most effective in smoothing the surface relief (minimizes the step density) under conditions such that the first few monolayers grow. As the number of grown-on layers is increased, the surface smoothing effect of the ion beam is decreased.

Multiple pulsed ion bombardment was modeled for successive switching-on of the ion beam in each successive growing layer at times corresponding to the same degree of filling of the layers [Fig. 4c]. As in the case of a single pulse, the lowest density of steps after ion bombardment is reached during growth of the first few monolayers.

6. DISCUSSION

The increase in the intensity of the diffracted electron beam after pulsed ion bombardment during epitaxy of a Si(111) surface points to a decrease in the surface density of scattering sites, these being the boundaries of islands and vacancy clusters. There are a number of phenomena with which the observed lowering of the surface roughness as a result of low-energy ion bombardment may be connected: 1) sputtering of the material, which corresponds to an effective lowering of the atom flux onto the surface during deposition from the molecular beam; 2) heating of the surface with corresponding changes in the diffusion processes due to the transfer of energy to the atoms from the incident ions; 3) changes in the mechanism of adatom diffusion as a consequence of deposition of an impurity onto the growing surface; 4) changes in the kinetics of growth of the monolayers as a consequence of generation of adatoms and surface vacancy clusters, and also disintegration of the islands forming during growth; 5) reconstruction of the surface, due to the liberation of energy by the accelerated particles.

We analyzed the role of the first two factors in Ref. 4, where we concluded that they do not contribute substantially to changes in the surface morphology for the total ion fluxes used in the experiment.

According to the data in the literature,[11] the deposition of group-III (In, Ga) and group-IV elements (Sn, Pb) during Si epitaxy from molecular beams is capable of increasing the diffusion coefficient of silicon atoms on the surface. These elements, by forming bonds with silicon atoms, in fact alter the mechanism of surface diffusion. For inert gases no effect on the epitaxy process has been noted.

Modeling of epitaxial growth with allowance for the generation of surface vacancies and adatoms by a pulsed ion beam without the introduction of superstructural transitions has shown that no significant change in the surface density of steps takes place for any of the temperatures, deposition rates, or ion current densities used in the experiment. This is probably because the ion flux density is low compared to the atom flux from the molecular beam.

The model calculation gives good agreement with experiment only under the assumption that the surface mobility of the adatoms increases because in the surface phase under bombardment by low-energy ions. Subsequent deposition from a molecular beam leads to efficient filling of the vacancy clusters and a lowering of the density of islands in comparison with epitaxy without pulsed ion bombardment.

The nonmonotonic character of the temperature dependence of the decrease in roughness of the growing surface during pulsed ion bombardment appears to be associated with two competing processes.

1. As the temperature is increased, the diffusion length of the adatoms increases and the relative contribution of surface reconstruction to the change in the surface diffusion coefficient decreases $\sim \exp(lE_3/kT)-1$. This is the reason for the noticeable weakening of the effect for $T>500$°C.

2. Lowering the temperature enhances the formation of islands of the following monolayer before formation of the previous monolayer is complete, which in turn leads to the development of surface relief. Under these conditions, the diffusion length is determined mainly by capture of adatoms at the boundaries of the islands and vacancy clusters, and not by surface reconstruction. As a result, for $T<300$°C, according to the modeling data, the contribution of surface reconstruction during pulsed ion bombardment to the increase in the adatom diffusion length approaches zero.

The observed dependence of the surface smoothing effect under ion bombardment on the number of grown-on monolayers is probably also connected with the development of relief on the growing surface. It is well known that growth from a molecular beam is accompanied by the accumulation of deviations from the planar surface as the thickness of the epitaxial layer increases.[5]

7. CONCLUSION

The studies reported here of morphological changes in a Si(111) surface during two-dimension layer-by-layer growth of silicon from a molecular beam under conditions of pulsed (0.25–1 s) bombardment by low-energy (80–150 eV) Kr^+ ions have made it possible to identify the chief mechanisms of layer growth from ion–molecular beams in the region of small total radiative fluxes (10^{11}–10^{12} cm^{-2}), for which the density of radiation-generated vacancies and adatoms is significantly lower than the surface density of atoms. The idea of the investigative method is to communicate additional energy to the surface atoms during a short time interval by means of an ion pulse, and then to track the dynamics of the subsequent structural changes.

Implementing the given method, we detected a surface smoothing effect during epitaxy of Si(111) for simultaneous pulsed bombardment with low-energy ions and identified the conditions for its manifestation. Comparison of the diffraction patterns obtained during ordinary growth and during growth with ion bombardment allowed us to determine the increase in the fraction of the (7×7) superstructure after pulsed ion bombardment. We have proposed a mechanism and developed a model of structural changes taking place on the surface due to ion bombardment during molecular-beam epitaxy. Monte Carlo modeling gave good agreement with the experimental data (if we admit the assumption of a decrease in the activation energy of adatom surface diffusion as

a result of ion bombardment), which is linked with the structural transformations taking place in conjunction with the increase in the fraction of the (7×7) superstructure.

On the basis of this model we have been been able to explain the temperature and phase dependence of the observed effect, and also the results of experiments on growth of silicon during multiple pulsed ion bombardment.

The investigated class of phenomena in fact further undergirds the development of the method of molecular-beam epitaxy with synchronization of structural transformations by pulsed ion bombardment. This method allows one to obtain sharper boundaries during growth of modulated structures, and also to control the size of islands during heteroepitaxy, e.g., Ge on Si, to create structures with quantum dots.[12]

The authors would like to thank V. A. Kudryavtsev for assistance with the experiments. This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project No. 96-02-19509) and the Federal Program and Subprogram ''Promising Technologies and Devices of Micro- and Nano-Electronics'' (Project No. 211/68/2-1).

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

[1] G. Rosenfeld, R. Servaty, C. Teichert, B. Poelsema, and G. Comsa, Phys. Rev. Lett. **71**, 895 (1993).
[2] P. Bedrossian, J. E. Houston, J. Y. Tsao, E. Chason, and S. T. Picraux, Phys. Rev. Lett. **67**, 124 (1991).
[3] S. W. Park, J. Y. Shim, and H. K. Baik, J. Appl. Phys. **78**, 5993 (1995).
[4] A. V. Dvurechenskiĭ, V. A. Zinov'ev, V. A. Markov, R. Gretzschel, and K.-H. Heinig, JETP Lett. **64**, 742 (1996).
[5] V. A. Markov, O. P. Pchelyakov, L. V. Sokolov *et al.*, Poverkhnost' **4**, 70 (1991).
[6] U. Köhler, J. E. Demuth, and J. Hamers, J. Vac. Sci. Technol. A **7**(4), 2860 (1989).
[7] D. D. Vvedensky and Shaun Clarke, Surf. Sci. **225**, 373 (1990).
[8] G. V. Hansson and M. I. Larsson, Surf. Sci. **321**, 1255 (1994).
[9] M. Horn Von Hoegen, J. Falta, and M. Henzler, Thin Solid Films **183**, 213 (1989).
[10] V. A. Zinoviev, L. N. Aleksandrov, A. V. Dvurechenskii, K.-H. Heinig, and D. Stock, Thin Solid Films **241**, 167 (1994).
[11] B. Voigtländer, A. Zinner, T. Weber, and H. P. Bonzel, Phys. Rev. B **51**, 7583 (1995).
[12] A. I. Yakimov, V. A. Markov, A. V. Dvurechenskiĭ, and O. P. Pchelyakov, JETP Lett. **63**, 444 (1996).

Translated by Paul F. Schippnick

Ferrimagnetic phase transition in the system $Li_{0.5}Fe_{2.5-x}Ga_xO_4$ in the vicinity of the multicritical point of the $x-T$ phase diagram

N. N. Efimova*) and M. B. Ustimenkova

Kharkov State University, 310077 Kharkov, Ukraine
(Submitted 14 February 1998)
Zh. Éksp. Teor. Fiz. **114**, 2065–2077 (December 1998)

In connection with the problem of identifying magnetic states in the vicinity of x_0 (the multicritical point of the $x-T$ diagrams of spin-glass systems) a study has been made of properties that can be exploited to determine the presence of a thermodynamic phase transition at the Curie point T_C and the distinctive features of the transition, specifically, the temperature dependence of the magnetic part of the specific heat $C_m(T)$, the temperature dependence of the low-field magnetization $\sigma_H(T)$, and (with a view toward examining critical behavior in a magnetic field) the magnetization isotherms $\sigma_T(H)$. The investigated object is the system of dilute ferrimagnetic spinels $Li_{0.5}Fe_{2.5-x}Ga_xO_4$, in which every type of magnetic state has spatially inhomogeneous cluster structures. The results obtained for a sample with $x=1.45$ indicate that the classical criteria of a ferrimagnetic second-order phase transition at $T_C=(97\pm2)$ K occur for $x\sim x_0$. The results of similar investigations for a sample with $x=1.6$, which exists in the cluster spin-glass state for $T<T_f=22$ K and in an uncorrelated cluster state of the superparamagnetic type for $T>T_f$, are also given for comparison with the preceding case. © *1998 American Institute of Physics.* [S1063-7761(98)01212-8]

1. INTRODUCTION

Spin glasses have been the object of intensive research for several decades now, and yet the identification of magnetic states in the vicinity of the multicritical point x_0 of the $x-T$ phase diagrams remains as pressing a problem as ever.[1,2] The crux of the problem is essentially the existence of long-range order for $x\sim x_0$: ferromagnetic or ferrimagnetic ($x<x_0$) and spin-glass ($x\geqslant x_0$). The whole debate around this issue has been prompted by the results of neutron-diffraction studies, which are highly conflicting from the standpoint of notions regarding the set of canonical criteria corresponding to, say, the onset of long-range ferrimagnetic order.[1] For example, states having an infinite correlation length r_c (i.e., $r_c^{-1}=0$) and zero spontaneous magnetization $\sigma_s=0$ have been reported, but states with $\sigma\neq0$ and finite r_c^{-1} have likewise been reported.[3–6]

The difficulties surrounding the interpretation of the experimental results, including neutron-diffraction and nuclear giant resonance (NGR) data, are largely attributable to the spatial inhomogeneity of the magnetic states, which have a cluster structure near x_0 (Refs. 1 and 6–8). Clustering effects are especially pronounced for dilute systems exhibiting short-range exchange.[1,6,8,9] Setting details aside, the opinions expressed in the scientific literature in regard to this issue can be divided into two groups. According to some authors,[6,10] for spatially inhomogeneous structures of the cluster type $r_c^{-1}(T)$ has two contributions:

$$r_c^{-1}(T)=r_{\rm cr}^{-1}(T)+r_0^{-1}(\Delta x).$$

The usual ''thermal'' contribution is $r_{\rm cr}^{-1}(T)=0$ for $T\leqslant T_C$ (T_C is the Curie temperature), as in ordinary homogeneous ferrimagnets, but the second contribution is $r_0^{-1}(\Delta x)\neq0$ at all temperatures $T>0$ K. This term is induced by the spatial inhomogeneity of the ferrimagnetic state, and its value is determined by how close the concentration x is to the percolation threshold. Based on these concepts, long-range ferrimagnetic order is established for $T\leqslant T_C$, even though the resultant reciprocal correlation length $r_c^{-1}(T)\neq0$ for $T\leqslant T_C$. Another point of view states, in effect, that macroscopic spontaneous magnetization does not occur in cluster systems near x_0, where $r_c^{-1}(T)\neq0$ holds, but ferrimagnetic order is preserved within the boundaries of large clusters having dimensions on the order of hundreds of angstroms, so that ferrimagnetic-like properties appear in the presence of a field H (Refs. 5 and 11).

Obviously, the question as to whether magnetic order exists in the vicinity of x_0 cannot be resolved unambiguously exclusively by the application of experimental methods highly sensitive to spatial inhomogeneity of the cluster type; the need here is to investigate a broader category of properties. In this paper we discuss the identification of cluster magnetic properties in the vicinity of x_0 from the perspective of establishing the existence or nonexistence of a particular set of classical criteria of thermodynamic phase transition from the ferrimagnetic to the paramagnetic state. The properties chosen for experimental investigation include the temperature dependence of the magnetic contribution to the specific heat $C_m(T)$ for $H=0$, the temperature dependence of the low-field magnetization $\sigma_H(T)$, and, with a view toward studying the critical behavior in a magnetic field, isotherms of the magnetization $\sigma_T(H)$ for $H\leqslant8$ kOe. We believe that such a set of experiments (including $H=0$ and $H\neq0$) can yield, first, independent information about the presence of

long-range ferrimagnetic correlations in the spatially inhomogeneous systems and, in addition, some idea as to anomalies of the properties in a magnetic field.

The chosen object of investigation is a Heisenberg spin-glass system exhibiting short-range interaction, specifically dilute ferrimagnetic spinels $Li_{0.5}Fe_{2.5-x}Ga_xO_4$ with competing negative intersublattice and intrasublattice interactions.[12,13] The multicritical point corresponds to $x_0=1.5$. According to neutron-diffraction data, for $x\geqslant 1.35$ the reciprocal correlation length $r_c^{-1}(T)$ initially decreases as the temperature is lowered, and then in the interval $T_f\leqslant T\leqslant T_1$ it essentially remains constant; T_f is the freezing point of the mixed state, and sufficiently far from x_0 ($x=1.35$) the other limit T_1 coincides with T_C as determined from the temperature dependence of the dynamic susceptibility $\chi_{ac}(T)$ (Ref. 14). The most interesting model object for investigating the existence of long-range ferrimagnetic order in the cluster system is a sample with $x=1.45$; for a step $\Delta x=0.05$ (≈ 1.7 mol. %) in the reentrant-spin-glass region of the $x-T$ diagram ($x<x_0$) this model corresponds to the maximum density of nonmagnetic Ga^{3+} ions at which, in principle, it is still possible for long-range ferrimagnetic order to exist.[13] Moreover, judging from the nature of the problems discussed in the literature (in particular, the influence of a magnetic field), we deemed it advisable from the standpoint of comparison to add another sample with $x=1.6$, where the density of nonmagnetic ions is close to x_0 on the spin-glass side, and transition from the paramagnetic state to the cluster spin-glass state takes place at $T_f=22$ K (Ref. 13).

2. MEASUREMENT PROCEDURE AND SAMPLES

The polycrystalline dilute spinel samples used to investigate the thermal and magnetic properties were prepared by standard ceramic technology from analytical-grade carbonates and oxides of the corresponding metals (5-h reaction in the solid phase in air at $T=1523$ K). The samples were certified as single-phase spinels within the error limits of x-ray analysis; their density was 80–90% of the x-ray value. Lithium ferrite $Li_{0.5}Fe_{2.5}O_4$ and lithium gallate $Li_{0.5}Ga_{2.5}O_4$ form a continuous series of solid solutions with a superlattice of the type 1:3 (Li^+:$Fe^{3+}+Ga^{3+}$) existing in an octahedral sublattice. At Ga^{3+} concentrations (dilutions) $x\sim x_0$ the distribution of Ga^{3+} and Fe^{3+} metal ions among the sublattices is nearly random. According to an estimation of the cation distribution using the magnetization at $T=4.2$ K and $H\to\infty$, the fractions of nonmagnetic Ga^{3+} ions at $x=1.45$ in the tetrahedral and octahedral sublattices, averaged over the sample, are 0.6 (60%) and 0.85 (56%), respectively. However, the spatial distribution of magnetic and nonmagnetic ions in this concentration range is extremely nonuniform (compositional disorder), a condition that is manifested in the development of specific anomalies of the magnetic properties as $x\to x_0$ (Refs. 13 and 14).

A ballistic magnetometer[13] having a sensitivity of 10^{-3} G·cm^3·g^{-1} was used to investigate the magnetization isotherms $\sigma_T(H)$ and polytherms $\sigma_H(T)$. The specific heat was measured within 1.5% limits by means of a vacuum adiabatic calorimeter equipped with an adsorption pump.[15]

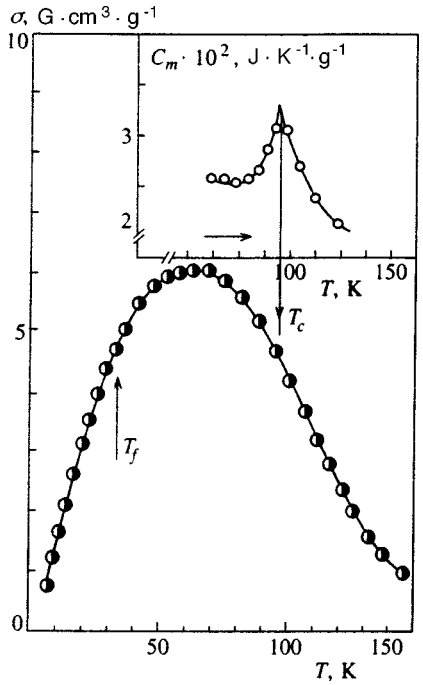

FIG. 1. Temperature dependence of the low-field magnetization $\sigma_{ZFC}(T)$ for a sample with $x=1.45$, $H=50$ Oe. The zero-field cooling (ZFC) regime corresponds to precooling of the sample from 300 K to 4.2 K without a magnetic field. Inset: Fragment of the temperature dependence of the magnetic contribution to the specific heat $C_m(T)$ for the same sample.

The magnetic contribution to the specific heat was isolated by a technique similar to one used earlier.[15] In all the experiments TSU-2 carbon thermometers were used to record the temperature.

3. EXPERIMENTAL RESULTS AND DISCUSSION

Specific Heat and Low-Field Magnetization, $x=1.45$

Figure 1 shows the temperature dependence of the low-field magnetization $\sigma_H(T)$ at $H=50$ Oe and a fragment of the temperature dependence of the magnetic contribution to the specific heat $C_m(T)$ (inset to Fig. 1) for the main object of investigation, i.e., the sample with $x=1.45$. It is evident that the $\sigma_H(T)$ curves spread out considerably in the high-temperature range, with the same kind of behavior observed over a wide range of static fields, beginning at $H=2$ Oe, as well as in the dynamic regime for $\chi_{ac}(T)$ and for $h_0<2$ Oe (Refs. 13 and 14). For $x\geqslant 1.4$ the $\sigma_H(T)$ curves acquire a specific bell shape because of the comparative proximity to the freezing point ($T_f=33$ K for $x=1.45$; Ref. 13) and to the Curie point T_C, as should be the case for $x=1.45$, judging from the behavior of $C_m(T)$. Consequently, the broad maximum of the $\sigma_H(T)$ curves does not mirror any possible transition, either at T_f or at T_C. At the same time, it follows from the data in Fig. 1 that in no way do the $\sigma_H(T)$ curves outwardly disclose a phase transition at the Curie point (in low fields this event is equivalent to the temperature dependence of the initial susceptibility $\chi_0(T)$). In general, behavior of this kind is typical of spatially inhomogeneous systems.[1,5] This result in conjunction with the finite values of $r_c^{-1}(T)$ is often viewed as reason to doubt the existence of a

true thermodynamic transition to the ferrimagnetic state, because even though the second derivatives of the Gibbs free energy for a second-order transition diverge in spatially inhomogeneous systems, they must have singularities.[2,16,17]

In our situation this effect is observed for the specific heat: At $T=97\pm2$ K the $C_m(T)$ curve has an anomaly whose form is typical of a second-order phase transition at T_C. If long-range correlations did not exist in the system anywhere in the crystal as a whole, such an anomaly could be observed only if most of the clusters had values of T_C close to each other and at the same time close to the temperature of the maximum of $C_m(T)$ (Ref. 17). This situation could occur, for example, in an ensemble of isolated particles of identical composition,[18] but not in real dilute systems, where cluster formation is a statistical process caused by compositional disorder, and the cluster subsystem must be characterized by a certain distribution function $f(T_{Ck})$ and, independently, the function $f(M)$, where T_{Ck} is the Curie temperature, and M is the magnetic moment of the cluster.[10,19]

According to existing experimental data, in dilute ferrimagnetic oxides, including Li–Ga spinels with $x>1.3$, intracluster ferrimagnetic ordering is preserved up to temperatures much higher than the Curie point T_C of the samples as a whole.[13,19] Judging from the results of the present study, this result (the existence of $f(T_{Ck})$ and the condition $T_{Ck}>T_C$) is supported both by the spreading of $\sigma_H(T)$ over a wide range of temperatures and by the high specific heat C_m for $T>T_C$ (Fig. 1). The latter consideration suggests that a large number of magnetic degrees of freedom is preserved at $T>T_C$. This conjecture is fully justified if a superparamagnetic state exists at $T>T_C$, because a system of noninteracting clusters can be a source of different types of magnetic perturbations.[20] Thus, the anomaly of the specific heat at $T=T_C$, in our opinion, is conclusive evidence that a correlated magnetic state, i.e., long-range ferrimagnetic order, exists below T_C for the crystal as a whole, and not merely within the boundaries of large clusters.

Critical Behavior in a Magnetic Field, $x=1.45$

Another way to determine the presence of a transition from the ferrimagnetic to the paramagnetic state at T_C is associated with the characteristic features of a second-order transition, specifically the occurrence of critical behavior in the vicinity of T_C for $H>0$ (Refs. 2 and 21–23). Assuming $T_C=97\pm2$ K, we analyze the behavior of the magnetization isotherms $\sigma_T(H)$ for the sample with $x=1.45$ in the vicinity of T_C in terms of their consistency with the magnetic equation of state in the critical region: $(H/\sigma)^{1/\gamma}=A(T-T_C)/T_C+B\sigma^{1/\beta}$, where γ and β are the critical exponents, and A and B are the critical amplitudes. Some of the experimental $\sigma_T(H)$ curves used for this analysis are shown in Fig. 2. As a visual aid, the initial segments of the $\sigma_T(H)$ curves are shown in the inset to Fig. 2. We begin with the mean-field approximation, which satisfactorily describes the critical behavior of homogeneous systems or, in our case, undiluted ferrimagnets.[21] We use the standard procedure for this purpose, where the experimental $\sigma_T(H)$ curves are displayed in

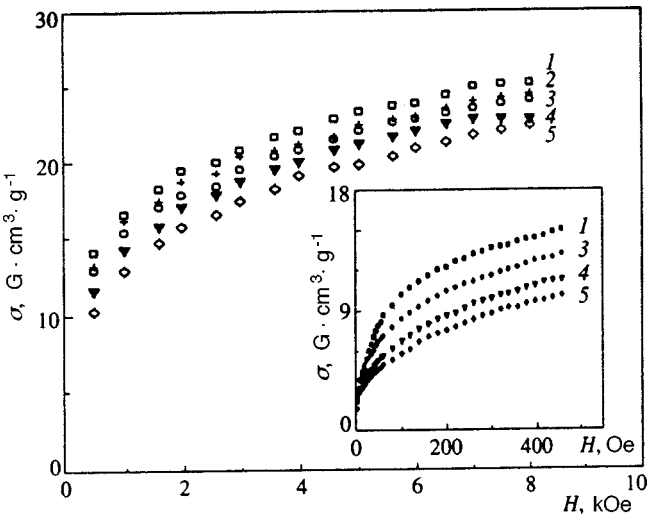

FIG. 2. Magnetization isotherms $\sigma_T(H)$ for a sample with $x=1.45$ at various temperatures: (*1*) 85 K; (*2*) 90 K; (*3*) 95 K; (*4*) 105 K; (*5*) 115 K. Inset: initial segments of the $\sigma_T(H)$ curves for $H\leqslant460$ Oe.

the form of Belov–Arrott graphs, i.e., in H/σ, σ^2 coordinates (Refs. 21–23). This endeavor yields the following results. The experimental $\sigma_T(H)$ curves are straight lines in the magnetic field interval ranging from the fields at which technical magnetization processes terminate to fields $H\approx260-460$ Oe in Belov–Arrott coordinates (Fig. 3). It is evident that the corresponding field interval increases in the paramagnetic region. The isotherm passing through the origin corresponds to $T_C=95$ K, which is in good agreement with previous $C_m(T)$ data. It follows from the results in the inset to Fig. 3 that for the critical isotherm at $T=T_C$ we have $\sigma\sim H^{1/3}$, i.e., the equation $\sigma\sim H^{1/\delta}$ also involves the mean-field value $\delta=3$. Thus, the critical behavior predicted by mean-field theory for homogeneous systems with critical exponents $\gamma=1$, $\beta=0.5$, and $\delta=3$ is observed for the sample with $x=1.45$ in fields $150\text{ Oe}<H<450$ Oe.

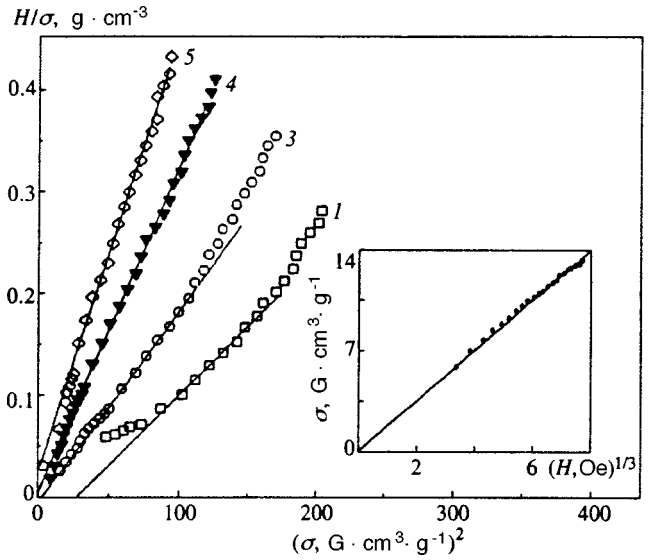

FIG. 3. Magnetization isotherms of Fig. 2 in Belov–Arrott coordinates, $H\leqslant460$ Oe, at various temperatures: (*1*) 85 K; (*3*) 95 K; (*4*) 105 K; (*5*) 115 K. Inset: Graph $\sigma(H^{1/3})$ at $T=95$ K, $H\leqslant260$ Oe.

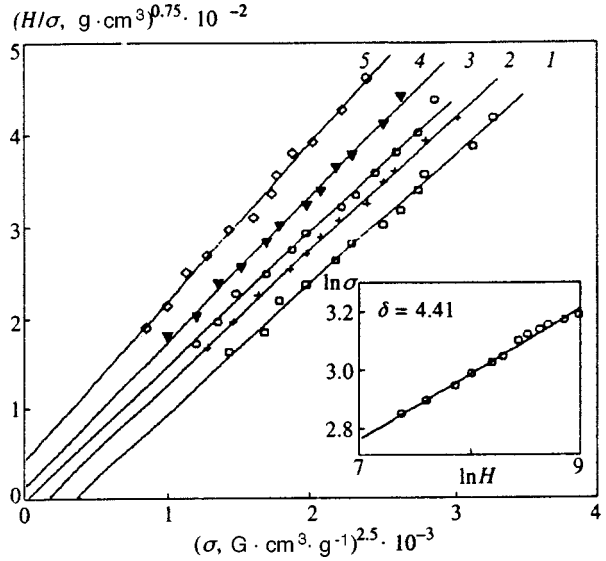

FIG. 4. Magnetization isotherms of Fig. 2 in Belov–Arrott coordinates in fields 1.5 kOe$<H<$8 kOe. The isotherms are numbered as in Fig. 2. Inset: isotherm $\sigma_T(H)$ at $T=95$ K, plotted on log–log scale.

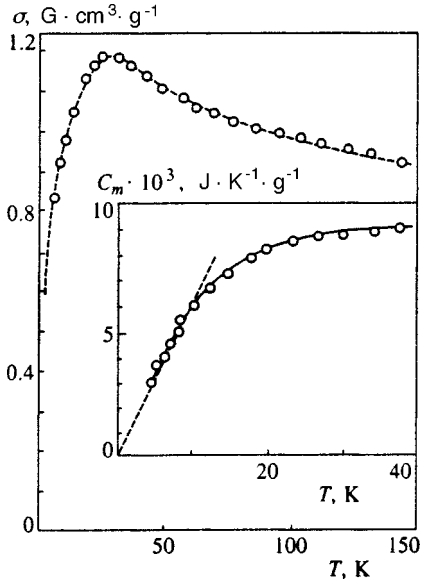

FIG. 5. Temperature dependences of the low-field magnetization $\sigma_{ZFC}(T)$ at $H=50$ Oe and (inset) the magnetic contribution to the specific heat $C_m(T)$ for a spin-glass sample with $x=1.6$.

Outside this field interval, however, the isotherms of Fig. 2 shown in Belov–Arrott coordinates deviate from linear. This behavior, i.e., the possible curvature of the Belov–Arrott graphs as a function of H, has been predicted in a generalization of the mean-field model to spatially inhomogeneous systems.[24–27] The authors of Refs. 25–27 describe these deviations by introducing a function $F(H)$, as yet undefined, in the mean-field equation for homogeneous systems: $H/\sigma=A'+B'\sigma^2+F(H)$, $A'=a'(T-T_C)$; it is associated with fluctuations of the thermodynamic coefficients A' and B'. In another approach widely used in experimental work, spatially inhomogeneous systems are assumed to correspond to a magnetic equation of state in which the critical exponents γ and β differ from the mean-field values.[23,28] We have chosen the second approach to analyze the magnetization isotherms in fields $H>460$ Oe; following Kaul et al.,[23,28] we display the experimental $\sigma_H(T)$ curves in the form of Arrott–Noakes graphs: $(H/\sigma)^{0.75}-\sigma^{2.5}$ (Fig. 4). It is evident that in fields 1.5 kOe$\leqslant H\leqslant$8.0 kOe the experimental isotherms in Arrott–Noakes coordinates obey a linear law, and the isotherm passing through the origin again corresponds to $T=95$ K. In this case, therefore, we have $\gamma=1.33$ and $\beta=0.4$. Converting the critical isotherm at $T=T_C=95$ K to log–log scale, we find $\delta=4.41$ (see the inset to Fig. 4). These values of the critical exponents satisfy the scaling relation[2] $\alpha=2(1-\beta)-\gamma$ for the realistic value $\alpha=-0.13$. We note that the same values of the critical exponents are also typical of another type of spatially inhomogeneous systems, amorphous magnets.[23]

By the same reasoning as in our discussion of the nature of the $C_m(T)$ anomaly we are fully justified in assuming that the observed critical behavior reflects the behavior of the system as a whole rather than intracluster processes. The only special feature that we have observed on the part of second-order transition between spatially inhomogeneous ferrimagnetic and paramagnetic states is a change in the critical behavior in the transition from the weak-field region to the moderate-field region, and we have not encountered any experimental evidence of such behavior. We shall return to this problem below. First, however, it is instructive to compare the behavior of the given $x=1.45$ sample with the spin-glass sample characterized by $x=1.6$.

The $C_m(T)$, $\sigma_H(H)$, and $\sigma_T(H)$ Curves, $x=1.6$

A state of the cluster spin-glass type occurs in Li–Ga spinels with $x>x_0=1.5$ (Ref. 13). Ferrimagnetic ordering exists within the cluster boundaries, both for $T<T_f$ and in the paramagnetic region for $T>T_f$ (Refs. 13 and 29). At $T>T_f$ the thermal energy exceeds the exchange energy between clusters, so that an uncorrelated cluster state of the superparamagnetic type occurs in spin-glass samples with $x>1.5$, and the cluster subsystem is described by certain distribution functions $f(T_{Ck})$ and $f(M)$. The sample with $x=1.6$ is used as a model of this type of state. The behavior of the spin-glass and ferrimagnetic samples is compared using the same set of properties for $x=1.6$ as for $x=1.45$.

Figure 5 shows the temperature dependences of the low-field magnetization $\sigma_H(T)$ and the specific heat $C_m(T)$. Clearly, the cluster inhomogeneity also influences the character of the spin-glass transition at T_f: It emerges abruptly for $T<T_f$ ($T\to 0$ K) and spreads out considerably for $T>T_f$, the low-field magnetization remaining high up to 150 K. The same pattern of behavior as in Fig. 5 is observed in external fields H weaker than those represented in Fig. 5.

The specific heat $C_m(T)$ (see the inset to Fig. 5), as in classical spin glasses,[1] is a linear function of the temperature at $T\to 0$ K and does not have an anomaly at T_f. In the given situation, however, a departure from linearity is already observed for $T<T_f$. The values of C_m are high for $T>T_f$, as is typical of spin glasses, where only a part (~ 0.4) of the total magnetic entropy is tapped for $T\leqslant T_f$ (Ref. 1).

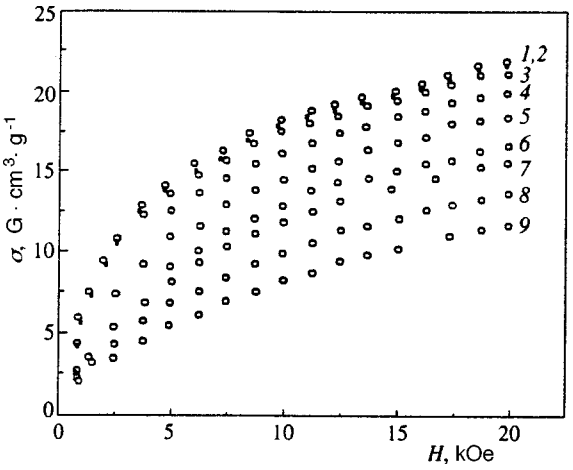

FIG. 6. Magnetization isotherms $\sigma_T(H)$ for a sample with $x=1.6$ at: (*1*) $T=20$ K; (*2*) 4.2 K; (*3*) 40 K; (*4*) 60 K; (*5*) 80 K; (*6*) 100 K; (*7*) 110 K; (*8*) 130 K; (*9*) 150 K.

Despite the special behavior characteristics described above, the occurrence of transition from the paramagnetic to the spin-glass state at T_f for the sample with $x=1.6$, as in the case of other samples of the system $Li_{0.5}Fe_{2.5-x}Ga_xO_4$ ($x>1.5$), is uniquely determined from the onset of long-term logarithmic relaxation of the nonequilibrium (zero-field cooling) magnetization σ_{ZFC} and from the existence of critical behavior lines $T_f(H)$ in a magnetic field; finally, the transition is described by the one-component Edwards–Anderson order parameter q_{EA} (Ref. 13).

Moreover, the presence of ferrimagnetic clusters in the magnetic subsystem appears, at first glance, to leave the magnetization isotherms $\sigma_T(H)$ exactly the same as before (for $x=1.45$), whether for $T<T_f$ or for $T>T_f$ (see Fig. 6). The magnetization is fairly high up to temperatures $T=150$ K, which is well in excess of T_f. From the standpoint of the stated problem it would be interesting to resolve the issue of the proper equation of state for the isotherms in Fig. 6. As a first attempt we have used the same procedure as for $x=1.45$. It follows from the data of Fig. 7 that the magnetization isotherms $\sigma_T(H)$ of the sample with $x=1.6$ in the temperature range $4.2\text{ K}<T<150\text{ K}$ and in the magnetic field range $3.5\text{ kOe}<H<20\text{ kOe}$ are straight lines in Arrott–Noakes coordinates, but none of them crosses the $\sigma^{2.5}$ axis in the positive range. This condition implies the absence of spontaneous magnetization.[21,23,28]

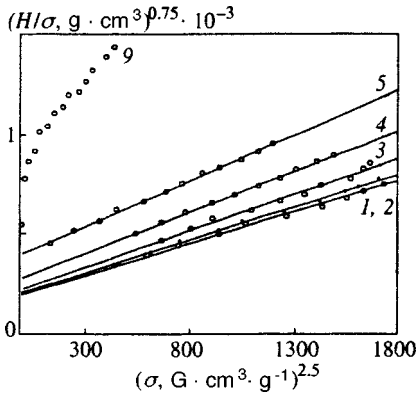

FIG. 7. Magnetization isotherms of Fig. 6 in Arrott–Noakes coordinates. The isotherms are numbered as in Fig. 6.

The results obtained for the sample with $x=1.6$ clearly show that even in fairly strong magnetic fields long-range ferrimagnetic order is not ''restored'' in the crystal if it did not exist in zero field $H=0$. This conclusion applies equally to the range $T<T_f$, where the exchange energy between clusters exceeds the thermal energy, and to the range $T>T_f$. Consequently, the assertion that a cluster ferrimagnet and a superparamagnet (uncorrelated cluster state) exhibit similar behavior in a magnetic field can apply exclusively to the similar forms of the isotherm $\sigma_T(H)$ and, in part, the polytherm $\sigma_H(T)$ of the magnetization. However, even a simple analysis of these dependences indicates a fundamental difference between the correlated ferrimagnetic and uncorrelated superparamagnetic cluster states, manifested in whether or not σ_s and, accordingly, T_C exist.

Model of a Spatially Inhomogeneous Magnetic Structure and Phase Transition at the Curie Point for $x\sim x_0$

In systems exhibiting short-range interaction due to compositional disorder (irregular distribution of magnetic and nonmagnetic ions in the lattice) the exchange interaction acquires spatial inhomogeneity, so that two exchange-coupled subsystems, clusters and a matrix, can be distinguished in the crystal.[10,13,14] The clusters correspond to regions having a high content of magnetic ions with strong exchange maintained between them, so that ferrimagnetic ordering occurs within the boundaries of each cluster. In the matrix, on the contrary, we find a large content of the nonmagnetic component. This competition between exchange interactions not only weakens the exchange, but also incurs frustrated bonds. Judging from existing experimental data,[13] including those in the present study, the type of magnetic ordering of the crystal as a whole depends on the state of the matrix; in particular, the existence of the ferrimagnetic or spin-glass state is obviously dictated by the concentration of frustrated bonds.

Using this model of the cluster ferrimagnetic state, in the example of the sample with $x=1.45$ we analyze the pattern of phase transformation at the Curie point in connection with the ferrimagnet–paramagnet transition in spatially inhomogeneous structures. The breakdown of long-range ferrimagnetic order in the crystal is associated with the disappearance of long-range correlations between spins of the matrix and, hence, between clusters. The value of T_C for the crystal as a whole is determined by the average exchange between the matrix and the clusters and between spins of the matrix. When $T>T_C$ holds, the latter are essentially completely disordered, and ferrimagnetism is preserved in clusters up to $T_{Ck}>T_C$. For $T>T_C$ is obvious that only local second-order phase transitions occur at $T=T_{Ck}$, corresponding to the breakdown of ferrimagnetic ordering of the spins in the clusters. This process takes place over a broad temperature range, which is determined by the distribution function $f(T_{Ck})$, and so it cannot impart to the thermodynamic properties the characteristic features of a macroscopic second-

order phase transition. This conclusion follows directly from our experimental data, including those for the sample with $x=1.6$.

Regardless of whether the critical behavior of spatially inhomogeneous systems is described in terms of changes (from the homogeneous case) in the critical exponents[23,28] or by introducing the function $F(H)$ (Refs. 24–27), in general it is governed primarily by the presence and character of spatial fluctuations of the magnetization and the exchange and also, according to Refs. 24–27, by the magnitude of the external field H. In this light the anomalies observed by us in the critical behavior of the sample with $x=1.45$ could well be indicative of general functional relations that can occur for the right combination of governing factors. In our case these factors are sharp spatial fluctuations of the exchange and magnetization, which result in the formation of the cluster-type magnetic structure discussed above. As a consequence, with regard for the specific features of the magnetization processes in such systems the following pattern of variation of the critical behavior can be envisioned for the sample with $x=1.45$. It follows from general energy considerations (the relation between the thermal and magnetic energies $E_H=\mathbf{M}\cdot\mathbf{H}$, where $\mathbf{M}$ is the magnetic moment of the cluster or the individual ion) and from the behavior of the magnetization isotherms $\sigma_T(H)$ (Fig. 2) that the contributions of the two subsystems (clusters and matrix) to the total susceptibility depend on the field H. If the field range where technical magnetization has essentially vanished is partitioned arbitrarily into three intervals corresponding to weak, moderate, and strong fields, the behavior of $\sigma_T(H)$ in weak fields is governed by clusters having large magnetic moments, and in strong fields it is determined by the matrix spins, because the moments of the clusters are already oriented in the direction of the field. In the interval of moderate fields the behavior of $\sigma_T(H)$ is determined to a greater or lesser degree by the exchange subsystems and, among other things, depends on the distribution function $f(M)$. In this sense (the experimentally recorded response to an external magnetic field) the inhomogeneous system plays the part of a homogeneous system in weak and strong fields.

The cooperative nature of the second-order phase transition at the Curie point T_C, i.e., the participation of both subsystems—clusters and matrix—in the process, can also be demonstrated within the scope of the same approach. This follows directly from the identical values of T_C obtained whether the response of the cluster subsystem (weak fields) or of the spatially inhomogeneous system as a whole ($H=0$, moderate fields) is recorded. We call attention to an alternative point of view concerning this problem in Ref. 23, where it is stated that only a fraction of the total number of spins is involved in second-order phase transitions in amorphous magnets.

We conclude our discussion of the nature of the ferrimagnetic phase transition near x_0 with a brief look at the absence of any kind of anomalies of the temperature dependence of the low-field magnetization $\sigma_H(T)$ and the initial susceptibility $\chi_0(T)$. According to preliminary data, the cause can be found in specific phenomena associated with the cluster subsystem. As a matter of fact, some of the clusters can exist in few-domain or single-domain states rather than in the superparamagnetic state.[18,29] The magnetization process of such clusters does not prevent any changes from taking place in passage through the Curie point of the sample, but their contribution to the magnetization (susceptibility) in the technical magnetization region can be large. As the field is increased, single-domain clusters can enter the superparamagnetic state.[18,30] This event could also account for the experimental fact that when the field H is increased for $T\geqslant T_C$ ($T\geqslant T_f$), a sharper drop in the magnetization is observed than in low fields. The curve representing the dynamic susceptibility χ_{ac} (very low fields) in the vicinity of T_C or T_f can also spread out because of the presence of a weak residual moment in the sample.

4. CONCLUSION

For the case of the dilute ferrimagnetic spinel $Li_{0.5}Fe_{2.5-x}Ga_xO_4$ ($x=1.45$) with short-range interaction we have shown that in the vicinity of the multicritical point ($x_0=1.5$) a ferrimagnetic second-order phase transition takes place at the Curie point, as characterized by standard thermodynamic criteria: An anomaly of the $C_m(T)$ curve, typical of second-order phase transitions, is observed at $T=T_C=(97\pm2)$ K, and in the vicinity of T_C critical behavior is observed in a magnetic field. On the whole, the experimental results argue against a magnetic state model that postulates the preservation of ferrimagnetic ordering strictly within the cluster boundaries for concentrations $x\sim x_0(x<x_0)$. The existence of long-range order in spatially inhomogeneous structures of the cluster type (ferrimagnetic for $x<x_0$ or spin-glass for $x>x_0$) is satisfactorily explained by a model of ferrimagnetic and spin-glass states that takes into account the two exchange-coupled subsystems—clusters and matrix—in the crystal. The crystal acquires long-range correlations as a result of exchange interaction between the matrix and clusters and between matrix spins. The average magnitude of this exchange determines T_C, and its variance determines T_f.

Further analysis of the model object, i.e., the spin-glass sample with $x=1.6$, where the uncorrelated cluster state occurs at $T>T_f$, graphically demonstrates the fallacy of the view that when long-range ferrimagnetic order does not exist in cluster systems at $H=0$, it can be ''restored'' in a sufficiently strong magnetic field.

Inasmuch as the present study has addressed the extreme situation of a system with short-range interaction, where compositional disorder is conducive to exchange having an extremely pronounced spatial inhomogeneity, the results can be extended to other spin-glass systems, including metallic, of the nature ofthe Au–Fe systems discussed in Ref. 11.

*)E-mail: Alexander.V.Vankevich@univer.kharkov.ua

[1] K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).
[2] V. S. Dotsenko, Usp. Fiz. Nauk **163**(6), 1 (1993) [Phys. Usp. **36**, 455 (1993)]; Usp. Fiz. Nauk **165**, 482 (1995).
[3] G. Aeppli, S. M. Shapiro, H. Maletta *et al.*, J. Appl. Phys. **55**, 1628 (1984).
[4] H. Maletta, G. Aeppli, and M. Shapiro, J. Magn. Magn. Mater. **31–34**, 1367 (1983).
[5] Ph. Mangin, D. Boumazouza, G. George *et al.*, Phys. Rev. B **40**, 11123 (1989).
[6] M. Apai, Y. Ishikawa, N. Saito *et al.*, J. Phys. Soc. Jpn. **54**, 781 (1985).
[7] A. P. Murani, Solid State Commun. **34**, 705 (1980).
[8] H. Maletta, J. Appl. Phys. **53**, 2185 (1982).
[9] J. Hubsch, G. Gavoille, and J. Bolfa, J. Appl. Phys. **49**, p. II, 363 (1978).
[10] R. J. Birgeneau, R. A. Cowley, G. Shirane *et al.*, Phys. Rev. B **21**, 317 (1980).
[11] P. A. Beck, Phys. Rev. B **32**, 7255 (1985).
[12] N. N. Efimova and Yu. A. Mamaluĭ, Ukr. Fiz. Zh. **20**, 1201 (1975).
[13] N. N. Efimova, Yu. A. Popkov, and N. V. Tkachenko, Zh. Éksp. Teor. Fiz. **90**, 1413 (1986) [Sov. Phys. JETP **63**, 827 (1986)]; Zh. Éksp. Teor. Fiz. **97**, 1208 (1990) [Sov. Phys. JETP **70**, 678 (1990)]; Fiz. Nizk. Temp. **14**, 981 (1988) [Sov. J. Low Temp. Phys. **14**, 539 (1988)]; Fiz. Nizk. Temp. **15**, 1055 (1989) [Sov. J. Low Temp. Phys. **15**, 584 (1989)]; Fiz. Nizk. Temp. **16**, 1565 (1990) [Sov. J. Low Temp. Phys. **16**, 881 (1990)].
[14] N. N. Efimova, Yu. A. Popkov, G. A. Takzeĭ *et al.*, Fiz. Tverd. Tela (St. Petersburg) **36**, 490 (1994) [Phys. Solid State **36**, 271 (1994)].
[15] N. N. Efimova, V. A. Pervakov, V. I. Ovcharenko *et al.*, Fiz. Tverd. Tela (St. Petersburg) **35**, 2838 (1993) [Phys. Solid State **35**, 1405 (1993)].
[16] V. Ya. Balagurov and V. G. Vaks, Zh. Éksp. Teor. Fiz. **65**, 1600 (1973) [Sov. Phys. JETP **38**, 799 (1973)].
[17] D. E. Khmel'nitskiĭ, Zh. Éksp. Teor. Fiz. **68**, 1960 (1975) [Sov. Phys. JETP **41**, 981 (1975)].
[18] Yu. I. Petrov, *Physics of Small Particles* [in Russian], Nauka, Moscow (1982), p. 358.
[19] N. V. Tkachenko, *Dissertation for the Degree of Candidate of Physico-mathematical Sciences* [in Russian], Khar'kov. Gos. Univ., Kharkov (1990).
[20] I. Ya. Korenblit and E. F. Shender, Usp. Fiz. Nauk **126**, 233 (1978) [Sov. Phys. Usp. **21**, 832 (1978)].
[21] K. P. Belov, *Magnetic Transitions*, Consultants Bureau, New York (1961).
[22] A. Arrott, Phys. Rev. Lett. **20**, 1029 (1968).
[23] S. N. Kaul, J. Magn. Magn. Mater. **53**, 5 (1985).
[24] G. Herzer, M. Fähnle, T. Egami *et al.*, J. Appl. Phys. **52**, p. II, 1794 (1981).
[25] M. Fähnle and H. Kronmüller, Phys. Status Solidi B **98**, 219 (1980).
[26] M. Fähnle, Phys. Status Solidi B **99**, 547 (1980).
[27] E. P. Wohlfarth, Proc. Intermag. Conf. IEEE Trans. Magn. **14**, 933 (1978).
[28] A. Arrott and J. E. Noakes, Phys. Rev. Lett. **19**, 786 (1967).
[29] N. N. Efimova, S. R. Kufterina, Yu. A. Popkov *et al.*, Fiz. Nizk. Temp. **22**, 1079 (1996) [Low Temp. Phys. **22**, (1996)].
[30] H. Pfeiffer and W. Schüppel, J. Magn. Magn. Mater. **130**, 92 (1994).

Translated by James S. Wood

Multiply connected Fermi sphere and fermion condensation

M. V. Zverev*) and M. Baldo†)

Russian Scientific Center ''Kurchatov Institute,'' 123182 Moscow, Russia
(Submitted 2 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 2078–2088 (December 1998)

We examine the structure of the ground state of a homogeneous Fermi liquid beyond the instability point of the Fermi-like quasiparticle momentum distribution in the effective-functional method with a strong repulsive effective interaction. A numerical study of the initial stage of rearrangement of the ground state, based on a simple effective functional, showed that there exists a temperature T_0, above which the behavior of the system is the same as in the theory of fermion condensation, and for $T<T_0$ the scenario of rearrangement of the ground state is different. At low temperatures an intermediate structure arises, with a multiply connected quasiparticle momentum distribution. The transition of this structure with growth of the coupling constant to a state with a fermion condensate is discussed. © *1998 American Institute of Physics.* [S1063-7761(98)01312-2]

1. INTRODUCTION

The applicability of Landau's Fermi liquid theory[1] to a description of the properties of strongly correlated Fermi systems has long been debated in the literature. For one-dimensional systems this theory does not apply. For them an alternative quasiparticle picture is provided by the concept, based on the Luttinger model,[2] of a Luttinger liquid[3] in which the single-particle Green's function does not contain a quasiparticle pole. After the discovery of HTSC compounds having quasi-two-dimensional structure and possessing properties contrary to Landau theory, the boundaries of the non-Fermi-liquid viewpoint have been extended to encompass also two-dimensional, strongly correlated Fermi systems.[4,5] However, the recently measured electronic spectra of such compounds[6–8] appear to speak of the existence of a single-particle pole in the electron Green's function. At the same time, the authors of Refs. 9–12 have discovered new possibilities in the quasiparticle approach. A new class of systems predicted in Refs. 11 and 12 with a fermion condensate possess a rich spectrum of properties,[12–15] including those which had hitherto been the prerogative of the non-Fermi-liquid picture. As has been shown in various models,[12–14,16] a state with a fermion condensate stems from a rearrangement of the ground state of the system of quasiparticles. Such a rearrangement occurs when as a change in the external parameters violates the necessary condition for stability in a quasiparticle system with a Fermi momentum distribution. In the present paper we consider a model of a homogeneous three-dimensional Fermi liquid in which a change in the parameters also results in a rearrangement of the ground state of the quasiparticle system, and we investigate the scenario of the initial stage of this rearrangement.

2. DOUBLY CONNECTED FERMI SPHERE AND FERMION CONDENSATE; EFFECTIVE QUASIPARTICLE FUNCTIONAL

To start with, let us recall the relation in Landau theory between the quasiparticle distribution $n_{\mathbf{p}}(T)$ and the quasiparticle spectrum $\varepsilon_{\mathbf{p}}(T)$:

$$n_{\mathbf{p}}(T)=\left\{1+\exp\frac{\varepsilon_{\mathbf{p}}(T)-\mu(T)}{T}\right\}^{-1} \tag{1}$$

[$\mu(T)$ is the chemical potential], which follows from the variational equation $\delta F/\delta n_{\mathbf{p}}=0$ (F is the free energy of the system), in which the entropy is given by the usual combinatorial expression.[17] Expression (1), on the one hand, is simply the Fermi quasiparticle energy distribution. On the other hand, this relation is an equation for the quasiparticle distribution in momentum space since the quasiparticle energy, being the variational derivative of the ground-state energy functional E_0 with respect to the quasiparticle distribution $\varepsilon_{\mathbf{p}}(T)=\delta E_0/\delta n_{\mathbf{p}}(T)$, is itself a functional of $n_{\mathbf{p}}(T)$.

Landau theory postulates that in a homogeneous and isotropic Fermi liquid the quasiparticle distribution in momentum space at $T=0$ has, as in a Fermi gas, the form of a filled Fermi sphere: $n_F^{(0)}(p)=\theta(p_F-p)$ (the bounding momentum p_F is related to the density of the system ρ by the formula $\rho=p_F^3/3\pi^2$). The low-temperature behavior of the quasiparticle spectrum responsible for this quasiparticle momentum distribution has the form[17]

$$\varepsilon_p(T)-\mu(T)=\xi(p)+O(T^2). \tag{2}$$

The function $\xi(p)$ grows monotonically in the vicinity of the Fermi momentum, changing sign at $p=p_F$. Its slope at this point—the group velocity $v_F=d\xi(p)/dp|_{p=p_F}$ of the quasiparticles at the Fermi surface—is determined by a phenomenological parameter of Fermi liquid theory, the effective mass $M^*=p_F/v_F$.

In a homogeneous strongly correlated Fermi system the momentum distribution of the quasiparticles, which minimizes the energy functional $E_0[n(p)]$ at $T=0$, of course, is not necessarily found at the inflection point $n_F^{(0)}(p)$ of the functional space $[n]$. For example, the authors of Refs. 9 and 10 constructed effective functionals $E_0[n(p)]$, which for certain values of the parameters attain their minimum on the doubly connected Fermi sphere

$$n_F^{(1)}(p)=\theta(p_1-p)-\theta(p_2-p)+\theta(p_3-p). \tag{3}$$

A completely different quasiparticle ground state corresponds to a system with a fermion condensate.[11–14] In order not to send the reader back to the original works, we will briefly elucidate the main idea of the concept of a fermion condensate. A homogeneous and isotropic system with a fermion condensate is described by the singular solution of Eq. (1), which corresponds to the segment of the spectrum that is linear in T,[11,12]

$$\varepsilon_p(T)-\mu(T)=T\nu_0(p)+o(T), \quad p_i<p<p_f. \tag{4}$$

In this low-temperature expansion, in contrast to the Fermi-liquid expression (2) there is no term independent of T. This means that at $T=0$ the quasiparticle spectrum in the interval $p_i<p<p_f$ has a plateau $\varepsilon_p\equiv\mu$. For $T>0$ the slope of the plateau is linear in T, and its position relative to the chemical potential $\mu(T)$ is determined by the function $\nu_0(p)$ associated with the quasiparticle momentum distribution in the condensate at $T=0$. Indeed, the singular solution of Eq. (1), which is easily obtained by substituting formula (4) into Eq. (1), has the form $n_p(T)=n_0(p)+O(T)$, where

$$n_0(p)=\{1+\exp[\nu_0(p)]\}^{-1} \quad (p_i<p<p_f) \tag{5}$$

is the momentum distribution of the condensate quasiparticles at $T=0$. Beyond the limits of the condensate region $n_0(p)=1$ for $p<p_i$ and $n_0(p)=0$ for $p>p_f$ (Refs. 11 and 12). The specific form of $n_p(T)$ and $\varepsilon_p(T)$ can be found if the functional dependence $E_0[n_p(T)]$ is known. References 12–14 and 16 investigated a series of effective functionals for which the ground state of the system with quasiparticle distribution $n_F^{(0)}(p)$ becomes unstable when the coupling constants exceed some critical value, and the minimum is attained in the singular solutions corresponding to the state with a fermion condensate. In the present paper we show that a scenario is possible in which the system transitions to a state with a fermion condensate through an intermediate structure, which corresponds to a multiply connected quasiparticle distribution of the type (3).

In this paper we examine the initial stage of rearrangement, and in this stage, as will become clear in what follows, all changes in the quasiparticle momentum distribution take place in a relatively narrow layer near the Fermi surface. Therefore we neglect the third and higher orders in the variation δn in the formula of Landau theory for the variation of the energy of the system. As is well known, this is the basis for the concept, used in many branches of many-body theory (especially widely in nuclear theory), of the effective functional, the main ingredient of which is the phenomenological effective interaction. In accordance with this approach, we will consider the simple effective functional of the energy of the system of quasiparticles

$$E_0(T)=\int \frac{p^2}{2M} n_p(T)\, d\tau+\frac{1}{2} \times\int V(\mathbf{p}-\mathbf{p}')\, n_p(T)\, n_{p'}(T)\, d\tau d\tau' \tag{6}$$

with effective repulsive interaction in the form

$$V(\mathbf{p}-\mathbf{p}')=\frac{V_0}{(\mathbf{p}-\mathbf{p}')^2+\alpha^2}. \tag{7}$$

In formula (6) $d\tau$ denotes integration with respect to $\mathbf{p}/(2\pi)^3$ and calculation of the trace over the spin variable.

The quasiparticle spectrum is found by calculating the variational derivative $\delta E_0/\delta n_p(T)$:

$$\varepsilon_p(T)=\frac{p^2}{2M}+\int V(\mathbf{p}-\mathbf{p}')\, n_{p'}(T)\, d\tau'. \tag{8}$$

The functional relation between $\varepsilon_p(T)$ and $n_p(T)$ given by formula (8) together with Eq. (1) and the normalization condition

$$\int n_p(T) d\tau=\rho \tag{9}$$

represent a system for calculating the quasiparticle distribution $n_p(T)$ and spectrum $\varepsilon_p(T)$. The present work uses the parameter value $\alpha=0.07\,p_F$ and examines the behavior of a system as a function of the parameter V_0 (for convenience everywhere below we use the dimensionless parameter $\gamma=MV_0/(4\pi^2 p_F)$).

Let us comment on some aspects of the numerical scheme. Equation (6) together with formula (1) is a nonlinear integral equation for the function $\varepsilon(p)$. This equation was solved on a grid with grid step $h_p=5\times10^{-5}p_F$. To convolve $n(p)$ with the effective interaction $V(p,p')$ numerically, we used a five-point variant of the Newton–Cotes quadrature formula with a five-point output filter. The nonlinear equation was solved using an iterative method with relaxation. The accuracy of the solution was determined by substituting it into the original equation. The acceptable error (maximum discrepancy between the left and right sides) was taken to be equal to $10^{-8}\varepsilon_F$. The importance of such an accurate calculation of the spectrum will become clear from the discussion of the calculated results which follows. The number of iterations necessary to achieve the chosen accuracy for the relaxation parameter $w=0.001$ (this value of the parameter proved to be optimal for stability of the iterative procedure), was roughly 30,000. Note also that the results turned out not to depend on from which point of the functional space the iterative procedure started (the trajectories along which the iterations reach the solution and, of course, the stability of the procedure depend on the starting point). Thus, for example, the solution for $\gamma=0.50$ at $T=10^{-7}$ was obtained twice: once starting from the solution for $\gamma=0.50$ at $T=10^{-5}$, and a second time, starting from the solution for $\gamma=0.48$ at $T=10^{-7}$ (here and everywhere below we assume the temperature T to be given in units of $\varepsilon_F^0=p_F^2/2M$).

3. STABILITY OF SINGLY CONNECTED AND DOUBLY CONNECTED FERMI SPHERES: DISCUSSION OF NUMERICAL RESULTS

We begin our discussion of the calculated results by clarifying at what value of $\gamma_c^{(0)}$ the necessary condition for stability of the ground state with quasiparticle distribution $n_F^{(0)}(p)$ at $T=0$ is violated. This condition arises[11,12] as a

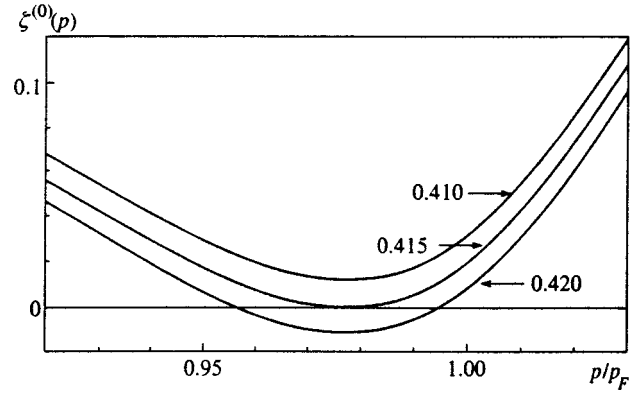

FIG. 1. The function $\zeta^{(0)}(p)$, calculated for $\gamma=0.410$, 0.415, and 0.420.

condition of nonnegativity of the variation of the energy E_0 of the ground state for any admissible variations of the distribution function $n(p)$:

$$\delta E_0=\int [\varepsilon(p)-\mu]\,\delta n(p)\,d\tau. \tag{10}$$

For the distribution $n_F^{(0)}(p)$ the admissible variations have the same sign as the difference $p-p_F$. Therefore, replacing the chemical potential μ in formula (10) by the energy $\varepsilon(p_F)$, we can reformulate the necessary condition for stability of $n_F^{(0)}(p)$ in the form of a requirement[11,12] that

$$s(p)=2M\frac{\varepsilon(p)-\varepsilon(p_F)}{p^2-p_F^2} \tag{11}$$

be positive for all values of the momentum p. If the function $s(p)$ first vanishes near p_F, then this is equivalent to the appearance of an inflection in the curve $\varepsilon(p)$ in the vicinity of the Fermi momentum. The derivative $d\varepsilon/dp$ for the distribution $n_F^{(0)}(p)$ is easily calculated. For the effective interaction (7), in dimensionless form it is equal to

$$\zeta^{(0)}(p)=\frac{M}{p_F}\frac{d\varepsilon}{dp}=\frac{p}{p_F}+\frac{\gamma p_F}{p}$$
$$-\frac{\gamma(p^2+p_F^2+\alpha^2)}{4p^2}\ln\frac{(p+p_F)^2+\alpha^2}{(p-p_F)^2+\alpha^2}. \tag{12}$$

From Fig. 1, which plots the curves $\zeta^{(0)}(p)$ for three values of γ, it can be seen that tangency to the straight line $\zeta=0$ obtains for $\gamma=\gamma_c^{(0)}\simeq 0.415$ at the point $p_c\simeq 0.97p_F$ [the nearness of p_c to p_F justifies replacing $s(p)$ by the function $\zeta^{(0)}(p)$]. For $\gamma>\gamma_c^{(0)}$ the ground state with quasiparticle distribution $n_F^{(0)}(p)$ becomes unstable, and it undergoes rearrangement.

The question of how the ground state is structured immediately beyond the transition point is answered by Fig. 2, which plots the results of a calculation of $n_p(T)$ for various values of T for $\gamma=0.45$. For $T\sim 10^{-3}$ the dependence $n_p(T)$ has the form of the quasiparticle momentum distribution in a system with a fermion condensate (below we will dwell on this point in more detail). For $T<10^{-3}$ a dip appears in the distribution $n_p(T)$, which deepens as the temperature is lowered, and at $T=10^{-7}$ it is already almost indistinguishable from the doubly connected Fermi sphere (3). The quasiparticle spectrum $\varepsilon_p(T)$ corresponding to this distribution, calculated for $T=10^{-7}$, is plotted in Fig. 3. In contrast to the spectrum of the fermion condensate, which has the form of a plateau ''lying on the chemical potential μ'' at $T=0$ (Refs. 11 and 12) and having a small slope for $T>0$, the quasiparticle spectrum of the doubly connected Fermi sphere is found to be equal to μ on the boundary p_1 of the inner sphere and on the boundaries p_2 and p_3 of the spherical shell. On the segment $p_1<p<p_3$ the deviation of $\varepsilon_p(T)$ from μ at the minimum of the spectrum reaches $2\times 10^{-4}\varepsilon_F$, and at the maximum, $2\times 10^{-6}\varepsilon_F$. Although the latter number is very small, recall that it is more than two orders of magnitude greater than the precision with which the spectrum is calculated.

Figure 4 depicts the behavior of a doubly connected Fermi sphere with increasing value of the coupling constant

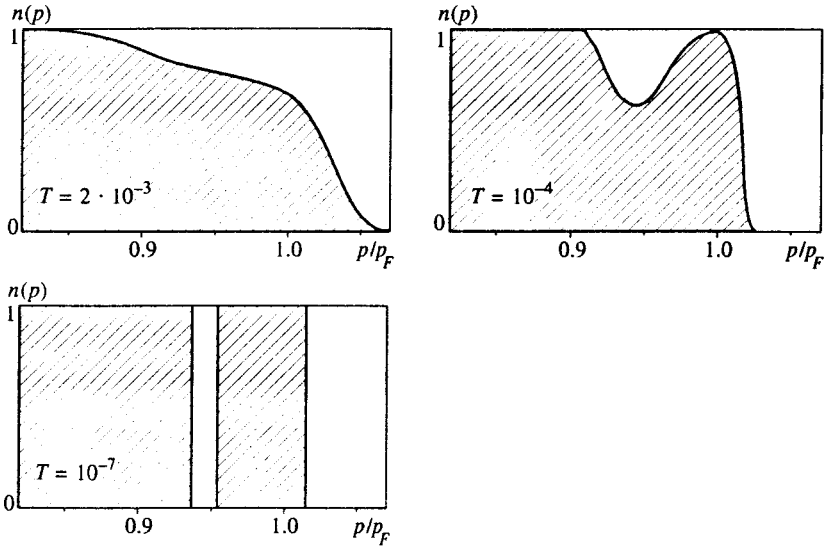

FIG. 2. Quasiparticle momentum distributions $n(p,T)$, calculated for different temperatures for $\gamma=0.45$.

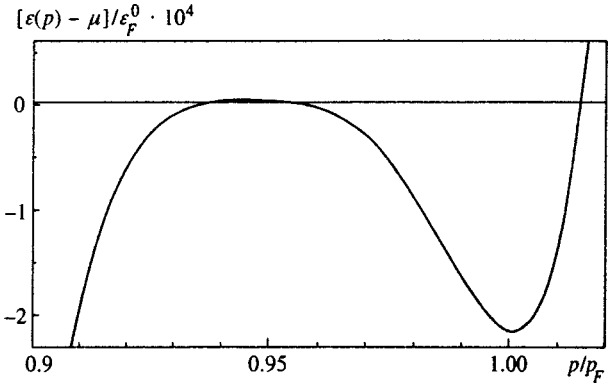

FIG. 3. Quasiparticle spectrum $[\varepsilon(p)-\mu]/\varepsilon_F^0$ for $\gamma=0.45$ at $T=10^{-7}$.

γ. The spherical shell appearing beyond the transition point starts out with a finite thickness, but the gap between it and the inner filled sphere starts out vanishingly small. As γ increases the outer shell thickens and separates from the inner sphere. What happens to this solution if γ continues to increase? In order to better understand this question, we will study the stability of such a layered Fermi sphere. Toward this end, we elucidate the function $\zeta^{(1)}(p)$ for the momentum distribution in the form (3) and observe where and when the critical change of sign of this function takes place. An elementary calculation gives

$$\zeta^{(1)}(p)=\frac{M}{p_F}\frac{d\varepsilon}{dp}=\frac{p}{p_F}+\sum_{i=1}^{3}(-1)^{i-1}\times\left\{\frac{\gamma p_i}{p}-\frac{\gamma(p^2+p_i^2+\alpha^2)}{4p^2}\ln\frac{(p+p_i)^2+\alpha^2}{(p-p_i)^2+\alpha^2}\right\}. \tag{13}$$

The function $\zeta^{(1)}(p)$, calculated for different values of γ, is depicted in Fig. 4. In this calculation we choose as the boundary momenta p_1,p_2,p_3 those momenta at which the absolute value of the derivative dn/dp takes its maximum for $T=10^{-7}$. It is easy to imagine that in the interval $\gamma_c^{(0)}<\gamma<\gamma_c^{(1)}\simeq 0.452$ the two points at which $\zeta^{(1)}(p)$ changes sign are located in such a way that the local minimum and maximum of $\varepsilon(p)$ corresponding to them lie in regions where $n(p)$ is equal respectively to zero and unity, i.e., the sign of the difference $\varepsilon(p)-\mu$ coincides with the sign of the variations $\delta n(p)$ permitted by the Pauli principle and the distribution satisfies the necessary condition for stability. But for $\gamma>\gamma_c^{(1)}$ the situation is already different, as is clear from Fig. 4. For such behavior of the function $\zeta^{(1)}(p)$ there exist regions where we have $\varepsilon(p)-\mu>0$, but $n(p)=1$. This means that the necessary condition for stability is violated in the system, since variations $\delta n(p)$ exist which lower the ground-state energy. This implies that for $\gamma=\gamma_c^{(1)}$ a new rearrangement of the ground state of the system of quasiparticles takes place. Figure 5 shows how the quasiparticle distribution is arranged beyond the second point $\gamma_c^{(1)}$ of the rearrangement for the case $\gamma=0.46$ for different values of T. The calculation shows that a new layering of the doubly connected Fermi sphere has already taken place for this value of the coupling constant, and for $T=10^{-7}$ the quasiparticle distribution $n(p)$ is close to a triply connected Fermi sphere:

$$n_F^{(2)}(p)=\theta(p_1-p)-\theta(p_2-p)+\theta(p_3-p)-\theta(p_4-p)+\theta(p_5-p). \tag{14}$$

The scenario of increasing layering with increasing value of the parameter γ does not stop at a triply connected Fermi sphere. Thus, for $\gamma=0.50$ the distribution $n_p(T)$ depicted in Fig. 6 consists of an small filled inner sphere with radius $\sim 0.85\,p_F$ at low temperatures T, surrounded by four filled spherical shells of thickness $\sim(0.3\text{–}0.4)\,p_F$, which are separated by empty spherical shells of thickness $\sim(0.1\text{–}0.2)\,p_F$. The quasiparticle spectrum corresponding to such distributions is depicted in Fig. 7. At $T=10^{-7}$ the spectrum intersects the chemical potential nine times—on the boundary of the inner sphere and on the boundaries of the spherical shells.

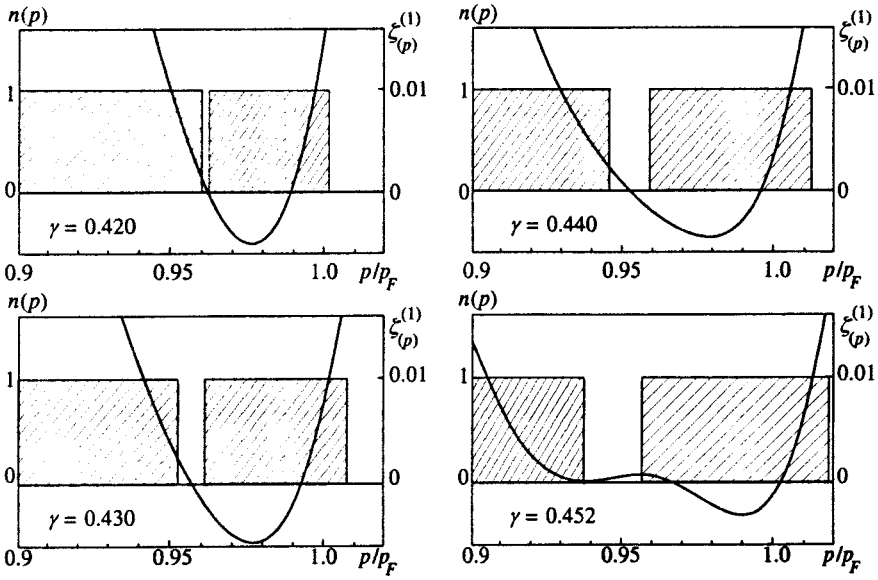

FIG. 4. Quasiparticle momentum distributions $n(p)$ and the function $\zeta^{(1)}(p)$, calculated for different values of the parameter γ.

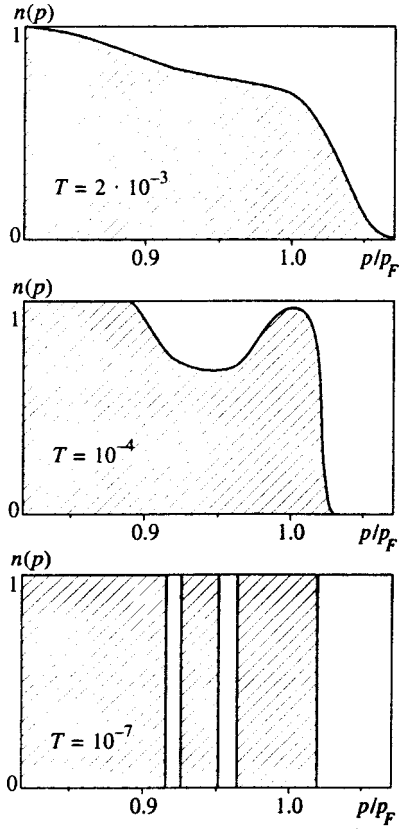

FIG. 5. The same as in Fig. 2, but for $\gamma=0.46$.

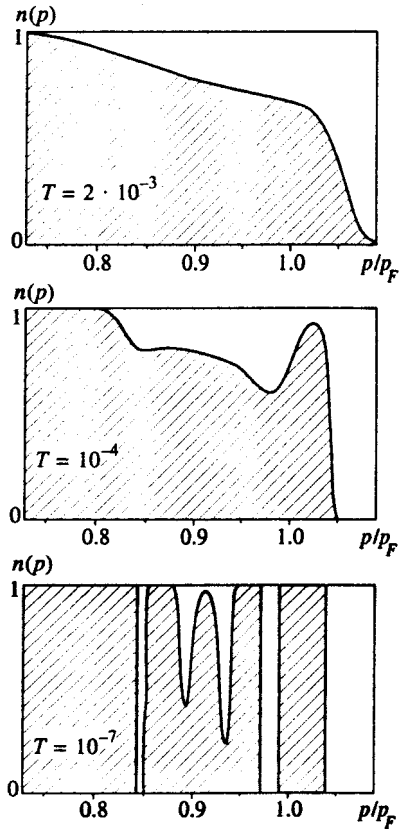

FIG. 6. The same as in Fig. 2, but for $\gamma=0.50$.

4. MULTIPLY CONNECTED FERMI SPHERE AS AN INTERMEDIATE STRUCTURE ON THE PATH TO A FERMION CONDENSATE

Thus, beyond the transition point $\gamma_c^{(0)}$ the scenario for rearrangement of the ground state of the system of quasiparticles at low temperatures $T<10^{-3}$ consists of a sequence of rearrangements taking place with increase of the coupling constant γ, as a result of each of which a new spherical shell arises in momentum space. What distinguishes a system with a multiply connected quasiparticle momentum distribution from a system with a fermion condensate, and what do they have in common? Let us recall the main features of the state with a fermion condensate and the rearrangement scenario corresponding to the emergence of such a state. First of all, we note the plateau in the quasiparticle spectrum $\varepsilon_p(T)$, lying on the chemical potential μ for $T=0$ in accordance with formula (4) and acquiring a slope that increases with temperature. The spectrum $\varepsilon_p(T)$ for a multiply connected Fermi sphere does not generally behave like this. At the lowest calculated temperature $T=10^{-7}$ the spectrum is found to be equal to the chemical potential only at the boundary points of the spherical shells of the multiply connected momentum distribution. Thus, the states with a multiply connected Fermi sphere have no macroscopic degeneracy as is the case in a system with a fermion condensate. At the same time, there are features in the density of states, associated with the existence of maxima and minima of the function $\varepsilon(p)$. These peculiarities gradually disappear with growth of T up to the temperature $T_0\sim 10^{-3}$, at which the last knee in the spectrum is smoothed out. For $T>T_0$ the difference $\varepsilon_p(T)-\mu(T)$ becomes linear in temperature, as in systems with a fermion condensate.

Another property of a system with a fermion condensate is the distribution $n_p(T)$, which for $T=0$ is given by expression (5). The region occupied by the fermion condensate obeys $0<n_p(T)<1$, which corresponds to a nonzero value of its entropy at $T=0$. This violation of the Nernst theorem disappears when correlations are included (e.g., superfluid correlations), which because of degeneracy slowly rearrange the ground state and make the entropy zero at $T=0$. The entropy of a state with a multiply connected Fermi sphere vanishes at $T=0$, since $n(p)$ only takes values of 0 and 1 there. With growth of T a multiply connected distribution changes rapidly—the shell boundaries broaden and the shells coalesce, transitioning at $T\sim T_0$ into a smoothly decreasing dependence, similar to the momentum distribution of a system with a fermion condensate. Along with a rapid change in $n_p(T)$, the entropy of the system grows rapidly as the temperature is raised. Calculations show that at $T=T_0$ the entropy of a system with a multiply connected Fermi sphere reaches the value $S_0\sim\Omega_0/\Omega$, which is equal to the ratio of the phase volume of the region of the multiply connected momentum distribution Ω_0 to the phase volume of the entire system Ω. Just such a value of the entropy would be characteristic at $T\sim T_0$ for a system with a fermion condensate occupying a phase volume Ω_0. For $T>T_0$ the entropy becomes linear in T, as in systems with a fermion condensate.[12,16]

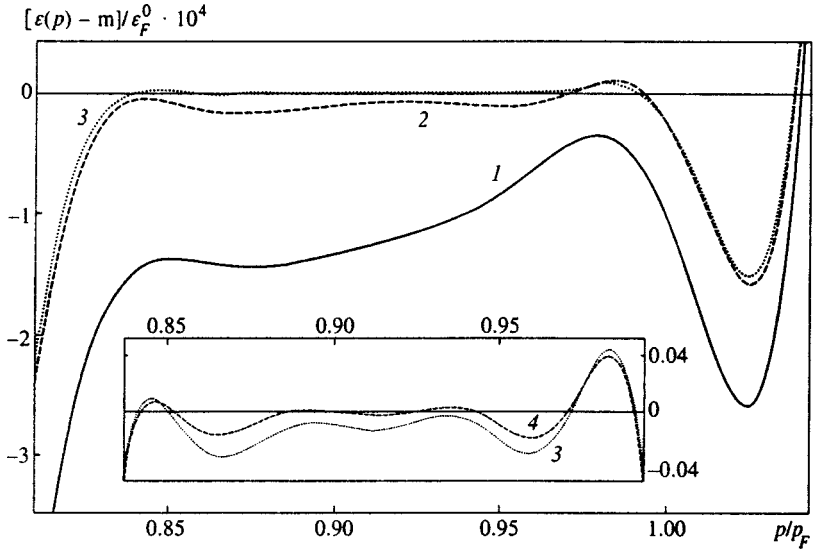

FIG. 7. Quasiparticle spectra $[\varepsilon(p,T)-\mu]/\varepsilon_F^0$, calculated for $\gamma=0.50$ at $T=10^{-4}$ (curve *1*), $T=10^{-5}$ (*2*), $T=10^{-6}$ (*3*), and $T=10^{-7}$ (*4*).

All of these curious trends in the behavior of the momentum distribution, entropy, quasiparticle spectrum, and density of states of a system with a quasiparticle momentum distribution in the form of a multiply connected Fermi sphere are worthy of separate, detailed consideration.

The scenario in which rearrangement takes place when a state with a fermion condensate forms is characterized by one critical value of the coupling constant γ_c, at which a fermion condensate arises in the system. With further growth of γ the phase volume of the fermion condensate increases, but beyond that no further qualitative changes take place in the system—increasing the coupling constant leads only to an increase in the relative phase volume of the fermion condensate, while the momentum distributions and quasiparticle spectra remain fundamentally similar.[11,12] The scenario of rearrangement found in the present calculations is different in different temperature intervals. For $T=0$ it is characterized by a sequence of values $\gamma_c^{(i)}$, after each of which is reached states arise having a higher connectedness of their momentum distribution. The number of such critical values decreases with increasing T, and for $T>T_0$ only $\gamma_c^{(0)}$ survives. This means that in the given model the scenario of rearrangement with formation of a multiply connected Fermi surface with increase of the temperature gives way gradually to the scenario of fermion condensation.

Unfortunately, the computer time needed to calculate the spectra increases rapidly as a function of the phase volume Ω_0 occupied by all the spherical shells. For this reason, in the present study we were able to calculate only for values of the coupling constant $\gamma<0.5$ corresponding to the initial stage of rearrangement. What happens with the system for larger values of γ? In order to better understand this question, we will use a mechanical analogy, treating the momentum p like the spatial coordinate r (Refs. 11 and 12). Thus, the problem of minimizing the ground-state energy functional at $T=0$ can be interpreted in terms of the mechanical analogy as a search for the equilibrium spatial distribution $\nu(r)$ of the particles moving in an external harmonic field $U(r)=kr^2/2$ with stiffness $k=1/M$ and interacting with each other via the forces (7), where the number of particles is fixed by the normalization condition (9). As long as the solution $\nu(r)$ of the mechanical problem exceeds $2/(2\pi)^3$ at just one point, it cannot be taken as the solution $n(p)$, since in that case it would contradict the Pauli principle. As the repulsion between the particles increases, the mechanical system obviously expands and therefore becomes more rarefied. And as soon as the distribution $\nu(r)$ becomes everywhere less than $2/(2\pi)^3$, it begins to correspond to the solution $n(p)$ of the initial problem. It is natural to expect that in the mechanical problem the distribution $\nu(r)$ should be smooth and monotonic; therefore the corresponding solution $n(p)$ is something different than the singular solution corresponding to fermion condensation. Thus, the structure observed in the present calculations having a multiply connected quasiparticle momentum distribution is probably intermediate along the path to fermion condensation. The transition from the phase with a multiply connected momentum distribution to the phase with a fermion condensate taking place as the coupling constant is increased requires separate study.

5. CONCLUSION

In conclusion, we repeat that we have examined the structure of the ground state of a homogeneous Fermi liquid using the method of the effective functional with a strong repulsive interaction characterized by two parameters, the radius in momentum space α and the coupling constant γ. A numerical study of this functional has shown that for a fixed value of the parameter α there exists a critical value of the coupling constant, $\gamma_c^{(0)}$, beyond which the ground state with a Fermi-like quasiparticle distribution becomes unstable and rearranges. The scenario of the initial stage of rearrangement with increasing γ is found to be different in different temperature ranges. For $T=0$ there exists a series of critical constants $\gamma_c^{(i)}$ corresponding to a sequence of transitions, as a result of each of which a new spherical shell of the quasiparticle momentum distribution $n(p)$ arises. The quasiparticle

spectrum $\varepsilon(p)$ corresponding to such a multilayered distribution, in contrast to the spectrum of systems with a fermion condensate, does not have a plateau in a finite region of momentum space, but is equal to the chemical potential on the spheres bounding the filled shells. The ground state with a multiply connected distribution does not possess a macroscopic degeneracy, and the entropy of this state does not vanish at zero temperature. As the temperature increases the shell boundaries rapidly broaden, and for $T \sim T_0 \approx 2 \times 10^{-3}$ there no longer remain any vestiges of the critical constants $\gamma_c^{(i)}$ with the exception of the single critical constant $\gamma_c^{(0)}$. For $T > T_0$ the rearrangement scenario is the scenario of fermion condensation. A qualitative analysis has shown that the found structure with a multiply connected quasiparticle momentum distribution is an intermediate structure and as the coupling constant is increased, it should give way to the fermion condensate.

The authors thank V. A. Khodel' for sustained interest in this work and numerous stimulating discussions, and also S. A. Artamonov, A. É. Bulatov, É. E. Sapershteĭn, S. V. Tolokonnikov and V. R. Shaginyan for fruitful discussions of the questions touched on in this work.

This work was carried out with the partial support of the Russian Fund for Fundamental Research (Grant No. 96-02-19292). One of us (M. V. Z.) is grateful for the hospitality extended him by the National Institute of Nuclear Physics (Catania, Italy), where most of this work was done.

*)E-mail: zverev@polyn.kiae.su
†)Istituto Nazionale di Fisica Nucleare, 95129 Catania, Italy.

[1] L. D. Landau, Zh. Éksp. Teor. Fiz. **30**, 1058 (1956) [Sov. Phys. JETP **3**, 920 (1956)]; *ibid*. **35**, 97 (1959) [Sov. Phys. JETP **8**, 70 (1959)].
[2] J. M. Luttinger, J. Math. Phys. **4**, 1154 (1963); D. C. Mattis and E. H. Lieb, J. Math. Phys. **6**, 304 (1965).
[3] F. D. M. Haldane, J. Phys. C **14**, 2585 (1981).
[4] P. W. Anderson, Phys. Rev. Lett. **64**, 1839 (1990); **65**, 2306 (1990).
[5] C. M. Varma, P. B. Littlewood, S. Schmitt-Rink *et al.*, Phys. Rev. Lett. **63**, 1996 (1989).
[6] Z.-X. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1955).
[7] T. Yokoya, A. Chainani, T. Takahashi *et al.*, Phys. Rev. Lett. **76**, 3009 (1996).
[8] D. H. Lu, M. Schmidt, T. R. Cummins *et al.*, Phys. Rev. Lett. **76**, 4845 (1996).
[9] M. de Llano and J. P. Vary, Phys. Rev. C **19**, 1083 (1979).
[10] M. de Llano, A. Plastino, and J. G. Zabolitsky, Phys. Rev. C **20**, 2418 (1979).
[11] V. A. Khodel', V. R. Shaginyan, JETP Lett. **51**, 626 (1990).
[12] V. A. Khodel, V. R. Shaginyan, and V. V. Khodel, Phys. Rep. **249**, 1 (1994).
[13] V. A. Khodel, J. W. Clark, and V. R. Shaginyan, Solid State Commun. **96**, 353 (1995).
[14] V. A. Khodel', V. R. Shaginyan, and P. Shuk, JETP Lett. **63**, 752 (1996).
[15] M. V. Zverev, V. A. Khodel', V. R. Shaginyan, and M. Baldo, JETP Lett. **65**, 863 (1997).
[16] P. Nozières, J. Phys. I **2**, 443 (1992).
[17] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, N. J., 1963).

Translated by Paul F. Schippnick

Phase transitions in biperiodic stripe domain structures of uniaxial magnetic films with a positive anisotropy constant

G. V. Arzamastseva, F. V. Lisovskiĭ,*) and E. G. Mansvetova

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 141120 Fryazino, Moscow Region, Russia

(Submitted 9 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2089–2110 (December 1998)

New phase transitions induced by a magnetic field and accompanied by a change in the symmetry or the period of the distribution of the magnetization vector are observed in biperiodic stripe domain structures of iron garnet films with a positive anisotropy constant. A symmetry classification of the observed types of domain structures is derived, and the form of the state diagram of the films is determined in the $H_\perp H_\parallel$ plane, where $H_\perp$ and $H_\parallel$ are the components of the magnetic field vector perpendicular and parallel to the normal to the surface.
 [S1063-7761(98)01412-7]

1. INTRODUCTION

It is a well-known fact that when the thickness L of a uniaxial magnetic film with a positive anisotropy constant β_u and with the easy magnetization axis oriented along the normal **n** to the surface (the normal also being the direction of the z axis) is gradually increased, at a critical thickness L_{cr}^* the initial equilibrium stripe domain structure acquires quasiharmonic distortions of the domain wall profile, causing the distribution of the magnetization vector **M** to become biperiodic. In thicker ($L>L_{cr}^{(1)}>L_{cr}^*$) films this modification of the distribution of **M** at the surfaces of the magnet is a precursor to the onset of chains of isolated cone-shaped surface domains within each stripe domain, which penetrate into the bulk of the film to a maximum depth $L/2$ and for which the component M_z has the opposite sign from the domains in which the given cone-shaped domains reside. With a further increase in the thickness ($L>L_{cr}^{(2)}>L_{cr}^{(1)}$) smaller cone-shaped domains with the opposite sign of the projection M_z begin to form within each of the original cone-shaped domains, and so on until, as a result, complex structures of the ''Russian nested dolls'' type are formed, consisting of a set of domains contained one inside another. Historically the term ''branching'' has been attached to the process, but is something of a misnomer in that only localized layering of the distribution **M(r)** into the depth of the film takes place, characterized by the alternating sign of M_z in consecutive layers. True branching of the domain walls near the surfaces of the film, accompanied by the formation of fractal-like configurations, is observed for thicknesses in the interval $L_{cr}^{(1)}>L>L_{cr}^*+\delta L$, where $\delta L<(L_{cr}^{(1)}-L_{cr}^*)$, i.e., from the inception of the first side branches on curved domain walls to the onset of the first cone-shaped domains (see, e.g., Refs. 1–6).

This behavior of the domain structure is inherent in magnets with any nonzero, positive value of the uniaxial anisotropy constant, from magnetoplumbites and hexaferrites ($\beta_u \gg 1$) to low-coercive-force alloys ($\beta_u \ll 1$). The only difference is that in zero external magnetic field the ground state of the films corresponds in the first case to labyrinth stripe domain structures with chaotic orientation of the domain walls and in the second case to ordered two-dimensional stripe domain lattices exhibiting translational symmetry. The ordering of the domain walls in films with small values of β_u is caused by the large deviation of the vector **M** from the normal to the surface.

For two decades the behavior of biperiodic stripe domain structures[1)] has been investigated experimentally on metal and alloy films[1,2,5,9–18] and on ferrite films having magnetoplumbite,[2,3,9,10,18–21] spinel,[22,23] garnet,[24] and hexagonal[2,4] structures by means of Lorentz electron microscopy,[2] powder patterns,[1–5] polarization optical microscopy,[9,24] and magnetooptical diffraction.[24] In that same time span a few attempts have been made to describe the properties of biperiodic stripe domain structures via the Ritz variational principle using various test functions and working within the concept of geometrical (unstructured) domain walls.[13,25–28] Interest in biperiodic stripe domain structures eventually died out and has resurfaced only recently in connection with the magnetooptical diffraction of light by such domain arrays.[29,30]

An analysis of the cited papers shows clearly that the behavior of biperiodic stripe domain structures in external magnetic fields has not been adequately studied. In particular, the possibility of the existence of various modifications of such nonuniform distributions of the magnetic moment and the question of phase transitions between them have been all but ignored. Our investigations fill this gap.

The paper is organized as follows. In Sec. 2 the experimental conditions are described, and brief information is given in regard to the procedure used to grow the experimental samples, their parameters, and the general properties of biperiodic stripe domain structures. Section 3 contains a preliminary analysis of the experimental results, which is essential to the material that follows. A symmetry classification of the experimentally observed biperiodic stripe domain structures is given in Sec. 4, drawn from the apparatus of the theory of two-dimensional space groups. The kinetics of

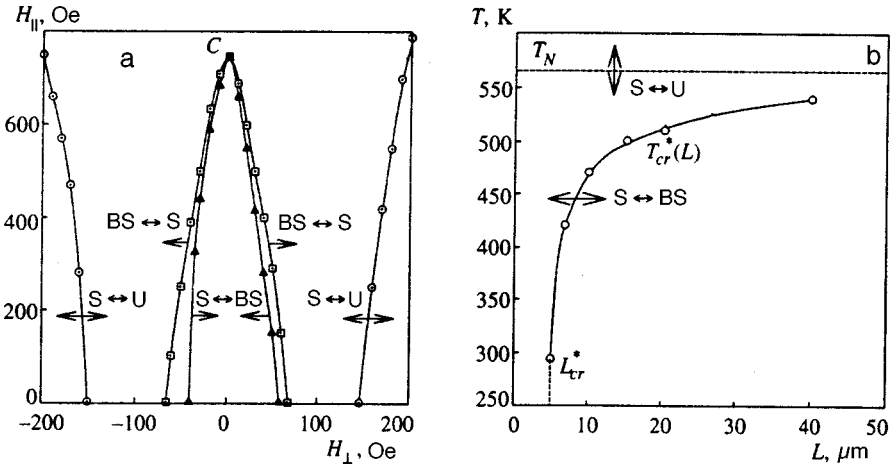

FIG. 1. State diagrams of uniaxial magnetic garnet films for orientational (a) and spontaneous (b) phase transitions. (U) Uniformly magnetized state; (S) simple stripe domain structures; (BS) biperiodic stripe domain structures; the arrows on the curves indicate the directions of variation of the transition-inducing parameter.

magnetic field-induced phase transitions in biperiodic stripe domain structures is described in Sec. 5. A brief discussion of the results and related problems is set forth in the Conclusion.

2. EXPERIMENTAL CONDITIONS AND GENERAL PROPERTIES OF BIPERIODIC STRIPE DOMAIN STRUCTURES

We have investigated the domain structures of uniaxial magnetic garnet films with compositions $Lu_{2.1}Bi_{0.9}Fe_{5-x}Mg_xO_{12}$ ($x=0-0.17$) and $Lu_{2.1}Bi_{0.9}Fe_5O_{12}$ (referred to from now on as type-I and type-II films) and thicknesses $L=1-40\,\mu$m, grown by liquid-phase epitaxy (using solvents Bi_2O_3 and $PbO-Bi_2O_3$, respectively) on (111)-oriented $Gd_3Ga_5O_{12}$ substrates. The growth conditions were similar to those used by Tamada *et al.*[31,32] The Néel temperature of the films was T_N was 560 K, and the saturation magnetization $4\pi M$ was 1800 G. The magnetization of the samples to saturation by fields along ($H_\parallel=H_z$) and perpendicular ($H_\perp=H_y$) to the normal **n** to the surface in type-I films required fields $H^*_\parallel=1500-1750$ Oe) and ($H^*_\perp=25-200$ Oe, respectively, and in type II fields, 1600–1750 Oe and 300–500 Oe, respectively, where $H^*_\parallel$ and $H^*_\perp$, as a rule, increased as functions of the film thickness. An approximate estimation[2)] of the uniaxial anisotropy constants, based on the relation

$$\beta_u \approx 1 - H^*_\perp/4\pi M,$$

gives $\beta_u \approx 0.88-0.98$ for type-I films and $\beta_u \approx 0.72-0.83$ for type-II films. The films also exhibited weak cubic anisotropy, which in our case merely imparted a slight asymmetry to the state diagram on the $H_\perp H_\parallel$ plane relative to the ordinate axis without producing any qualitative changes in the response of the domain structure to an external magnetic field (see Ref. 33).

High specific Faraday rotation in the visible range (more than 1 deg/μm) permitted reliable visual determination of the type of domain structure of the films by means of a polarization microscope and from diffraction patterns; the light source was an incandescent lamp in the first case and a helium-neon laser with a working wavelength $\lambda_L=0.63\,\mu$m in the second case. The radiation intensity in the individual diffraction maxima were recorded by a moving photosensitive probe with the application of spatial filtering, which was also used to enhance the contrast and sharpness of subtle details of the domain images. The domain structures were photographed (in TIFF format) by a Kodak DC 120 digital camera which provided a resolution of 1280×960 pixels and were processed by means of standard computerized photoretouch software. Two pairs of orthogonally positioned coils were used to generate the magnetic fields $H_\parallel$ and $H_\perp$ with maximum strengths of 10 kOe and 2 kOe, respectively.

In zero external field a simple (monoperiodic) stripe domain structure existed in films having a thickness smaller than 5 μm, and a biperiodic stripe domain structure with harmonic surface modulation of the domain wall profile was found in thicker films. The amplitude of the distortions of the domain wall profile increased with the film thickness, but conical near-surface domains did not form, even in the thickest films ($L=40\,\mu$m) attainable by liquid-phase epitaxy. The boundaries of the regions in which biperiodic domain structures exist on the LT and $H_\perp H_\parallel$ planes could be determined by observing the behavior of the domain structures as the films were heated (from room temperature) and as the strength and orientation of the external magnetic field were varied (see Fig. 1). There is clearly an upper temperature limit $T^*_{cr}(L)$ for modulation of the domain wall profile, i.e., a simple stripe domain structure is encountered at $T>T^*_{cr}$. The modulation of the domain wall profile at $T=T^*_{cr}$ vanishes as a second-order phase transition (see Ref. 34). As the film thickness is increased, the transition temperature increases monotonically, asymptotically approaching the Néel temperature T_N but always below it (see Fig. 1b). Figure 1b shows the results of a series of type-II films; the results are similar for type-I films.

A similar situation is encountered for oriented phase transitions (for $T=$const) induced by the magnetic field $H_\perp$, as illustrated by Fig. 1a, which shows a portion of the state diagram[3)] for a type-I film having a thickness of approximately 10 μm (from now on designated film No. 1) at $T=293$ K. The region of biperiodic stripe domain structures on the state diagram in coordinates $\{H_\perp, H_\parallel\}$ always lies inside the region of stability of simple stripe domain structures, the transformation from one type of domain structure

to the other taking place as a first-order phase transition. One exception is the point C corresponding to the apex of the biperiodic region. In Fig. 1a the point C is located on the ordinate, which is the symmetry axis of the state diagram; as the film rotates about the normal under the influence of the weak cubic anisotropy, symmetry is lost in general, and C is no longer on the ordinate. If the film thickness L tends to $L_{cr}^{*}+0\simeq 5\ \mu$m, the region of biperiodic stripe domain structures shrinks to a point; as the film thickness is increased, this region broadens, and the BS$\rightarrow$S curves on the diagram of Fig. 1a approach the S$\leftrightarrow$U curves, but a gap always exists between them.

The generation of stripe domain structures in uniaxial films from the homogeneous state as the magnitude of the external magnetic field $H_{\perp}$ is lowered with $H_{\parallel}=0$ is accompanied by symmetry breaking with respect to in-plane translations and reflections and with respect to rotations about the normal to the surface, so that the magnet in the domain phase must be described by three order parameters: translational, orientational, and ''modulational,'' which are characterized by (respectively) the degree of periodicity of the distribution of the magnetization vector, the degree of parallelism of the domain walls, and the maximum angle of deviation of the vector **M** from the magnetic field vector.[35] Only orientational and modulational long-range orders exist for $T\neq 0$, because the translation group is not discrete. Long-range translational order is broken by the onset of bound or free magnetic dislocations in the stripe domain structure. In the first case the state of the magnet corresponds to the so-called Berezinskiĭ–Kosterlitz–Thouless (BKT) phase, in which the ''translational'' correlation function decreases (in the far zone) according to a power law, and the domain walls are completely ordered; in the second case the state corresponds to the liquid-crystal phase, for which the translational correlation function decreases exponentially, and the orientational correlation function decreases by a power law.[35–38]

3. PRELIMINARY ANALYSIS OF THE EXPERIMENTAL RESULTS

The experiments were performed at $T=293$ K for various values of $H_{\perp}$ and $H_{\parallel}=$ const. For a given value of $H_{\parallel}$ the saturation field $H_{\perp}>H_{\perp}^{*}$ was determined, then the current in the corresponding pair of coils of the electromagnet was gradually reduced to zero, the polarity of the current was switched, and the film was again brought to the magnetically saturated state by a field $H_{\perp}$ in the opposite direction to the first. Next, all the observations were repeated for the ascending branch of the hysteresis loop. Throughout the entire region of existence of simple and biperiodic domain structures the latter were well ordered, but contained defects in the form of free and bound magnetic dislocations, i.e., the state of the magnet corresponded to the liquid-crystal phase or the BKT phase.[35–38]

In the phase transition, induced by the magnetic field $H_{\perp}$, from the uniformly magnetized state to the phase with a simple stripe domain structure of period d the boundaries between the domains were oriented parallel to the vector $\mathbf{H}_{\perp}$, i.e., the domain array was characterized by the reciprocal lattice vector $\mathbf{b}_1=(2\pi/d)\mathbf{e}_x$. Two other parameters besides d must be used to describe the biperiodic domain structures: Λ_1 and Λ_2 (in general $\Lambda_1\neq\Lambda_2$), which are equal to the periods of surface modulation of the domain wall profile at the interfaces of the film with free space and with the substrate, respectively. The space of reciprocal lattice vectors becomes two-dimensional in this case, where $\mathbf{b}_2=(2\pi/\Lambda_i)\mathbf{e}_y, i=1,2$.

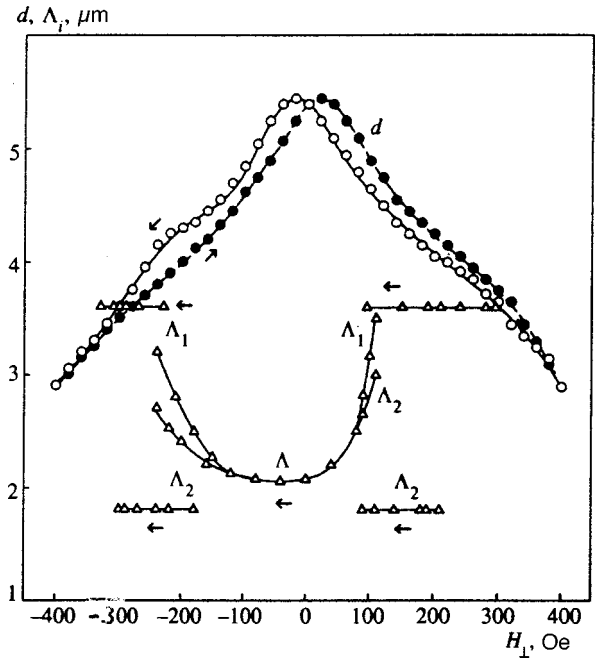

FIG. 2. Geometrical parameters of biperiodic stripe domain structures in type-II film No. 2 of thickness 16 μm versus the magnetic field $H_{\perp}$.

The distribution of the magnetization in domain structures of any type was nonuniform both along the thickness and in the developed planes of the films. However, a polarization microscope with high optical resolution and standard spatial and polarization filtering techniques could be used to obtain qualitative information about the variation of the profile of the thickness distribution of the magnetization and to distinguish the image elements needed to test the reliability of the working hypotheses.

An analysis of the behavior of the domain arrays and their corresponding diffraction patterns leads to the following conclusions.

1. In films of any thickness the harmonic instability of the domain wall profile in the initial simple stripe domain structure evolves independently into two developed surfaces (usually for different values of $H_{\perp}$) as the field $H_{\perp}$ is decreased, because the parameters of the near-surface layers of epitaxial films differ at the boundaries with the substrate and with free space (see, e.g., Refs. 39 and 40). The periods of the harmonic distortions on the two surfaces Λ_1 (at the film–free space interface) and Λ_2 (at the film–substrate interface) can differ substantially (more than twofold; see Fig. 2). A further decrease in the field $H_{\perp}$ causes the penetration depth of the surface distortions of the domain wall profile to increase, and they begin to influence one another. For films of thickness $L\lesssim 20\ \mu$m this mutual influence leads to total spa-

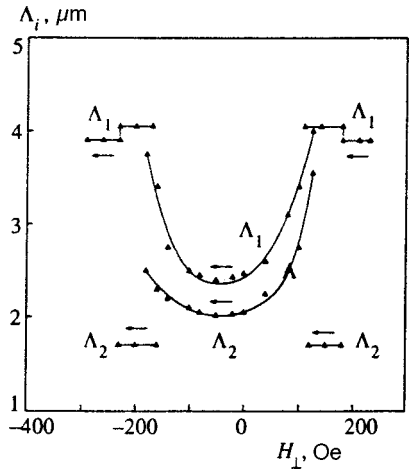

FIG. 3. Geometrical parameters of biperiodic stripe domain structures in type-II film No. 3 of thickness 26 μm versus the field $H_\perp$.

tial ''synchronization'' of the harmonic distortions: They acquire a common period ($\Lambda_1=\Lambda_2=\Lambda$), as shown in Fig. 2 and phase opposition, i.e., the phase of the function modulating the domain wall profile near one surface of the film differs by π from the phase of the analogous function near the other surface. In thicker films the increase in the penetration depth of the surface harmonic distortions of the domain wall profile does not induce their spatial synchronization ($\Lambda_1\neq\Lambda_2$), but the difference in the periods $\Delta\Lambda=\Lambda_1-\Lambda_2$ decreases significantly from its value for noninteraction; see Fig. 3. We note that one of the possible types of biperiodic domain structures—with antiphase modulation of the profile of the distribution of **M** in adjacent domain walls—is most likely stable only when there is no interaction between surface structures at the interfaces of the film with the substrate and with free space.

The amplitudes of the harmonic distortions on the two surfaces of the film differ, with more developed bends of the domain walls existing at the film–substrate interface. In the median plane of the film the domain walls remain vertical and even (unmodulated), as in a simple stripe domain structure. As the field $H_\perp$ is increased, the bends of the domain walls disappear first at the film–substrate interface and then at the free surface. In a finite field $H_\perp$ the peaks (and troughs) of the modulating functions in adjacent domain walls shift in opposite directions from their position for $H_\perp=0$. Moreover, the teeth of the ''saw'' in the same domain wall but on opposite sides of the film deviate in opposite directions under the influence of $H_\perp$. Reversing the sign of $H_\perp$ causes the slopes of the sawtooth domain walls to reverse directions. Effects such as these occur because the curved domain walls create periodic stay fields $\mathbf{H}_d$ with an alternating y-component above the film surfaces. When the stay fields interact with $\mathbf{H}_\perp$, the latter lengthens the parts of the domain walls for which $\mathbf{H}_{dy}\uparrow\uparrow\mathbf{H}_\perp$ and shortens the parts for which $\mathbf{H}_{dy}\uparrow\downarrow\mathbf{H}_\perp$.

2. The generation and disappearance of surface distortions of the domain wall profile and the variation of the periods Λ_1 and Λ_2 take place in a stripe domain structure whose period d depends on the orientation and magnitude of the external magnetic field. The smooth variation of the average period d of the stripe domain structure as the fields $H_\perp$ and $H_\parallel$ are varied is attributable to the generation, motion, or annihilation of magnetic dislocations; this fact was first brought to light by Palatnik and Lukashenko.[17] It is obvious that the smooth variation of the average domain wall modulation period Λ (or the modulation periods Λ_1 and Λ_2) should also be attributable exclusively to the presence of magnetic dislocations, which lift the prohibition imposed by translation invariance on continuous variation of the reciprocal lattice vectors in ideal periodic structures.[4)] We note that only the field component $H_\perp$ has a strong influence on d, Λ_1, and Λ_2; throughout the entire region of existence of stripe domain structures these parameters depend weakly on the field component $H_\parallel$, increasing slightly as $|H_\parallel|$ is increased.

3. Biperiodic domain structures in the region of stability can exist in several different modifications, which can be classified by symmetry. Such a classification, of course, cannot be made on the basis of the apparatus of three-dimensional space groups, because the finite thickness of the films prohibits translations in the direction of the normal to the surface (along the z axis), and the function $\mathbf{M}(z)$ does not have any kind of symmetry in general. For the purpose of classification, nonetheless, the apparatus of the theory of Shubnikov (black-and-white) two-dimensional space groups can be used to analyze the symmetry of the distribution of the vector **M** on either surface of the film (or in any plane in the interior of the film, parallel to these surfaces). This approach is admissible not only for stepped (black-and-white) distributions of the magnetic moment, when the domain walls are infinitely narrow, and $\mathbf{M}=\pm|\mathbf{M}|\mathbf{e}_z$, but also for ''gray'' domain structures with a smooth function $\mathbf{M}(\mathbf{r})$, because the symmetry operations that produce abstract black-and-white groups can act on objects of any nature.

4. Since the phase is any stable state of the system that differs in symmetry or degree of ordering from other possible stable states, the transitions induced by the external magnetic field between different modifications of biperiodic domain structures must be regarded as phase transitions. This means that the liquid-crystal phase of a magnet with biperiodic stripe domain structures is polymorphic and differs radically from the analogous phase with simple stripe domain structures.

5. The nuclei of new states in phase transitions between different modifications of biperiodic stripe domain structures are magnetic dislocations, which therefore bring to the initial domain array not only new spatial periods d and Λ, but also a new symmetry. In contrast, the magnetization reversal (change of sign of the projection of the vector **M** to the direction of the vector $\mathbf{H}_\perp$) of films having the investigated composition, observed for $\mathbf{H}_\perp<5$ Oe, involves dislocations whose motion in the weak-field region is inhibited by the coercive force and is fostered by processes of rotation (flopping) of the magnetization vectors.

We close this section with the observation that domain structures having monopolar domain walls occur in our investigated films with $\beta_u<1$ in zero field; the walls are characterized by the fact that throughout the entire volume of the film the angles of deviation of the magnetization vectors θ

$=\arccos(|\mathbf{M}|^{-1}(\mathbf{M}\cdot\mathbf{e}_0))$ from a certain designated direction with unit vector $\mathbf{e}_0\perp\mathbf{n}$ does not exceed the angle $\theta_{\max}<\pi/2$, which depends on β_u. Formally we have $\theta_{\max}\to 0$ in the limit $\beta_u\to 0$. In transition from one domain to the next the vectors $\mathbf{M}$ merely ''rock'' about the average position defined by the unit vector $\mathbf{e}_0$. In films having purely uniaxial magnetic anisotropy and the easy magnetization axis parallel to $\mathbf{n}$, all directions in a plane perpendicular to the normal are equivalent, so that the ''selection'' of the direction $\mathbf{e}_0$ depends on the history of the domain structure and how it is generated. If the domains are generated from the uniformly magnetized state as the result of a decrease in the field $\mathbf{H}_\perp^{(+)}$ (from the saturating value), the unit vector $\mathbf{e}_0$ will be parallel to the vector $\mathbf{H}_\perp^{(+)}$. This is also the situation when the field $\mathbf{H}_\perp^{(+)}$ is reduced to zero, and when $\mathbf{H}_\perp^{(+)}=0$ holds, the distribution $\mathbf{M}^{(+)}(\mathbf{r})$ is characterized by restriction of the allowed orientations of the vectors $\mathbf{M}$ to within a solid angle $\Omega^{(+)}$ smaller than 2π sr, with the axis along $\mathbf{H}_\perp^{(+)}$. The oppositely directed (relative to the initial) field $\mathbf{H}_\perp^{(-)}$ corresponds to the analogous [representing the distribution $\mathbf{M}^{(-)}(\mathbf{r})$] solid angle $\Omega^{(-)}<2\pi$, but with the axis in the opposite direction. The solid angles $\Omega^{(+)}$ and $\Omega^{(-)}$ are nonoverlapping, i.e., a smooth transition between the states $\mathbf{M}^{(+)}(\mathbf{r})$ and $\mathbf{M}^{(-)}(\mathbf{r})$ is impossible. Consequently, the limiting hysteresis loop with respect to the field $H_\perp$ must have a jump corresponding to the change of sign of the projection of the vector $\mathbf{M}$ onto the vector $\mathbf{H}_\perp$.

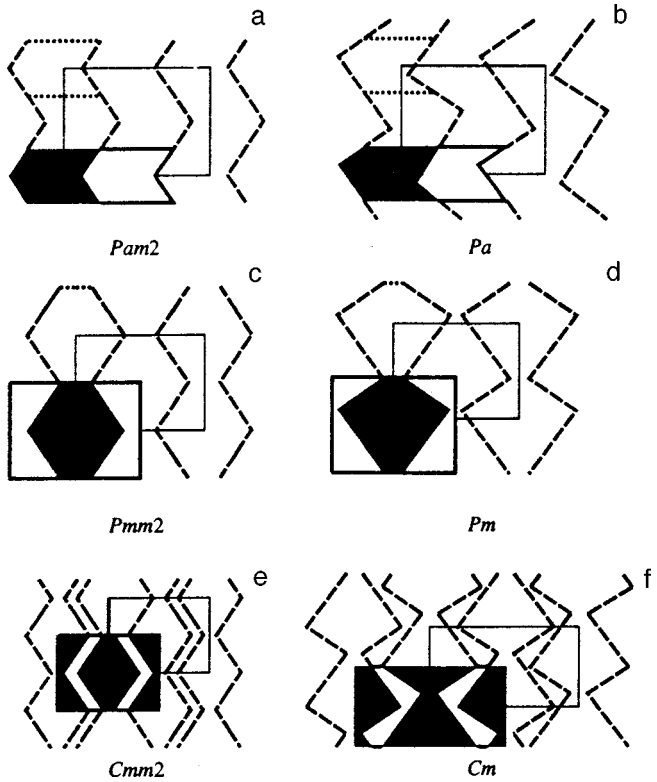

FIG. 4. Symmetry classification of the experimentally observed biperiodic domain structures: (a) symmetric in-phase; (b) asymmetric in-phase; (c) symmetric antiphase; (d) asymmetric antiphase; (e) bisymmetric double; (e) mirror-symmetric double.

4. SYMMETRY ANALYSIS OF BIPERIODIC STRIPE DOMAIN STRUCTURES

The results of the symmetry classification of the experimentally observed biperiodic domain structures are given in Fig. 4, which shows schematically the motif-forming elements (in the form of black-and-white figures outlined by a thick solid line), the domain walls (dashed lines), and the rectangular Bravais cells (thin solid lines). Auxiliary dotted lines are drawn in some figures to accentuate individual symmetry operations. The black-and-white representation of the motif-forming elements symbolically portrays the nonuniform distribution of the vector $\mathbf{M}$ on either of the two surfaces of the film (or in any plane parallel to them), black representing regions with one sign of the projection M_z (for any value of the latter) and white representing regions with the opposite sign of M_z; the dashed lines (domain walls) correspond to the set of points at which $M_z=0$. In this symbolic representation the symmetry of the biperiodic domain structures formed in the investigated films is described by one of six two-dimensional space groups of the rectangular system with a primitive (*Pam*2, *Pa*, *Pmm*2, *Pm*) or centered (*Cmm*2, *Cm*) Bravais cell. Structures with symmetry *Pa*, *Pm* and *Cm* can exist in dual modifications (not shown in the figure) corresponding to two antiparallel directions of the vector $\mathbf{H}_\perp$.

In the structure with symmetry *Pam*2 shown in Fig. 4a the distortions of the profile of the distribution of the magnetization vector $\mathbf{M}$ are in phase in all the domain walls on the given surface of the film[5)] and are symmetric; we therefore call this structure a symmetric in-phase biperiodic domain structure. This kind of profile of the magnetization distribution can occur only for $H_\perp=0$: Any arbitrarily small field $H_\perp$ will transform such a structure by a second-order phase transition into one of two modifications of an asymmetric in-phase (more precisely, quasi-in-phase) biperiodic domain structure with symmetry *Pa* (see Fig. 4b).

Certain zones of the region of stability of biperiodic domain structures on the $H_\perp H_\parallel$ plane, adjacent to the regions of stability of simple stripe domain structures, can also contain the structures shown in Figs. 4c and 4d, which we call (respectively) symmetric and antisymmetric antiphase biperiodic domain structures.

If the field $H_\parallel$ has a nonzero magnitude greater than a critical value, translations in the direction of the x axis ($\mathbf{e}_x\perp\mathbf{H}_\perp$) experience period doubling in certain intervals of the field $H_\perp$, and the structures shown in Fig. 4e or Fig. 4f are formed, which can be called (respectively) bisymmetric double and mirror-symmetric double biperiodic domain structures. Such domain arrays make up a sequence of alternating zones of in-phase and antiphase biperiodic domain structures (symmetric or asymmetric).

Photographs of actual observed domain structures in film No. 2 are shown in Fig. 5. The photographs in Fig. 5 represent biperiodic domain structures of the type: a) symmetric in-phase ($\mathbf{H}=0$, $d=5.5\,\mu$m, $\Lambda=2.1\,\mu$m); b) asymmetric in-phase ($H_\parallel=0$, $H_\perp=20$ Oe, $d=5.5\,\mu$m, $\Lambda=2.1\,\mu$m); c) symmetric antiphase ($H_\parallel=0$, $H_\perp=300$ Oe, $d=3.6\,\mu$m, $\Lambda=\Lambda_1=3.6\,\mu$m); d) mirror-symmetric double ($H_\parallel=300$ Oe, $H_\perp=100$ Oe, $d=13\,\mu$m, $\Lambda=2.7\,\mu$m.

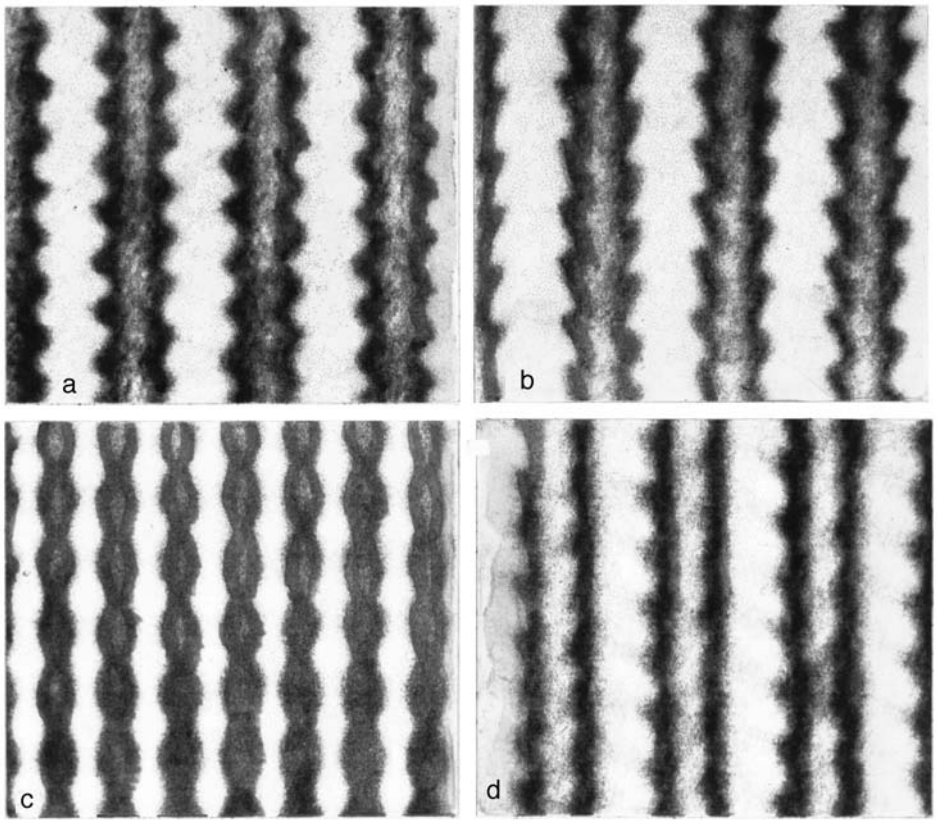

FIG. 5. Photographs of observed domain structures in film No. 2.

Experiments on the observation of magnetooptical light diffraction by two-dimensional domain arrays play a decisive role in the identification and symmetry classification of different modifications of the observed biperiodic domain structures. For an abstract two-dimensional biperiodic lattice with periods d and Λ along the x and y axes of a Cartesian coordinate system with the scattering of light incident on the film along the z axis ($\mathbf{e}_z \| \mathbf{n}$) diffraction maxima appear in the directions $\mathbf{k}_{(p,q)} = \mathbf{k} + p\mathbf{b}_1 + q\mathbf{b}_2$, where $\mathbf{k}$ is the wave vector of the primary light beam, p and q are integers, and $\mathbf{b}_1$ and $\mathbf{b}_2$ are the reciprocal lattice basis vectors, which are defined by the expressions $\mathbf{b}_1 = (2\pi/d)\mathbf{e}_x$ and $\mathbf{b}_2 = (2\pi/\Lambda)\mathbf{e}_y$. In general, the intensities $J_{p,q}$ of the various diffraction maxima depend both on the type and symmetry of the lattice and on the type and shape of its elements.

If circular magnetic dichroism is disregarded, the domain arrays represent pure phase diffraction gratings, which implement a periodic modulation of the angle of rotation of the polarization plane of the linearly polarized primary light beam as a result of the Faraday effect. The only components of the film-scattered light that contribute to the formation of the diffraction pattern are those in which the polarization plane is rotated from the primary beam through an angle $\pm\pi/2$ (the sign changes in each transition from a domain with $M_z > 0$ to a domain with $M_z < 0$), whereas components having the same polarization plane as the primary beam do not acquire a phase difference after passing through domains with $M_z > 0$ and $M_z < 0$ and so they do not interfere. It follows, therefore, that when light is scattered by the domain arrays, the diffracted radiation in all diffraction maxima of order $p \neq 0$ and $q \neq 0$ is polarized orthogonally relative to the principal ($p = 0$ and $q = 0$) maximum (see, e.g., Refs. 41 and 42). When light is scattered by a symmetric simple stripe domain structure, in which domains with $M_z > 0$ and $M_z < 0$ are of identical width ($d_+ = d_- = d/2$), the intensity of the even diffraction maxima becomes equal to zero.[6)] When the symmetry of a simple stripe domain structure is broken (e.g., by the application of a field $H_\|$), extinction of the even diffraction maxima no longer occurs.

Taking into account the symmetry of the experimentally observed structures and these properties of the magnetooptical diffraction of plane-polarized light by the domain arrays, we can determine the form of the diffraction patterns for various modifications of biperiodic stripe domain structures on the basis of fairly straightforward considerations. A symmetry approach similar to that used in x-ray structural analysis for the determination of Laue classes[44] shows that the possible types of optical diffraction patterns for the space groups *Pam*2, *Pa*, *Pmm*2, *Cmm*2, and *Cm* in the case of normal incidence on the film belong to the same symmetry point group $\underline{mm}\underline{2}$ (see, e.g., Ref. 45). For $d_+ = d_- = d/2$, where d_+ and d_- are the average widths of domains with $M_z > 0$ and $M_z < 0$, respectively, the extinction of maxima with $p = 2n$ and any q (other than the principal diffraction maximum) will be observed for in-phase symmetric biperiodic domain structures (Fig. 4a), and the extinction of maxima with $p = 2n+1$ and any q [except maxima with $(p, q) = (\pm 1, 0)$] will be observed for antiphase biperiodic domain structures (Fig. 4c). Indeed, if we hypothetically delineate bands containing modulated domain walls in each domain, then secondary Huygens–Fresnel sources situated in

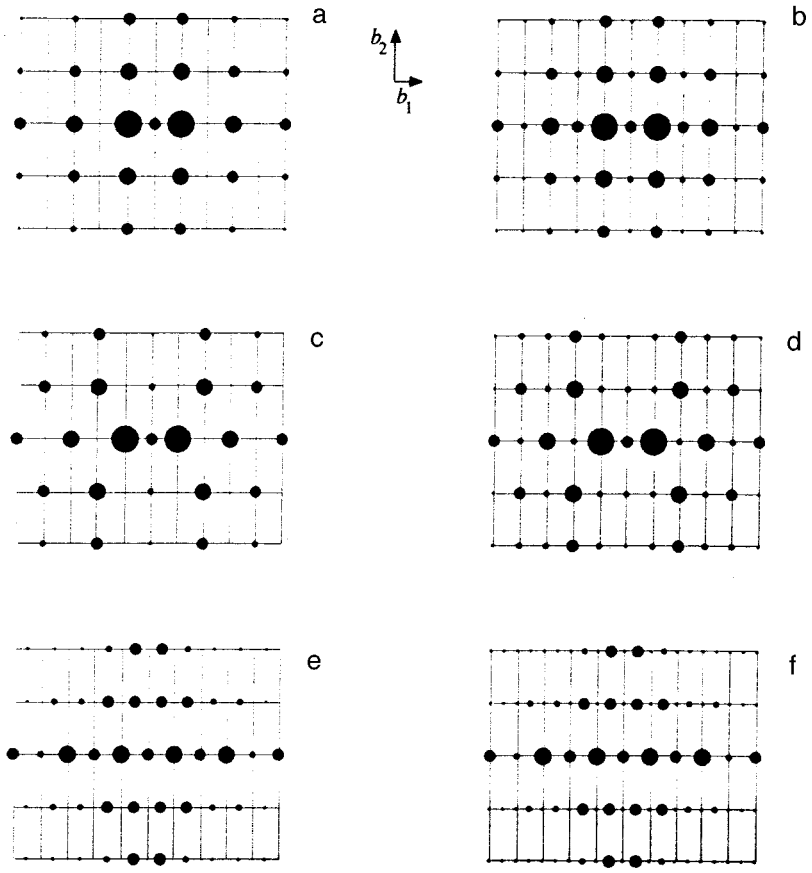

FIG. 6. Schematic representation of the observed diffraction maxima for biperiodic domain structures of various symmetries.

adjacent bands will be in phase in the first case and antiphased in the second case, resulting in the extinction of the corresponding maxima (see Figs. 6a and 6c). If the condition $d_+ = d_- = d/2$ does not hold, any diffraction maxima can be observed, as shown in Figs. 6b and 6d.

In light scattering by the structures shown in Figs. 4e and 4f the directions to the diffraction maxima are given by the relations $\mathbf{k}_{(p,q)} = \mathbf{k} + p\mathbf{b}_1' + q\mathbf{b}_2$, where $\mathbf{b}_1' \simeq \mathbf{b}_1/2$, and $\mathbf{b}_1$ is the corresponding reciprocal lattice vector for in-phase or antiphase ''nested dolls'' structures generating the given double biperiodic domain structure. For bisymmetric double biperiodic domain structures we observe the extinction of all diffraction maxima with $p = 2n$ and any $q \neq 0$ (Fig. 6e), where $|n| \geqslant 1$; such extinction does not occur for mirror-symmetric biperiodic structures (see Fig. 6f).

The diffraction patterns in Fig. 6, which are plotted solely on the basis of a symmetry analysis of the Bravais lattices and motif-forming elements, are qualitative in nature and are not presumed to accurately represent the relative intensities of the various diffraction maxima.

Photographic negatives of diffraction patterns actually observed in the experiments for film No. 2 are shown in Fig. 7. The photographs correspond to biperiodic domain structures of the following types: in-phase a) symmetric ($\mathbf{H} = 0$) and b) asymmetric ($H_\parallel = 0$, $H_\perp = 20$ Oe); c) double bisymmetric ($H_\parallel = 300$, $H_\perp = 100$ Oe); d) in-phase symmetric and double symmetric ($H_\parallel = 250$ Oe, $H_\perp = 60$ Oe); e) in-phase asymmetric and antiphase symmetric ($H_\parallel = 0$, $H_\perp = 200$ Oe). In the latter case spatial filtering has been used to eliminate ambient light (screening the central diffraction maximum and the lower part of the diffraction pattern). The pincushion distortions in the photographs are the result of curvature of the image field.

It follows from a comparison of Figs. 6 and 7 that the form of the actual observed diffraction patterns is consistent with the results of our qualitative symmetry analysis.

5. KINETICS OF PHASE TRANSITIONS IN BIPERIODIC STRIPE DOMAIN STRUCTURES

We consider the specific characteristics of phase transitions involving biperiodic domain structures during cyclic variation of the field $H_\perp$ with $H_\parallel = \text{const}$ in the case of film No. 2, which has a thickness of 16 μm and whose state diagram (for the range of fields $|H_\parallel| \leqslant 500$ Oe is shown in Fig. 8. The solid curves represent the boundaries at which the various magnetic states become unstable[7] for the descending branch of the limiting hysteresis loop (as the perpendicular field varies from $+H_\perp^{(\max)}$ to $-H_\perp^{(\max)}$, where $|H_\perp^{(\max)}| > 400$ Oe corresponds to magnetic saturation), and the dashed curves represent the same for the ascending branch, for which the curve numbers are primed ($1', 2', \ldots, 6'$). The symbols U↔S refer to the second-order phase transition curves between the uniformly magnetized state and a simple striped domain structure; the numbering of the other curves will be explained below. Individual curves that have no bearing on the ensuing discussion are not shown in the state diagram.

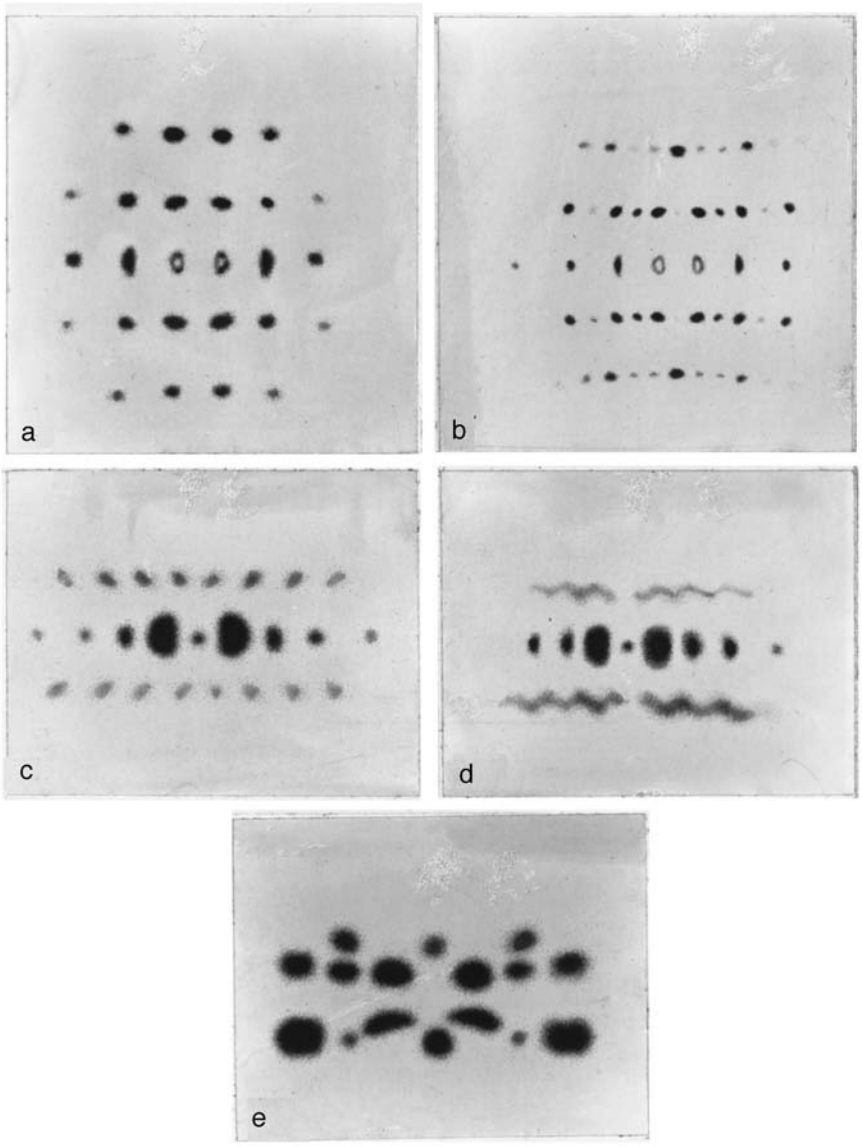

FIG. 7. Photographic negatives of experimentally observed diffraction patterns for film No. 2.

We first describe the evolution of the experimentally observed domain structures for film No. 2 in the case $H_{\parallel}=0$, $H_{\perp}=\text{var}$, which is shown in Fig. 2. At the very beginning a nearly symmetric (judging from the form of the observed diffraction patterns) antiphase biperiodic domain structure is formed (with the maximum possible modulation period of the domain wall profile for the given surface, $\Lambda_1=\Lambda_{1\max}$) at the interface between the film and free space from the region of a simple stripe domain structure when the field $H_{\perp}$ is gradually reduced to some critical value. This instability on the diagram of Fig. 8 corresponds to curves *1* and *1′*. In a slightly weaker field (curves *2* and *2′* in Fig. 8) a similar bending instability of the domain walls sets in at the film–substrate interface (with the minimum possible period for the given surface, $\Lambda_2=\Lambda_{2\min}<\Lambda_{1\max}$). In a certain range of the field $H_{\perp}$ the two surface structures exist independently and essentially have no influence on each other; the modulation periods of the domain wall profiles in both structures, Λ_1 and Λ_2 remain constant for all practical purposes. Then, owing to the increase in the penetration depth of modulation of the domain wall profiles, coupling arises between the surface biperiodic structures, transforming the antiphase symmetric biperiodic domain structure into one of the dual modifications of an in-phase asymmetric biperiodic domain structure, specifically the modification corresponding to the initial direction of the field $H_{\perp}$ (curves *3* and *3′* in Fig. 8). This process is accompanied by the onset of a very pronounced field dependence of the modulation periods Λ_1 and Λ_2, a sudden change and merging of their values (see Fig. 2), and a sharp increase in the amplitude of the functions modulating the profiles of the domain structures. Total spatial synchronization of the quasiharmonic distortions of the profile is then attained on both surfaces of the film ($\Lambda_1=\Lambda_2=\Lambda$, where $\Lambda_{1\min}<\Lambda<\Lambda_{2\max}$), and a single in-phase asymmetric biperiodic domain structure is formed (the corresponding curves are not shown in the diagram of Fig. 8), remaining stable down to zero value of the field $H_{\perp}$.

After the polarity of the field is changed and its magnitude is increased somewhat (to values no higher than 5 Oe for film No. 2), the magnetization of the film suddenly re-

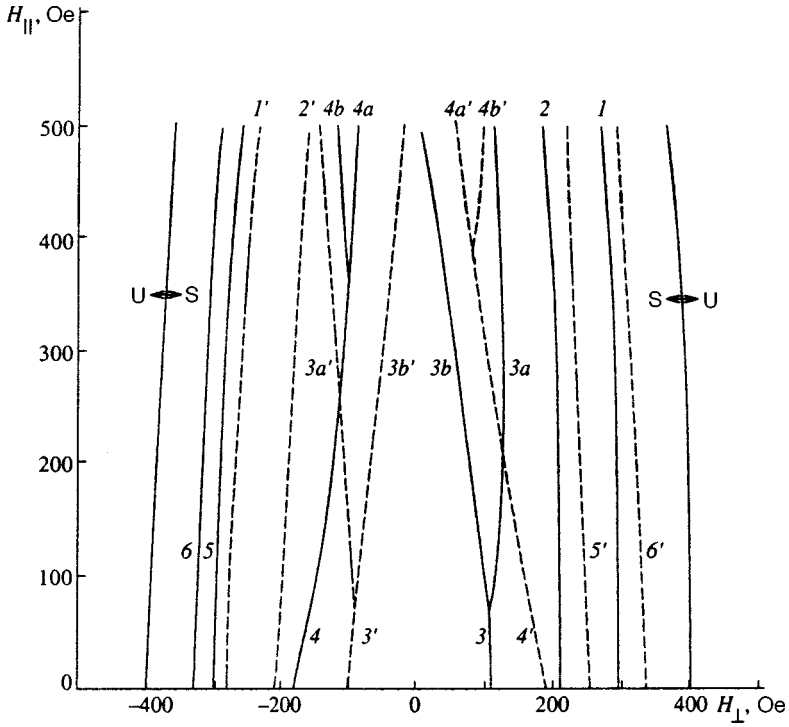

FIG. 8. State diagram of film No. 2 in the $H_{\perp}H_{\|}$ plane.

verses (i.e., the projection M_z changes sign in all domains to the original direction of the magnetization vector of the field $H_{\perp}$) with an accompanying transition from an in-phase asymmetric biperiodic domain structure to an in-phase symmetric structure.[8)] When the direction of the field $H_{\perp}$ is reversed (from the initial direction) and its magnitude is increased, all of the transitions described above occur similarly but in the reverse sequence, exhibiting the typical hysteresis of a first-order transition. Curves *4* and *5* (*4′* and *5′*) in Fig. 8 refer to (respectively) the generation and annihilation of an antiphase symmetric biperiodic domain structure at the free surface, and curve *6* (or *6′*) refers to the disappearance of this structure at the film–substrate interface, i.e., transition to a simple stripe domain structure. We note that only for transitions between the symmetric and asymmetric modifications of in-phase (symmetry groups *Pam*2 and *Pa*) and antiphase (symmetry groups *Pmm*2 and *Pm*) biperiodic domain structures do second-order transitions take place, but during the cyclic variation of $H_{\perp}$ they also occur in unequal intervals for different directions of the field, because the field dependence of the period *d* of the ''nested dolls'' stripe domain structure also exhibits hysteresis (see Fig. 2).

A single ''synchronized'' biperiodic domain structure is never formed in sufficiently thick films ($L>15\,\mu$m); two uncoupled systems of domain walls having a modulated profile and different modulation periods Λ_1 and Λ_2 but stripe domains of the same period *d* are observed over the entire range of existence of the structures. This effect is illustrated in Fig. 3a, which shows Λ_1 and Λ_2 as functions of $H_{\perp}$ for $H_{\|}=0$ in the case of the type-II film No. 3 of thickness $L=26\,\mu$m. One other detail distinctly emerges in this graph: the disappearance or onset of modulation at one of the surfaces of the film induces a jump in the graph of $\Lambda_i=f(H_{\perp})$ for the other surface. This phenomenon is attributable to the interrelationship of the parameters characterizing the biperiodic domain structure (*d*, Λ_1, and Λ_2), so that even in a fixed field the disappearance of domain wall modulation at one of the surfaces induces an abrupt change in the period of the stripe domain structure *d*, implying a change in the domain wall modulation period at the other surface of the film.

The symmetry of biperiodic domain structures generated in films having this composition and any thickness at the interface of the film with free space from a region with a simple stripe domain structure as the field $H_{\perp}$ is gradually diminished puts them in the class of antiphase asymmetric (with very little difference from symmetric—see Fig. 5c) biperiodic domain structures (with modulation of the domain wall profile Λ_1). The surface domain structures produced in a somewhat weaker field with the domain wall modulation period Λ_2 at the film–substrate interface have the same symmetry for any film thickness. We note that even for the thinnest films (with $L\simeq L^{*}_{cr}$) in which the existence of biperiodic domain structures is observed the $H_{\perp}$-induced transition from a simple to a biperiodic stripe structure does not take place throughout the entire bulk of the film all at once, but is localized near the interface of the film with free space. This means that the penetration depth of the ''frozen'' surface soft mode responsible for the given phase transition at the point of destabilization of the simple stripe domain structure for films of the investigated composition is less than $L^{*}_{cr}/2\simeq 2.5\,\mu$m.

The observed evolution of two-dimensional domain arrays $H_{\|}=0$ essentially remains unchanged in the presence of a finite field $H_{\|}$ as long as the magnitude of the latter does not exceed a critical value $H_{\|}^{(cr1)}$, which is approximately 70 Oe for film No. 2. For $H_{\|}>H_{\|}^{(cr1)}$ the chain of phase tran-

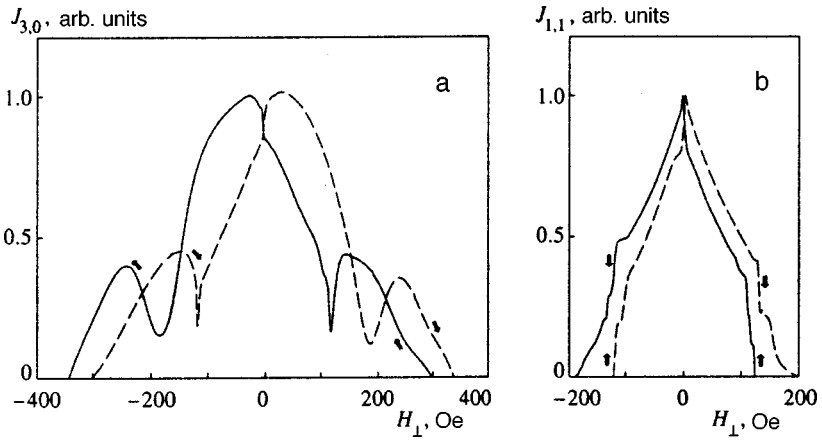

FIG. 9. Intensity of diffracted radiation for the maxima $J_{3,0}$ (a) and $J_{1,1}$ (b) versus the field $H_\perp$ at $H_\parallel = 0$ for film No. 2. The arrows on the curves indicate the direction in which the field is varied.

sitions accompanying cyclic magnetization reversal acquires new links; specifically, as $H_\perp$ is decreased, the transformation of antiphase biperiodic domain structures into in-phase biperiodic structures proceeds through ''buffer'' states corresponding to a mirror-symmetric double biperiodic domain structure (with period doubling relative to the adjacent antiphase or in-phase biperiodic domain structures, $d' = 2d$). In the diagram of Fig. 8 this process is manifested by splitting of each of the curves *3* and *3′* in two: *3a*, *3b* and *3a′*, *3b′*. An analogous effect is observed when $|H_\perp|$ is increased with $H_\parallel > H_\parallel^{(cr2)}$ (the second critical value[9)] of the field $H_\parallel$ for film No. 2 is ≈ 400 Oe) and is manifested in the state diagram by splitting of the curves *4* and *4′* in two: *4a*, *4b* and *4a′*, *4b′*.

The hypothesis of polymorphism of the biperiodic domain structures is also supported in experiments aimed at studying the dependence of the intensities of the diffraction maxima $J_{p,q}$ on the field $H_\perp$ for $H_\parallel = 0$. The $J_{3,0}(H_\perp)$ curves for film No. 2 (Fig. 9a) exhibit sharp (particularly for a decreasing field) troughs in the regions corresponding to domain walls between antiphase and in-phase biperiodic domain structures (cf. Fig. 2), along with jumps in the vicinity of zero-field points[10)] due to magnetization reversal of the film. The $J_{1,1}(H_\perp)$ curves in Fig. 9b exhibit similar features. A diffraction maximum of $J_{2,0}$ (see Fig. 10) is observed in the region of variation of $H_\perp$ corresponding to the zone of stability of antiphase biperiodic domain structures. A comparison with Sec. 4 shows that the laws encountered in the behavior of the experimental $J_{p,q}(H_\perp)$ curves are fully consistent with the qualitative conclusions based on our symmetry analysis of the diffraction process.

6. CONCLUSION

An analysis of the results of the reported investigations shows conclusively that several types of regular biperiodic stripe domain structures can exist in uniaxial magnetic films having a small positive anisotropy constant ($0 < \beta_u < 1$) with the easy magnetization axis directed along the normal to the surface in a certain interval of film thicknesses and in a certain region of variation of the direction and magnitude of the external magnetic field; the structures differ from one another in their symmetry and (or) the period of one of the uniaxial translations. The possible types of distributions of the magnetization vector in any plane parallel to the surface of the film for the generated biperiodic stripe domain structures belong, in terms of symmetry, to one of six two-dimensional space groups of the rectangular system with a primitive (*Pam*2, *Pa*, *Pmm*2, *Pm*) or centered (*Cmm*2, *Cm*) Bravais lattice. The various types of domain structures can be identified either by simple visual observation using a polarization microscope or by analysis of the optical diffraction patterns. As regards symmetry, the latter belong to the same point group *mm*2, but differ either in the intensity distribution of the diffracted radiation among the diffraction maxima $J_{p,q}$ of even and odd orders p and q or in the ratio of the magnitudes of the reciprocal lattice vectors $\mathbf{b}_1$ and $\mathbf{b}_2$.

When the magnitude or direction of the magnetic field vector is varied, first-order or second-order phase transitions take place between the different types of biperiodic stripe domain structures, bound or free magnetic dislocations playing the dominant role in the nucleation processes. The state of the magnet as a whole corresponds to the Berezinskiĭ–Kosterlitz–Thouless phase in the first case and to the liquid-crystal phase in the second case.

The region in which the various types of biperiodic stripe domain structures exist in films of thickness $L_{cr}^{(1)} > L$

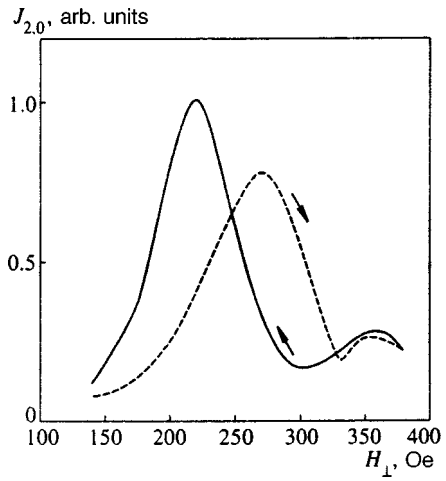

FIG. 10. Intensity of diffracted radiation for the maxima $J_{2,0}$ versus the field $H_\perp$ at $H_\parallel = 0$ for film No. 2. The arrows on the curves indicate the direction in which the field is varied.

$>L_{cr}^{*}$ on the $H_{\perp}H_{\parallel}$ plane for orientational phase transitions and $H_{\parallel}T$ for spontaneous phase transitions near the Curie (Néel) point is always contained inside the region in which simple (monoperiodic) stripe domain structures exist. This condition also enables us to draw conclusions as to the critical distributions of the magnetization vector in thicker films ($L>L_{cr}^{(1)}$), in which chains of floating cone-shaped domains emerge in the interior of each stripe domain of the biperiodic domain structures far from the corresponding phase transition lines for both surfaces of the magnet. In fact, since we have $L_{cr}>L_{cr}^{*}$, the region of stability of the domain structures containing conical domains on the state diagrams must be contained within the region of biperiodic stripe domain structures, i.e., direct transitions between the uniform state and states with floating conical domains do not occur. This assertion is true not only for magnets, but also for ferroelectrics, because the distribution of the electric polarization vector **P** in the latter obeys exactly the same laws as the distribution of the vector **M** in magnets (see, e.g., Refs. 46 and 47).

On the other hand, in the Russian and foreign scientific literature there is a rather prevalent opinion that structures containing floating conical domains are universal for sufficiently thick films, i.e., such structures are also stable in the immediate vicinity of the lines of spontaneous and orientational second-order phase transitions. Theoretical calculations utilizing the Ritz method have been published and predict a monotonic decrease in the critical thickness $L_{cr}^{(1)}$ as the second-order phase transition lines are approached and, hence, an increase in the probability of the existence of structures containing floating conical domains (see, e.g., Ref. 48). Such calculations are based on totally inadequate models employing the concepts of geometrical (unstructured) domain walls and uniformly magnetized domains. In real stripe domain structures, however, the boundaries between the domains are structured entities of finite extent with a vortexlike distribution of the transverse components of the magnetization vector **M** in them, even far from the phase transition lines.[49] All the same, if the half-period of the domain structure $d/2$ is much greater than the width of the domain walls δ (as is true only far from second-order phase transition lines), the distribution of the magnetization in the domains is essentially uniform, and the *z*-component of the demagnetizing field $H_{dz}=\pm 4\pi M_z$ is dominant at the surfaces of the film. Only under these conditions does the logic behind the reasoning of advocates of the universality of structures containing floating conical domains in thick films hold up without fault. As second-order phase transition lines are approached, the vortices inside the domain walls grow in volume and absorb regions with a uniform distribution of the vector **M**, so that the distribution of the transverse components of the magnetization vector in the film becomes highly nonuniform throughout the entire volume of the film.[49] In this case the following relations hold in the immediate vicinity of the second-order phase transition lines for thick films ($d_c/L=\epsilon\ll 1$, where d_c is the critical period of the domain structure):[35,46,47]

$$M_z \simeq A\cos\frac{(\pi-\epsilon)z}{L}\cos\frac{2\pi x}{d_c}, \quad H_{dz}\simeq -\frac{\pi\epsilon^2}{1+4\pi/\beta_u}M_z,$$

$$M_x \simeq \frac{2\pi\epsilon}{1+4\pi/\beta_u}A\sin\frac{(\pi-\epsilon)z}{L}\sin\frac{2\pi x}{d_c},$$

$$H_{dx}\simeq M_x\beta_u.$$

Here A is the order parameter, which vanishes at phase transition lines. It follows from these equations that the ratio of the amplitudes of the demagnetizing fields H_{dz} and H_{dx} is equal to $\epsilon/(2\beta_u)$, i.e., in the vicinity of second-order phase transition lines at the surfaces of the film the *x*-component of the demagnetizing field is dominant, and the *z*-component, which is responsible for the onset of floating conical domains, becomes negligible.

The authors are grateful to L. I. Antonov and I. E. Dikshteĭn for a discussion of the paper and for helpful consultations.

This work has received financial support from the Russian Fund for Fundamental Research (Project Code 96-02-16082-a).

*)E-mail: lisf@dataforce.net

1)We have intentionally chosen the terms ''biperiodic *stripe* domain structures'' to emphasize the distinction from biperiodic two-dimensional arrays of bubble domains of various shapes and symmetries, which are not discussed in this paper. The simplest example of such arrays are hexagonal lattices of circular bubble domains; more intricate configurations encountered in dynamic self-organization under the influence of a pulsating magnetic field are described, for example, in Refs. 7 and 8.

2)These estimates should be regarded with caution, because they are valid only for films with $\beta_u \ll 1$.

3)Phase transition lines associated with a hexagonal lattice of bubble domains are not shown in the diagram, as they do not have any direct bearing on the problems discussed here; for the complete state diagram see Ref. 33.

4)Indeed, for an infinite stripe domain or domain of finite length with pinned ends, states with arbitrarily close but different values of Λ are separated from each other by an energy barrier, and a smooth transition from one state to another is impossible. But the length of a magnetic dislocation can vary continuously under the influence of a magnetic field, lifting the prohibition against a smooth variation of Λ. The dislocation-generated ''deformations'' of the magnetostatic field are long-range effects and, in turn, modify the values of Λ in a sufficiently large neighborhood of the dislocation core. Moreover, the onset, change of symmetry, and annihilation of near-surface distortions of the domain wall profile produce insignificant variations of the period of the ''nested dolls'' stripe domain structure.

5)At the other surface of the film they are also in phase, but antiphase relative to the distortions of the domain wall profile on the first surface (if the distortions at the two surfaces are tightly coupled; see Sec. 3).

6)In sufficiently thick films, owing to multiple interference, even diffraction maxima can also be observed for $d_+=d_-=d/2$, but their polarization plane coincides with the polarization plane of the principal diffraction maximum.[43] Since the transmission plane of the analyzer in experiments on the observation of magnetooptical diffraction is always made orthogonal to the transmission plane of the input polarizer, the presence of even diffraction maxima associated with multiple interference can be disregarded in all the discussion to follow.

7)For several reasons the magnetic dithering method, which is necessary in order to establish thermodynamic-equilibrium initial states for different types of domain structures, could not be used in the experiments, so that the diagram in Fig. 8 refers to metastable states for which the positions of the boundaries of their destabilization depends on the prior history.

8)The observed abruptness of the magnetization reversal in weak fields in all the investigated films does not imply abruptness of the transition between asymmetric and symmetric in-phase biperiodic structures. Once created, a symmetric in-phase periodic domain structure makes a smooth transition to both dual modifications of an asymmetric in-phase structure under the

influence of a field $H_\perp$ in the corresponding direction, i.e., the phase transition described here is second-order, corroborating the results of the symmetry analysis (the group *Pa* is a subgroup of the group *Pam*2).

[9)]For thermodynamic-equilibrium mirror-symmetric double biperiodic domain structures, the bifurcation points of curves *3* and *4′* (also *4* and *3′* must coincide and be situated on the abscissa axis. In quasistatic magnetization reversal of the films a thermodynamic equilibrium state does not occur, so that the field $H_\parallel$ acquires critical values, which differ from each other depending on whether the field $|H_\perp|$ is increasing or decreasing.

[10)]The fact that the magnetization reversal of films of the investigated composition takes place in very weak fields ($|H_\perp|<5$ Oe) has been further corroborated by additional experiments performed by us on optical mode conversion in the waveguide propagation of light (with a wavelength of 1.15 μm) in the films and by experiments on the attenuation of magnetostatic surface waves (at a frequency of 1 GHz), which are known to be nonreciprocal with respect to the direction of the magnetization vector.

[1] W. S. Paxton and T. G. Nilan, J. Appl. Phys. **26**, 994 (1955).
[2] H. Williams and R. Sherwood, in *Magnetic Properties of Metals and Alloys* [Russian translation], Izd. Inostr. Lit., Moscow (1961).
[3] J. Kaczer and R. Gemperle, Czech. J. Phys., Sect. B **11**, 157 (1961).
[4] G. S. Kandaurova, Izv. Vyssh. Uchebn. Zaved. Fiz. **5**, 12 (1964).
[5] L. S. Palatnik, L. I. Lukashenko, Yu. V. Zolotnitskiĭ *et al.*, Zh. Éksp. Teor. Fiz. **59**, 1177 (1970) [Sov. Phys. JETP **32**, 643 (1970)].
[6] S. Chikazumi, *Physics of Ferromagnetism: Magnetic Characteristics and Practical Applications*, 2nd ed. (Clarendon Press, Oxford-New York, 1997).
[7] F. V. Lisovskiĭ, E. G. Mansvetova, E. P. Nikolaeva *et al.*, Zh. Éksp. Teor. Fiz. **103**, 213 (1993) [JETP **76**, 116 (1993)].
[8] F. V. Lisovskiĭ and E. G. Mansvetova, JETP Lett. **58**, 784 [836] (1993).
[9] J. Kaczer, M. Zeleny, and P. Suda, Czech. J. Phys., Sect. B **13**, 579 (1963).
[10] Ya. Katser, Zh. Éksp. Teor. Fiz. **46**, 1787 (1964) [Sov. Phys. JETP **19**, 1204 (1964)].
[11] L. S. Palatnik, L. I. Lukashenko, and A. G. Ravlik, Fiz. Tverd. Tela (Leningrad) **7**, 2829 (1965) [Sov. Phys. Solid State **7**, 2285 (1965)].
[12] R. Gemperle and A. Gemperle, Phys. Status Solidi **26**, 207 (1968).
[13] B. Wyslocki, Acta Phys. Pol. **34**, 327 (1968).
[14] A. Hubert, Z. Angew. Phys. **26**, 35 (1968).
[15] L. S. Palatnik and L. I. Lukashenko, Dokl. Akad. Nauk SSSR **192**, 1268 (1970) [Sov. Phys. Dokl. **15**, 622 (1970)].
[16] J. Kaczer, IEEE Trans. Magn. **MAG-6**, 442 (1970).
[17] L. S. Palatnik and L. I. Lukashenko, Fiz. Met. Metalloved. **29** 782 (1970).
[18] G. S. Kandaurova and L. G. Onoprienko, *Fundamental Problems of the Theory of Magnetic Domain Structure* [in Russian], Izd. UrGU, Sverdlovsk (1977).
[19] G. S. Kandaurova and Ya. S. Shur, Izv. Akad. Nauk SSSR, Ser. Fiz. **30**, 1030 (1966).
[20] Ya. S. Shur and G. S. Kandaurova, Fiz. Met. Metalloved. **23** 627 (1967).
[21] G. S. Kandaurova, Fiz. Tverd. Tela (Leningrad) **10**, 2311 (1968) [Sov. Phys. Solid State **10**, 1818 (1968)].
[22] L. I. Koshkin, T. I. Nestrelyaĭ, and T. A. Dunaeva-Mitlina, Fiz. Tverd. Tela (Leningrad) **11**, 1216 (1969) [Sov. Phys. Solid State **11**, 987 (1969)].
[23] L. I. Koshkin, T. I. Nestrelyaĭ, V. P. Gavrilin *et al.*, Izv. Akad. Nauk SSSR, Ser. Fiz. **34**, 1256 (1970).
[24] G. A. Jones, T. M. Lacey, I. B. Puchalska, J. Appl. Phys. **53**, 7870 (1982).
[25] J. Goodenough, Phys. Rev. **102**, 356 (1956).
[26] A. Hubert, Phys. Status Solidi **24**, 669 (1967).
[27] R. Szymczak, J. Appl. Phys. **39**, 875 (1968).
[28] R. Gemperle and N. Zeleny, Phys. Status Solidi **26**, 839 (1968).
[29] R. M. Grechishkin, Yu. M. Zubkov, and D. I. Sementsov, Pis'ma Zh. Tekh. Fiz. **15**(9), 45 (1989) [Sov. Tech. Phys. Lett. **15**, 349 (1989)].
[30] R. M. Grechishkin, Yu. M. Zubkov, D. I. Sementsov *et al.*, Kristallografiya **36**, 1484 (1991) [Sov. Phys. Crystallogr. **36**, 842 (1991)].
[31] H. Tamada, M. Kaneko, and T. Okamoto, J. Appl. Phys. **64**, 554 (1988).
[32] H. Huahui, S. Jan, D. Hanming *et al.* IEEE Trans. Magn. **25**, 3497 (1989).
[33] F. V. Lisovskiĭ, E. G. Mansvetova, and Ch. M. Pak, Zh. Éksp. Teor. Fiz. **111**, 283 (1997) [JETP **84**, 156 (1997)].
[34] F. V. Lisovskiĭ, E. G. Mansvetova, and Ch. M. Pak, JETP Lett. **60**, 136 (1994).
[35] I. E. Dikshteĭn, F. V. Lisovskiĭ, E. G. Mansvetova *et al.*, Zh. Éksp. Teor. Fiz. **86**, 1473 (1984) [Sov. Phys. JETP **59**, 863 (1984)].
[36] V. L. Berezinskiĭ, Zh. Éksp. Teor. Fiz. **59**, 907 (1970) [Sov. Phys. JETP **32**, 493 (1970)].
[37] V. L. Berezinskiĭ, Zh. Éksp. Teor. Fiz. **61**, 1144 (1971) [Sov. Phys. JETP **32**, 610 (1971)].
[38] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[39] J. E. Davies, E. A. Giess, and J. D. Cupstis, Mater. Res. Bull. **10**, 65 (1975).
[40] R. D. Henry and E. C. Whitcomb, Mater. Res. Bull. **10**, 681 (1975).
[41] B. Kuhlow, Optik **53**, 115 (1979).
[42] B. Kuhlow, Optik **53**, 149 (1979).
[43] B. Kuhlow and M. Lambeck, J. Magn. Magn. Mater. **4**, 337 (1977).
[44] H. Lipson and W. Cochran, *The Determination of Crystal Structure*, 3rd ed., Bell, London (1966).
[45] F. V. Lisovskiĭ, E. G. Mansvetova, and Ch. M. Pak, Zh. Éksp. Teor. Fiz. **108**, 1031 (1995) [JETP **81**, 567 (1995)].
[46] V. V. Tarasenko, E. V. Chenskiĭ, and I. E. Dikshteĭn, Zh. Éksp. Teor. Fiz. **70**, 2178 (1976) [Sov. Phys. JETP **43**, 1136 (1976)].
[47] E. V. Chenskiĭ and V. V. Tarasenko, Zh. Éksp. Teor. Fiz. **83**, 1089 (1982) [Sov. Phys. JETP **56**, 618 (1982)].
[48] W. A. Barker and G. A. Gehring, J. Phys. C: Solid State Phys. **19**, 259 (1986).
[49] L. I. Antonov, S. V. Zhuravlev, E. V. Lukasheva *et al.*, Fiz. Met. Metalloved. **74**, 23 (1992).

Translated by James S. Wood

Field dependence of Young's modulus in a gadolinium single crystal

V. Yu. Bodryakov

Ural State Technical University (UPI), 620002 Ekaterinburg, Russia

V. M. Zverev*)

P. N. Lebedev Physical Institute, Russian Academy of Sciences, 117924 Moscow, Russia

S. A. Nikitin

Moscow State University (M. V. Lomonosov), 119899 Moscow, Russia
(Submitted 13 April 1998; resubmitted 15 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 2111–2121 (December 1998)

The field dependence of Young's modulus along the hexagonal *c* axis is measured in a gadolinium single crystal over a wide range of temperatures and magnitudes of the magnetic field aligned with the *c* axis. It is found that the isotherms of the field dependence of the ΔE effect in gadolinium are well approximated by a linear dependence on the square of the magnetization in both strong and weak strong magnetic fields, and also above and below the spin reorientation temperature. It is shown that the experimental trends obtained near the ferromagnetic transition can be interpreted within the approach based on the Landau theory of second-order phase transitions. The parameters of such an approach are determined for gadolinium on the basis of the experimental data. © *1998 American Institute of Physics.* [S1063-7761(98)01512-1]

1. INTRODUCTION

Studies of magnetoelastic interactions in rare-earth (RE) metals are of great significance for understanding the nature of magnetic phase transitions in these metals and also for practical applications.[1] One effective method for investigating magnetoelastic interactions in rare-earth metals is to examine the peculiarities of the temperature and field dependence of the elastic moduli and internal friction in the region of magnetic phase transitions. Such studies make it possible not only to determine important experimental parameters of rare-earth metals, but also to verify the applicability of theoretical approaches to the description of magnetoelastic phenomena in these metals (see, e.g., Ref. 2). Although a substantial amount of experimental material has been accumulated on the elastic properties and internal friction of rare-earth metals (see, e.g., Refs. 3–8), the field dependence of anomalies of the elastic properties and internal friction in the region of the magnetic phase transitions has not been examined in sufficient detail and has not received an adequate theoretical explanation.

In particular, in regard to gadolinium, detailed isotherms of the field dependence of Young's modulus $C_{33}(T,H)$ have been measured over a wide temperature range, but only in a geometry in which the magnetic field H is oriented in the basis plane.[7] In this case a complicated field dependence was obtained for the modulus $C_{33}(T,H)$, which has still not been interpreted theoretically. This is apparently because of the difficulty of calculating the field-dependent contribution of the domain structure to the modulus $C_{33}(T,H)$ in the indicated geometry.

The influence of a magnetic field directed along the hexagonal axis of a gadolinium single crystal on the temperature dependence of Young's modulus $C_{33}(T,H)$ was investigated in Ref. 3 in a comparatively narrow temperature range below the spin reorientation temperature. References 9 and 10 were dedicated to an interpretation of the observed trends. These studies, first of all, did not examine the field dependence of Young's modulus $C_{33}(T,H)$ in the indicated geometry in the temperature range above the spin reorientation temperature and, second, they did not determine the experimental dependence of the field isotherms of Young's modulus $C_{33}(T,H)$ on the square of the magnetization above or below the spin reorientation temperature. The latter is especially important because it represents a convenient means of quantitatively comparing the experimental trends with the theoretical predictions and thereby opens up the possibility of determining the parameters of the theory on the basis of such a comparison.

Gadolinium is a convenient test case for a theoretical study of the magnetic contribution to the elastic moduli in the case when the magnetic field is directed along the hexagonal *c* axis. This is because the spontaneous magnetization vector in a gadolinium single crystal is directed along this axis in the temperature interval $T_{SR}<T<T_C$, where $T_{SR}\simeq 230$ K is the spin reorientation temperature and $T_C=293$ K is the Curie temperature (see, e.g., Ref. 11). The superposition of a comparatively weak magnetic field $H>0.5$ kOe parallel to the hexagonal axis means that in this temperature range the paraprocess alone determines the variation of the magnetization and of the magnetic contribution to the elastic moduli. This circumstance allows us to neglect the contribution of the domain structure to the elastic moduli, which substantially simplifies the theoretical approach used to describe the temperature and field dependence of the elastic moduli of gadolinium near the Curie temperature. Below, we

base such an approach on the phenomenological Landau theory of second-order phase transitions and the method developed in Ref. 12, which are free of model-dependent notions about the nature of ferromagnetism and magnetoelasticity.

The aim of the present paper is to measure the field dependence of Young's modulus $E(T,H)=C_{33}(T,H)$ along the hexagonal c axis in a gadolinium single crystal over a wide range of temperatures and magnetic field strengths with the magnetic field aligned with the c axis, to establish the experimental dependence of the field isotherms of the ΔE effect on the square of the magnetization, to interpret the observed trends near the ferromagnetic transition within the framework of an approach based on the Landau theory of second-order phase transitions, and to determine the parameters of such an approach for gadolinium on the basis of the experimental results.

2. EXPERIMENTAL TECHNIQUE AND RESULTS OF MEASUREMENTS

To solve the problems formulated above, we performed detailed measurements of Young's modulus $E(T,H)$ along the hexagonal c axis in a gadolinium single crystal at the acoustic frequencies 1–3 kHz over a wide range of temperatures $80\leqslant T\leqslant 350$ K and magnetic field strengths $H\leqslant 12.9$ kOe with the magnetic field directed along the c axis. The modulus was measured by the method described in Ref. 13, which employs an electrostatic means of exciting natural vibrations in the acoustic frequency range in a console-mounted sample having the shape of a thin rod.

The experimentally measured values of the temperature dependence of Young's modulus $E(T,H)$ are plotted in Fig. 1 for several values of the magnetic field. The straight line $E_0(T)$ in this figure is an extrapolation of Young's modulus according to a linear temperature law from the paramagnetic region into the temperature range below T_C neglecting magnetic phase transitions. The basis of this extrapolation is the experimentally established linear temperature dependence of Young's modulus $E_0(T)$ in the paramagnetic state above the Curie temperature. The zig-zag curve (curve *1*) in the figure, obtained in the absence of a magnetic field, agrees with the results of Refs. 3–6, according to which anomalies in the temperature dependence of the modulus $E(T,0)$ taking place in the vicinity of the temperatures $T_C\simeq 293$ K and $T_{SR}\simeq 230$ K correspond to transition from the paramagnetic state to the ferromagnetic state and to the spin reorientation transition. In this case, near the Curie temperature a well-defined negative ΔE effect takes place, i.e., $\Delta E(T,0)=E(T,0)-E_0(T)<0$. Attention should also be directed at the significant difference in the magnitudes of the anomaly of the modulus $E(T,0)$ near the spin reorientation temperature which is apparent from a comparison of curve *1* in Fig. 1 and Fig. 1 of Ref. 7. This difference may be due to a difference in the purity of the samples used in the experiments. As was noted in Ref. 7, in superpure gadolinium samples the anomaly of the modulus $E(T,0)$ near the spin reorientation temperature is smaller.

In a finite magnetic field the data depicted in Fig. 1 in the temperature range $T>T_{SR}$ complement the low-temperature results presented in Fig. 6 of Ref. 3 and the results plotted in Fig. 2 of Ref. 6, obtained for one value of the magnetic field, $H=25$ kOe. According to Fig. 1, with increase of the magnetic field the anomalies in the temperature dependence of $E(T,H)$ near T_C and T_{SR} smooth out while the magnitude of the modulus $E(T,H)$ grows substantially. In this case, a magnetic field $H\simeq 1.1$ kOe liquidates the negative ΔE effect near the Curie temperature. In the temperature range $T_{SR}<T<T_C$ the indicated trends belong to the paraprocess since, according to our measurements, the contribution of the domain structure to Young's modulus taking place in the indicated geometry in weak fields $H<0.5$ kOe is less than 0.1 %.

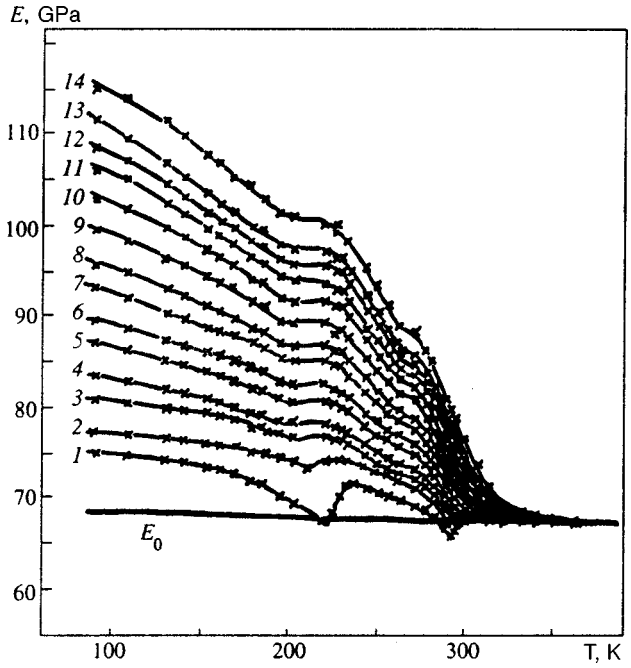

FIG. 1. Temperature dependence of Young's modulus $E(T,H)$, measured along the hexagonal c axis in a gadolinium single crystal for different values of the magnetic field $\mathbf{H}\|\mathbf{c}$ (in kOe): *1* — 0, *2* — 0.3, *3* — 1.1, *4* — 1.5, *5* — 2.3, *6* — 2.9, *7* — 3.8, *8* — 4.5, *9* — 5.7, *10* — 6.9, *11* — 8.0, *12* — 9.0, *13* — 10.2, *14* — 12.9. The straight line $E_0(T)$ is an extrapolation of the value of Young's modulus from the paramagnetic state.

To investigate the field dependence of Young's modulus $E(T,H)$ in gadolinium in more detail and establish qualitatively new regularities. It is of interest to experimentally investigate the dependence of the change in Young's modulus in a magnetic field $\Delta E(T,H)=E(T,H)-E(T,0)$ as a function of the square of the magnetization, $M^2(T,H)$, both above and below the spin reorientation temperature. Figure 2 plots the results of such a study, i.e., the isotherms of the field dependence of the ΔE effect in gadolinium along the hexagonal c axis as a function of the square of the magnetization, $M^2(T,H)$, at several temperatures. To construct such isotherms we used additional data on the magnetization in a field of the same sample on which Young's modulus was measured. Analysis of Fig. 2 allows us to formulate two important results. First, as can be readily seen, the experimental data lie very neatly along straight lines both in the region of comparatively weak magnetic fields (lines *3–14*),

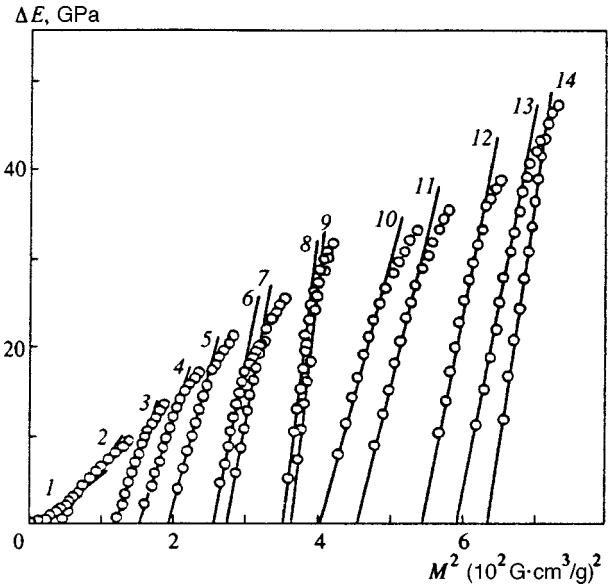

FIG. 2. Isotherms of the field dependence of the ΔE effect along the c axis in a gadolinium single crystal in a magnetic field $\mathbf{H}\|\mathbf{c}$ as a function of the square of the magnetization $M^2(T,H)$ at different temperatures T (in K): *1* — 300, *2* — 289, *3* — 278, *4* — 271, *5* — 261, *6* — 251, *7* — 239, *8* — 223, *9* — 218, *10* — 192, *11* — 173, *12* — 150, *13* — 120, *14* — 81.

where the change in the square of the magnetization $M^2(T,H)-M^2(T,0)$ induced by the field is small in comparison with the square of the spontaneous magnetization, $M^2(T,0)$, at a given temperature, and in the region of strong magnetic fields, where the magnetization is determined mainly by the field (lines *1* and *2*). Second, and especially important, the isotherms of the field dependence of the ΔE effect in gadolinium are well approximated by a linear dependence on the square of the magnetization both above and below the spin reorientation temperature. Thus, a definite similarity is observed in the trends of the field dependence of the ΔE effect in gadolinium as a function of the square of the magnetization in the temperature intervals corresponding to the two different phase transitions.

In the following sections we quantitatively analyze the experimentally observed trends in the field dependence of the ΔE effect in gadolinium near the ferromagnetic transition within the framework of the approach based on the Landau theory of second-order phase transitions, and determine the parameters of this approach for gadolinium from an analysis of the obtained experimental data.

3. TEMPERATURE AND FIELD DEPENDENCE OF THE ELASTIC MODULI OF A FERROMAGNET NEAR THE CURIE TEMPERATURE

In order to analyze the temperature and field dependence of the elastic moduli of a ferromagnet according to the Landau theory of second-order phase transitions with the accuracy needed here, it is necessary to go beyond the approximations of Refs. 12 and 14 and use an expansion for the free energy density in the form (see, e.g., Ref. 15)

$$F_M(T,M,\hat{u})=F_0(T,\hat{u})+\frac{1}{2}\Big\{\alpha(\hat{u})[T-T_C(\hat{u})]+\frac{1}{2}\alpha_2(\hat{u})[T-T_C(\hat{u})]^2\Big\}M^2+\frac{1}{4}\{b(\hat{u})+\alpha_3(\hat{u})[T-T_C(\hat{u})]\}M^4+\frac{1}{6}a_5(\hat{u})M^6, \quad (1)$$

where T is the temperature, M is the magnetization, and $\hat{u}$ is the strain tensor.

Expression (1) corresponds to the geometry of the experiment described above near the Curie temperature, when the magnetic field H and the magnetization M are directed along the hexagonal c axis of the gadolinium single crystal. Relation (1) yields the following expression for the elastic tensor (tensor of elastic moduli) of a ferromagnet in a constant magnetic field:

$$C_{ij}^H(T,H)=C_{ij}^0(T)-\frac{C_{ij}^{(1)}}{1+\xi(T,H)}+\frac{1}{1+\xi(T,H)}\times\left[C_{ij}^{(2)}+\frac{C_{ij}^{(3)}}{1+\xi(T,H)}\right]\left(1-\frac{T}{T_C}\right)+\left[C_{ij}^{(4)}+\frac{C_{ij}^{(5)}}{1+\xi(T,H)}+\frac{C_{ij}^{(6)}}{[1+\xi(T,H)]^2}\right]\times\frac{b}{\alpha T_C}M^2(T,H). \quad (2)$$

Here $C_{ij}^0(T)=(\partial^2F_0/\partial u_i\partial u_j)_T$ is the elastic tensor of the paramagnetic state in the absence of a magnetic field and $C_{ij}^{(n)}$ are the magnetoelastic tensors (tensors of the magnetoelastic coefficients), for which relation (1) yields

$$C_{ij}^{(1)}=\frac{(\alpha T_C)^2}{2b}\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln T_C}{\partial u_j},$$

$$C_{ij}^{(2)}=-\frac{(\alpha T_C)^2}{2b}\left[\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln\alpha}{\partial u_j}+\frac{\partial\ln T_C}{\partial u_j}\frac{\partial\ln\alpha}{\partial u_i}-\frac{2\alpha_2T_C}{\alpha}\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln T_C}{\partial u_j}\right],$$

$$C_{ij}^{(3)}=-\left(\frac{\alpha^2\alpha_3T_C^3}{2b^2}\right)\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln T_C}{\partial u_j},$$

$$C_{ij}^{(4)}=-\frac{(\alpha T_C)^2}{2b}\left[\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln\alpha}{\partial u_j}+\frac{\partial\ln T_C}{\partial u_j}\frac{\partial\ln\alpha}{\partial u_i}+\frac{1}{T_C}\frac{\partial^2T_C}{\partial u_i\partial u_j}-\frac{\alpha_2T_C}{\alpha}\frac{\partial\ln T_C}{\partial u_i}\frac{\partial\ln T_C}{\partial u_j}\right],$$

$$C_{ij}^{(5)}=\frac{(\alpha T_C)^2}{2b}\left[\frac{\partial \ln T_C}{\partial u_i}\frac{\partial \ln b}{\partial u_j}+\frac{\partial \ln T_C}{\partial u_j}\frac{\partial \ln b}{\partial u_i}-\frac{2\alpha_3 T_C}{b}\frac{\partial \ln T_C}{\partial u_i}\frac{\partial \ln T_C}{\partial u_j}\right],$$

$$C_{ij}^{(6)}=a_5\left(\frac{\alpha T_C}{b}\right)^3\frac{\partial \ln T_C}{\partial u_i}\frac{\partial \ln T_C}{\partial u_j}. \tag{3}$$

The dependence of the tensor (2) on the magnetic field is described not only by the last term on the right-hand side of relation (2), which is proportional to $M^2(T,H)$, but also by the dimensionless parameter[16]

$$\xi(T,H)=\frac{H}{2bM^3(T,H)}. \tag{4}$$

To obtain formula (2) we used the equation of the magnetic state of a ferromagnet corresponding to Eq. (1):

$$\frac{H}{M}=\alpha(T-T_C)+\frac{1}{2}\alpha_2(T-T_C)^2+[b+\alpha_3(T-T_C)]M^2+a_5M^4. \tag{5}$$

The second term on the right-hand side of Eq. (2) corresponds to the result of Ref. 12 and in the limit of zero field $H=0$ describes the ''jump'' of the elastic moduli at the Curie point since, according to Eqs. (4) and (5), $\xi(T,0)=\infty$ for $T>T_C$ and $\xi(T,0)=0$ for $T<T_C$. The following terms on the right-hand side of Eq. (2) for $H=0$ describe the variation of the slope angle of the temperature dependence of the elastic moduli, which is proportional to $M^2(T,0)$, during the transition to the ferromagnetic state, and correspond to the result of Ref. 14. An expression similar to (2) was studied in Refs. 16 for a finite magnetic field in connection with the Stoner model of a weakly ferromagnetic metal.

Formulas (2) and (5) allow one to examine the temperature and field dependence of the elastic moduli of a ferromagnet near the Curie temperature. Let us first consider the range of temperatures below the Curie temperature ($T<T_C$). Then in the case of a weak magnetic field,

$$H^{2/3}\ll\frac{\alpha T_C}{b^{1/3}}\left(1-\frac{T}{T_C}\right), \tag{6}$$

Eqs. (2) and (5) yield the following dependence of the elastic tensor on the temperature and magnetic field:

$$C_{ij}^H(T,H)=C_{ij}^0(T)-C_{ij}^{(1)}+C_{ij}^{(7)}\left(1-\frac{T}{T_C}\right)+\left[C_{ij}^{(1)}+C_{ij}^{(8)}\times\left(1-\frac{T}{T_C}\right)\right]\left(1-\frac{T}{T_C}\right)^{-3/2}\frac{b^{1/2}H}{2(\alpha T_C)^{3/2}}, \tag{7}$$

where

$$C_{ij}^{(7)}=\sum_{n=2}^{6}C_{ij}^{(n)}=-\frac{(\alpha T_C)^2}{2b}\left[\frac{\partial \ln T_C}{\partial u_i}\frac{\partial \ln(\alpha^2/b)}{\partial u_j}+\frac{\partial \ln T_C}{\partial u_j}\frac{\partial \ln(\alpha^2/b)}{\partial u_i}+\frac{1}{T_C}\frac{\partial^2 T_C}{\partial u_i\partial u_j}+\left(\frac{3\alpha_3 T_C}{b}-\frac{3\alpha_2 T_C}{\alpha}-\frac{2\alpha a_5 T_C}{b^2}\right)\frac{\partial \ln T_C}{\partial u_i}\frac{\partial \ln T_C}{\partial u_j}\right],$$

$$C_{ij}^{(8)}=-C_{ij}^{(2)}-2C_{ij}^{(3)}+2C_{ij}^{(4)}+C_{ij}^{(5)}.$$

Inside the brackets in the last term on the right-hand side of expression (7) we have kept the term proportional to the small parameter $1-T/T_C\ll 1$ since, as will become clear below, the inequality $C_{33}^{(1)}\ll C_{33}^{(8)}$ holds in the case of gadolinium. This distinguishes expression (7) from the formula that can be obtained in this limit on the basis of Refs. 12 and 14.

Let us turn now to the case of a strong magnetic field, where the following inequalities are satisfied:

$$\frac{2\alpha T_C}{3b^{1/3}}\left|1-\frac{T}{T_C}\right|\ll H^{2/3}\ll\frac{b^{5/3}}{a_5}, \tag{8}$$

The right-hand inequality (8) here denotes the smallness of the term a_5M^4 in Eq. (5) in comparison with bM^2. Under these conditions the dependence of the elastic moduli (2) on the temperature and magnetic field is described by the following expression:

$$C_{ij}^H(T,H)=C_{ij}^0(T)-\frac{2}{3}C_{ij}^{(1)}-\frac{2}{9}C_{ij}^{(1)}\times\left(1-\frac{T}{T_C}\right)\frac{\alpha T_C}{b^{1/3}H^{2/3}}+C_{ij}^{(9)}\frac{b^{1/3}H^{2/3}}{\alpha T_C}, \tag{9}$$

where

$$C_{ij}^{(9)}=C_{ij}^{(4)}+\frac{2}{3}C_{ij}^{(5)}+\frac{4}{9}C_{ij}^{(6)}.$$

If the additional condition on the magnitude of the magnetic field

$$\left[\frac{2C_{ij}^{(1)}}{9C_{ij}^{(9)}}\right]^{1/2}\frac{\alpha T_C}{b^{1/3}}\left|1-\frac{T}{T_C}\right|^{1/2}\ll H^{2/3} \tag{10}$$

is fulfilled, relation (9) takes the form

$$C_{ij}^H(T,H)=C_{ij}^0(T)-\frac{2}{3}C_{ij}^{(1)}+C_{ij}^{(9)}\frac{b^{1/3}H^{2/3}}{\alpha T_C}. \tag{11}$$

We stress here that formula (11) is a departure from the approximation of Ref. 12, according to which the elastic modulus is independent of the magnetic field for $T\simeq T_C$. It follows from formula (11) that the magnetic field

$$H_C=\frac{(\alpha T_C)^{3/2}}{b^{1/2}}\left[\frac{2C_{ij}^{(1)}}{3C_{ij}^{(9)}}\right]^{3/2} \tag{12}$$

compensates the ''negative'' jump of the elastic modulus at $T=T_C$.

Finally, let us consider the temperature range corresponding to the paramagnetic state ($T>T_C$). Here in the case of a weak magnetic field

$$H^{2/3} \ll \frac{\alpha T_C}{b^{1/3}}\left(\frac{T}{T_C}-1\right), \tag{13}$$

the dependence of the elastic tensor on the temperature and magnetic field, according to Eqs. (2) and (4), takes the form (compare with Refs. 4 and 17)

$$C_{ij}^H(T,H)=C_{ij}^0(T)-\left[2C_{ij}^{(1)}\left(\frac{T}{T_C}-1\right)^{-1}+2C_{ij}^{(2)}-C_{ij}^{(4)}\right] \times\left(\frac{T}{T_C}-1\right)^{-2}\frac{bH^2}{(\alpha T_C)^3}. \tag{14}$$

We will use the above relations in the following section to analyze the experimental data on the field dependence of Young's modulus $E(T,H)$ in gadolinium near the Curie temperature.

4. DISCUSSION AND CONCLUSIONS

We turn now to an interpretation of the experimental trends depicted in Figs. 1 and 2 for gadolinium. It is important to bear in mind that the measured values of Young's modulus $E(T,H)$ along the hexagonal c axis are determined by the component of the elastic tensor $C_{33}^H(T,H)$. Here it can be seen that the temperature behavior of Young's modulus $E(T,0)$ in the limit of zero field $H=0$, described the segment of the zig-zag curve (curve *1*) in Fig. 1 near the Curie temperature ($T\leqslant T_C$), can be approximated by formula (6), according to which we have

$$E(T,0)=E_0(T)-C_{33}^{(1)}+C_{33}^{(7)}\left(1-\frac{T}{T_C}\right), \tag{15}$$

where $E_0(T)=C_{33}^0(T)$. Comparison of formula (15) with the data in Fig. 1 makes it possible to find the magnetoelastic coefficients $C_{33}^{(1)}\simeq 1.2$ GPa and $C_{33}^{(7)}\simeq 30.8$ GPa for gadolinium, which determine respectively the jump of the modulus $E(T,0)$ at $T=T_C$ and the variation of the slope of the temperature dependence $E(T,0)$ below T_C. The influence of the magnetic field on the temperature dependence of the modulus $E(T,H)$ in the temperature range $T>T_{SR}$ visible in Fig. 1 corresponds qualitatively to the regularities described by formulas (7), (9), and (11), where, with growth of the magnetic field, the jump in the elastic modulus $E(T,H)$ at the Curie point washes out and the slope of the temperature dependence of $E(T,H)$ grows with the field if $C_{33}^{(8)}>0$, $C_{33}^{(9)}>0$. In order to discuss these trends quantitatively, we will consider in more detail the results displayed in Fig. 2. Toward this end, we rewrite formulas (7) and (11) with the help of Eq. (4) in a form that contains an explicit dependence on $M^2(T,H)$. Then, in the case of a weak field (6), with the help of formula (7) we find the following formula for the field dependence of the ΔE effect:

$$\Delta E(T,H)=C_{33}^H(T,H)-C_{33}^H(T,0) =\left[C_{33}^{(1)}+C_{33}^{(8)}\left(1-\frac{T}{T_C}\right)\right] \times\frac{M^2(T,H)-M^2(T,0)}{2M^2(T,0)}, \tag{16}$$

and in the case of a strong field (8), given condition (10), formula (11) gives

$$\Delta E(T,H)=C_{33}^H(T,H)-C_{33}^H(T,0)=C_{33}^{(9)}\frac{bM^2(T_C,H)}{\alpha T_C}. \tag{17}$$

Thus, in both limiting cases (weak and strong field), if the additional condition (10) is met, the field dependence of the ΔE effect turns out to be proportional to $M^2(T,H)$ at different temperatures. This conclusion corresponds in Fig. 2 to traces *1–7*. In this case, traces *1* and *2* correspond, obviously, to a strong field since for them the variation of $M^2(T,H)$ with increasing field is large in comparison with $M^2(T,0)$ for a fixed temperature $T\simeq T_C$. The slope of these curves, according to formula (17), is equal to

$$\tan\varphi_n=C_{33}^{(9)}\frac{b}{\alpha T_C}, \quad n=1,2. \tag{18}$$

On the contrary, curves *3–7* in Fig. 2 correspond to the weak-field limit, where the variation of $M^2(T,H)$ with increasing field is small in comparison with the spontaneous value $M^2(T,0)$. The slope of these curves, according to formula (16), is equal to

$$\tan\varphi_n=\left[C_{33}^{(1)}\left(1-\frac{T}{T_C}\right)^{-1}+C_{33}^{(8)}\right]\frac{b}{2\alpha T_C}, \quad n=3-7. \tag{19}$$

Since it is clear that the slope of curves *3–7* does not depend on the temperature, this is possible according to formula (19) if

$$C_{33}^{(1)}\left(1-\frac{T}{T_C}\right)^{-1}\ll C_{33}^{(8)}. \tag{20}$$

Under these conditions, Eq. (19) yields

$$\tan\varphi_n=C_{33}^{(8)}\frac{b}{2\alpha T_C}, \quad n=3-7. \tag{21}$$

The experimental data begin to deviate from curves *3–7* in Fig. 2 when the increment of $M^2(T,H)$ in a field becomes noticeable in comparison with $M^2(T,0)$. In this case, the trend described by formula (16) gives way to the trend described by formula (17), corresponding to traces *1*, *2*, which fill the role of asymptotic limits for the experimental data. This trend can also be made out in Fig. 2.

The experimental results depicted in Fig. 2 allow us to obtain quantitative estimates of the parameters and statements of the phenomenological approach fleshed out above for the case of gadolinium. First of all, from Fig. 2 we can find the ratio $\alpha T_C/b\simeq(1.1-1.3)\times 10^7$ G^2 using data on $M^2(T,0)$ plotted along the abscissa for known temperatures, and the value of the mass density of gadolinium $\rho\simeq 7.87$ g/cm^3 from Ref. 11. This, in turn, allows one to estimate the

parameter $b \simeq (0.56-0.66)\times 10^{-4}$ G^{-2} using the experimental value of the Curie constant $C=\alpha^{-1}=0.4$ K from Fig. 1 of Ref. 18. Next, formulas (18) and (21) together with experimental data on the slopes $\tan\varphi_n \simeq 1.29\times 10^5$ of curves *1* and *2* in Fig. 2 and $\tan\varphi_n \simeq 0.47\times 10^6$ of curves *3–7* allow one to calculate the magnetoelastic coefficients $C_{33}^{(9)} \simeq (1.4-1.7)\times 10^2$ GPa and $C_{33}^{(8)} \simeq (1.0-1.2)\times 10^3$ GPa. Bearing the above results in mind, it is possible to rewrite inequality (20) for gadolinium in the form

$$(1.0-1.2)\times 10^{-3} \ll 1-T/T_C. \tag{22}$$

Obviously, inequality (22) is satisfied for temperatures corresponding to curves *3–7*. This justifies the use of formulas (21) to describe these curves.

The values obtained above and formula (12) allow one to estimate the value of the magnetic field $H_C \simeq 0.8$–1.1 kOe at which the negative ΔE effect disappears at the Curie point. The value of H_C so obtained is in good agreement with the experimental data shown in Fig. 1.

Let us now discuss conditions (8) and (10) defining a strong magnetic field, applied to our experiment. The values obtained above allow one to rewrite the left-hand inequality of inequalities (8) in the form

$$(1.3-1.4)\times 10^3 |1-T/T_C|^{3/2} \text{ kOe} \ll H. \tag{23}$$

Thus, for $T=300$ K the designation ''strong field'' means $H \gg 5$ kOe, and for $T=289$ K—$H \gg 2$ kOe. These conditions are realized in our experiment for traces *1* and *2* in Fig. 2. For gadolinium inequality (10) takes the form

$$(19-24)|1-T/T_C|^{3/4} \text{ kOe} \ll H. \tag{24}$$

Obviously, for $T=300$ K and 289 K inequality (23) is stronger than inequality (24). Therefore, relations (17) and (18) correspond to traces *1* and *2* in Fig. 2. In contrast, for traces *3–7* the opposite inequality (23) is satisfied, i.e., the fields used in our experiment are weak at the temperatures corresponding to these curves, and these curves are described by formulas (16) and (21).

Thus, the new experimental trends in the field dependence of the ΔE effect along the hexagonal *c* axis in single-crystal gadolinium discovered in this work can be understood in the temperature range above the spin reorientation temperature within the framework of an approach based on the Landau theory of second-order phase transitions. Since similar experimental trends were also detected below the spin reorientation temperature, we think that the phenomenological approach developed in this work can be generalized to this temperature range with comparatively little effort (see also Refs. 9, 10, and 19).

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-17318-a) and the Federal Program for State Support of Leading Scientific Schools of Russia (Project No. 96-15-96429).

*)E-mail: zverev@sci.lpi.ac.ru

[1] K. P. Belov, *Rare-Earth Magnets and Their Applications* [in Russian] (Nauka, Moscow, 1980).
[2] A. K. Zvezdin, V. M. Matveev, A. A. Mukhin, and A. I. Popov, *Rare-Earth Ions in Magnetically Ordered Crystals* [in Russian] (Nauka, Moscow, 1985).
[3] M. Long, A. R. Wazzan, and R. Stern, Phys. Rev. **178**, 775 (1969).
[4] T. J. Moran and B. Lüthi, J. Phys. Chem. Solids **31**, 1735 (1970).
[5] H. Klimer and M. Rosen, Phys. Rev. B **7**, 2054 (1973).
[6] S. B. Palmer, E. W. Lee, and M. N. Islam, Proc. R. Soc. London, Ser. A **338**, 341 (1974).
[7] D. C. Jiles and S. B. Palmer, J. Phys. F: Metal Phys. **10**, 2857 (1980).
[8] S. A. Nikitin, A. M. Tishin, S. K. Godovikov, V. Yu. Bodriakov, and I. A. Avenarius, J. Magn. Magn. Mater. **125**, 190 (1993).
[9] L. M. Levinson and S. Shtrikman, J. Phys. Chem. Solids **32**, 981 (1971).
[10] F. Freyne, Phys. Rev. B **5**, 1327 (1972).
[11] S. A. Nikitin, *Magnetic Properties of Rare-Earth Metals and Their Alloys* [in Russian] (Moscow State Univ. Press, Moscow, 1989).
[12] K. P. Belov, G. I. Kataev, and R. Z. Levitin, Zh. Eksp. Teor. Fiz. **37**, 938 (1959) [Sov. Phys. JETP **10**, 670 (1960)].
[13] G. I. Kataev and V. V. Shubin, Acta Phys. Pol. A **68**, 115 (1985).
[14] G. I. Kataev, Fiz. Met. Metalloved. **11**, 375 (1961).
[15] V. P. Silin, D. Wagner, and V. M. Zverev, Phys. Lett. A **192**, 421 (1994).
[16] V. M. Zverev and V. P. Silin, Zh. Éksp. Teor. Fiz. **89**, 642 (1985) [Sov. Phys. JETP **62**, 369 (1985)]; Fiz. Met. Metalloved. **61**, 1055 (1986).
[17] B. W. Southern and D. A. Goodings, Phys. Rev. B **7**, 534 (1973).
[18] S. Arajs and R. V. Colvin, J. Appl. Phys. Suppl. **32**, 336 (1961).
[19] K. P. Belov, A. K. Zvezdin, A. M. Kadomtseva, and R. Z. Levitin, *Orientational Transitions in Rare-Earth Magnets* [in Russian] (Nauka, Moscow, 1979).

Translated by Paul F. Schippnick

Particle transport in a disordered medium: numerical experiment

V. A. Schweigert*) and A. L. Aleksandrov

Institute of Theoretical and Applied Mechanics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 18 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 2122–2129 (December 1998)

Using the multicenter Schrödinger equation for calculating the transmittance of a flat layer of randomly distributed point scattering centers through which a particle passes, we show that when the scattering length for one center is comparable to the particle wavelength λ or is larger, the Ioffe–Regel hypothesis holds. (According to this hypothesis, as the scatterer number density increases, the transmittance of the layer becomes constant, while the value of the particle's effective mean free path remains of order λ.) When the scattering length is small compared to λ, the Ioffe–Regel hypothesis does not hold. As the scattering length decreases, the accuracy of the approximation of the effective scattering potential gradually increases, and, depending on the strength of the potential, particles may either tunnel or diffuse; the effective mean free path can be much smaller than λ. © *1998 American Institute of Physics.* [S1063-7761(98)01612-6]

1. INTRODUCTION

Apart from being fundamental, the problem of the quantum transport of particles in a disordered medium has lately attracted a lot of interest in applied science, which is due mainly to the transition to nanometer technology in the fabrication of computer chips and other devices. If the characteristic dimensions of a device are smaller than the electron wavelength, the commonly used physical and mathematical models describing the operation of such devices break down. New phenomena, such as weak localization and the universal manner of conductance fluctuations, have also been detected recently in studies of mesoscopic electron transport, where the inelastic scattering length is much larger than all other characteristic lengths. All this has drawn a lot of attention to the development of new approaches in the description of mesoscopic charge-carrier transport, approaches based on the numerical solution of the Schrödinger equation for a large number of scattering centers.

The ideas about the origin of particles in highly disordered media are contradictory. According to the Ioffe–Regel hypothesis,[1] the mean free path l cannot be smaller than the wavelength λ. This led Mott to the idea (see Ref. 2) of minimum metallic conductivity, according to which the conductivity and diffusion coefficient reach their minimum values as the concentration of the scattering centers increases.

A more modern approach is based on the scaling theory developed for the case where the size d of a region of the disordered medium is much larger than λ. According to this theory, as the disorder in the medium becomes more pronounced, Anderson localization sets in, and near a critical value of the scattering center concentration the conductivity and diffusion coefficient decreases according to a power law.[3] The question of the validity of a one-parameter scaling theory is currently widely debated,[4] and the detection of anomalously large statistical fluctuations of conductance in disordered media points to the limits of this theory.[5]

Rigorous theoretical approaches are limited to the weak localization region ($l \ll \lambda$). Hence, for the case where all the characteristic lengths are close in order of magnitude ($d \sim l \sim \lambda$), numerical experiments prove to be the most effective method of investigation. In particular, a numerical experiment that would verify the Ioffe–Regel hypothesis would be very interesting.

The existing approaches to modeling the transport process[6] focus on calculations of the critical exponent and use the tight-binding Hamiltonian with diagonal disorder. Despite the efficiency (due to the sparsity of the matrix) of the tight-binding approximation, we believe that such an approach is biased. Physically it corresponds to the case of transport of particles in a bound state, where the Green's function decreases exponentially over the average distance between particles. It is difficult to relate the results obtained through the use of the tight-binding Hamiltonian to the classical concepts of transport theory (e.g., the mean free path and the diffusion coefficient), so that it is advisable to have an alternative approach.

The current techniques used in describing the operation of nanoelectronic devices (quantum wells, superlattices, etc.) ordinarily employ the effective-mass approximation. This leaves many problems, e.g., the problem of matching at the boundary between two media, unresolved. However, we can discard the effective-mass approximation and think of the medium as a set of zero-range potentials for which the solution of the Schrödinger equation is possible by the method described in this paper.

2. THE ALGORITHM

When studying the mesoscopic electron transport, we can ignore inelastic scattering. It then becomes possible to solve the Schrödinger equation for the scattering problem, in

which a plane incident wave is specified and an outgoing spherical wave exists at infinity. Mesoscopic transport occurs in semiconductors and insulators at low temperatures, when the electron wavelength is much larger than the characteristic range of the potential, which makes it possible to use the approximation of zero-range potentials.

Suppose that a plane wave with a wave vector $\mathbf{k}$ is incident on a layer consisting of N_c point scatterers described by zero-range potentials (we set $\hbar=m=1$). Then the solution of the time-independent Schrödinger equation is a linear combination of the incident wave and the outgoing spherical waves from each scatterer:

$$\psi(\mathbf{r})=\exp(i\mathbf{k}\cdot\mathbf{r})+\sum_j^{N_c} A_j G(R_j), \quad R_j=|\mathbf{r}-\mathbf{r}_j|, \tag{1}$$

where the coefficients A_j can be found from the effective boundary conditions at the surface of each scatterer:[7]

$$\left.\frac{\partial \ln(R_j\psi)}{\partial R_j}\right|_{R_j\to 0} = -\alpha_j, \tag{2}$$

where $1/\alpha_j$ is the scattering length at the jth center with coordinates $\mathbf{r}_j$, and $G(R)=\exp(ikR)/R$ is the Green's function of the Schrödinger equation in empty space. In our calculations we were forced to limit ourselves to consideration of the particles incident along the z axis perpendicular to a sample with finite transverse dimensions $L_x=L_y\simeq(2-4)\lambda$. The sample contained N_c randomly distributed point potentials. We were forced also to use periodic spatial conditions for the scattering medium. Then, the Green's function is the sum of vacuum Green's functions shifted by periods $L_{x,y}$:

$$G(\mathbf{r})=\sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\frac{\exp(ik|\mathbf{r}-\mathbf{r}_{mn}|)}{|\mathbf{r}-\mathbf{r}_{mn}|}, \tag{3}$$

$$\mathbf{r}_{mn}=mL_x\mathbf{e}_x+nL_y\mathbf{e}_y.$$

The series (3) converges very slowly, and in our calculations we used its Fourier transform with respect to the transverse coordinates. Near a scattering center the asymptotic behavior of the Green's function is given by the formula

$$G(R_j)|_{R_j\to 0}=\frac{\exp(ikR_j)}{R_j}+C, \tag{4}$$

where the constant C is the sum of the series (3) without the terms with $m,n=0$ at $\mathbf{r}=0$; it is determined by the geometrical dimensions $L_{x,y}$ and the wavelength. We found it convenient to calculate this constant in the form of the difference the Fourier transform of (3) and the Fourier transform of $\exp(ikr)/r$, which at $\mathbf{r}=0$ yields a converging difference of an infinite integral and a divergent sum:

$$C=\frac{i}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{dp_x\,dp_y}{(k^2-\mathbf{p}^2)^{1/2}} - \frac{2\pi i}{L_xL_y}\sum_{P_x}\sum_{P_y}\frac{1}{(k^2-\mathbf{p}^2)^{1/2}},$$

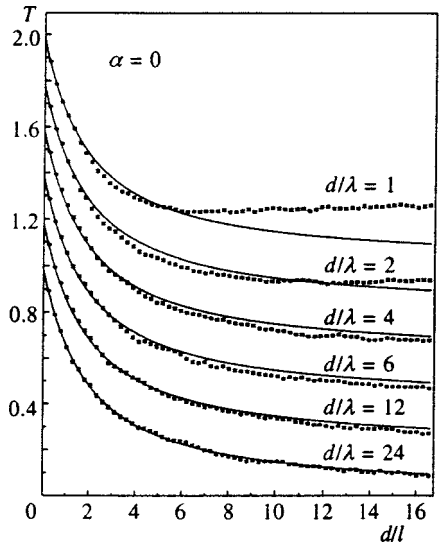

FIG. 1. Dependence of the classical (solid curves) and quantum (dots) layer transmittances on the concentration of scatterers with $\alpha=0$ at a fixed layer thickness for different values of the wavelength λ (to make the diagram more graphic, the curves are moved apart along the axis of ordinates by 0.2). Here and in the figures that follow the vertical notches indicate the statistical error.

$$P_{x,y}=\frac{2\pi n_{x,y}}{L_{x,y}}, \quad P_z, p_z=0.$$

Substituting (1) and (4) in (2), we arrive at system of complex-valued linear equations for the N_c coefficients A_j, which can be solved numerically by Gaussian elemination:

$$(ik+\alpha_j+C)A_j+\sum_{l\neq j}^{N_c} G(\mathbf{r}_j-\mathbf{r}_l)A_l=-\exp(ikz_j).$$

The transmittance T is given by the formula

$$T=\frac{1}{kL_xL_y}\int_0^{L_x}\int_0^{L_y} J_z\,dx\,dy, \quad \mathbf{J}=\frac{i}{2}(\psi\nabla\psi^*-\psi^*\nabla\psi)$$

or, in final form,

$$T=1+\frac{2\pi}{kL_xL_y}\sum_j^{N_c}(A_jF_j-A_j^*F_j^*),$$

where

$$F_j=i\exp(-ikz_j)+\frac{2\pi}{L_xL_y} \times\sum_{P_x}^{|\mathbf{P}|<k}\sum_{P_y}\sum_l^{N_c} A_l^*\frac{\exp(i\mathbf{P}\cdot(\mathbf{r}_j-\mathbf{r}_l)-q(z_j-z_l))}{q},$$

$$q=|k^2-\mathbf{P}^2|^{1/2}.$$

The zero-range potential approximation provides a method for solving the Schrödinger equation numerically for a fairly large number of scatterers (we chose N_c as high as 1000), which makes it possible to model particle transport in a disordered medium for a broad range of parameter values.

3. RESULTS OF CALCULATIONS

The results depicted in Figs. 1–3 were obtained by av-

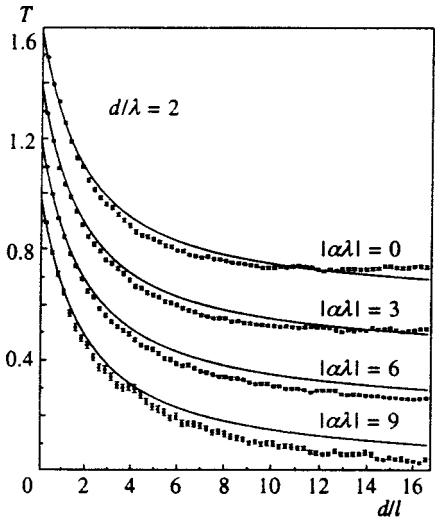

FIG. 2. Dependence of the classical (solid curves) and quantum (dots) layer transmittance on the concentration of unlike scatterers for different scattering lengths at fixed layer thickness and wavelength (to make the diagram more graphic, the curves are moved apart along the axis of ordinates by 0.2).

eraging the transmittance over various configurations of the scattering centers and simultaneously estimating the statistical error. In different trials involving the same (constant) number of scattering centers, the coordinates of these centers were produced by a random number generator. The common number of trials was several hundred.

Three cases were considered for a fixed absolute value of the scattering length $1/\alpha$: attractive scattering centers ($\alpha<0$), repulsive scattering centers ($\alpha>0$), and unlike scattering centers (the sign of each α_j is chosen at random with equal fractions of both values). The single-center scattering cross section has the form[7]

$$\sigma_{\alpha,k}=\frac{4\pi}{\alpha^2+k^2}. \tag{5}$$

The scatterer concentration n is related to the mean free path l through the formula

$$n=\frac{\alpha^2+k^2}{4\pi l},$$

and it is convenient to select the following parameters (with the dimensions of length) as the main parameters of the problem: λ, $1/\alpha$, l, and the thickness d of the scatterer layer. The transmittance T of the layer is a function of only three independent dimensionless parameters, e.g., λ/l, d/l, and $\alpha\lambda$ (naturally, other combinations are also possible).

We examined the dependence of the transmittance T on d/l for different fixed values of the two other independent parameters. Physically this corresponds to two situations: in the first case the scatterer concentration varies but λ, α, and d are fixed, i.e., the parameter l varies (with the dimensionless parameters d/λ and $\alpha\lambda$ fixed); in the second case the thickness d of the scatterer layer varies but λ, α, and l are fixed, i.e., the parameters λ/l and $\alpha\lambda$ are fixed.

In the classical limit $\lambda/l\to 0$, particle transport is described by a transport equation with an isotropic scattering indicatrix and a cross section (5). For $d\ll l$, the transmittance of the layer linearly decreases with increasing thickness, and for $d\gg l$, the diffusive transport sets in and T is proportional to l/d. To be able to compare our results with the quantum mechanical results, we calculated the classical transmittance $T_{\text{class}}(d/l)$ numerically, solving the transport equation by the Monte Carlo method.

Figures 1 and 2 depict the results of calculating the T vs. l dependence of a fixed layer thickness l, which is convenient for numerical verification of the Ioffe–Regel hypothesis. When $\lambda/l\ll 1$, the classical and quantum transmittances coincide with the statistical calculation error (the lower curve in Fig. 1). As λ/l increases, the quantum transmittance first becomes smaller than the classical, i.e., the interference of the scattered waves reduces the effective mean free path, which agrees with the Yakubov's theoretical results[8] obtained by the diagrammatic technique. When $\lambda>(6\text{-}7)l$ holds, the quantum transmittance of the layer ceases to decrease with increasing scatter concentration (the upper curves in Fig. 1), which corroborates the Ioffe–Regel hypothesis.[1] For small values of $|\alpha\lambda|$, the transmittance

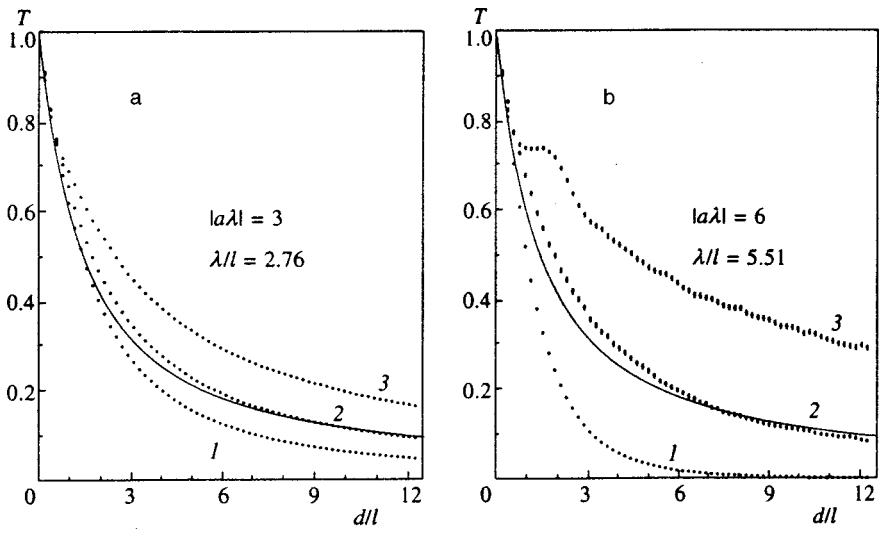

FIG. 3. Transmittance vs. layer thickness for two values of wavelength (a and b) at fixed scattering lengths and mean free paths for repulsive (curves *1*), unlike (curves *2*), and attractive (curves *3*) centers. The solid curves represent the dependence on the layer thickness of the classical transmittance for an effective mean free path $l_{\text{eff}}=0.77l$.

reaches a plateau as the scatterer concentration increases. This phenomenon exists for all types of center, and in the case of repulsive centers a minimum in transmittance is observed within a broad range of values of $\alpha\lambda$ for a scatterer concentration corresponding to $\lambda/l \approx 7$.

The Ioffe–Regel hypothesis probably becomes invalid as the scattering length decreases (Fig. 2). At least for $|\alpha\lambda| \gg 1$ and unlike scatterers, the quantum transmittance monotonically decreases with increasing d/l and is always smaller than the calculated transmittance in the classical approximation (the lower curves in Fig. 2). Calculations with $|\alpha\lambda| \gg 1$ show that for all types of scatterer the effective classical mean free path, which characterizes the transmittance of the layer, can be much smaller than λ.

The curves representing the dependence of the transmittance on the layer thickness for a fixed scatterer concentration are depicted in Fig. 3 for repulsive (curves *1*), unlike (curves *2*), and attractive (curves *3*) centers. As the value of the parameter $|\alpha\lambda|$ increases, the differences in the behavior of the curves *1–3* become more pronounced. A good explanation for such behavior is provided by the effective-potential approximation,[9] which can be used in the $k \ll n^{1/3} \ll |\alpha|$ range. According to this approximation, the effect of scatterers is equivalent to the presence of an average potential $U = 4\pi n/a$ (or, which is the same, to a shift of the mobility edge inside the layer by U). In this case the transport mode is determined by the sign of α.

For repulsive centers, as $\alpha\lambda$ and λ/l increase and the condition $k^2 < U$ holds, the tunneling mode sets in, with the transmittance of the layer exponentially decreasing as the layer thickness increases (curves *1* in Fig. 3).

For unlike centers we have $U = 0$, and the mobility edge in the layer is not shifted. The transport mode is close to diffusive and the transmittance of the layer is (approximately) inversely proportional to the layer thickness (curves *2* in Fig. 3). The effective mean free path l_{eff}, at which the classical transmittance corresponds to curve *2*, is somewhat smaller than l corresponding to the cross section (5) and is much smaller than λ.

For attractive centers, as $|\alpha\lambda|$ increases, the curves representing the dependence of transmittance on layer thickness first acquire an inflection point (curves *3* in Fig. 3), which is followed by the formation of small local maxima whose height increases with $|\alpha\lambda|$. The appearance of these maxima can also be interpreted by the effective-potential approximation. The attractive scatterers form a potential well with a transmittance

$$\left[\left(\frac{k}{2k_*} - \frac{k_*}{2k}\right)^2 \sin^2(k_* d) + 1\right]^{-1},$$

which reaches its maximum value at resonant layer thicknesses

$$d_m = \frac{m\pi}{k_*}, \quad k_* = \sqrt{k^2 - U}.$$

This expression for d_m provides a good description for the positions of the minima in the transmittance curves in the $k \ll n^{1/3} \ll |\alpha|$ range.

4. CONCLUSION

We have proposed a method for calculating the transmittance of a particle with positive energy through a layer of a disordered medium. The method is based on solving the three-dimensional Schrödinger equation for multicenter scattering. The medium is interpreted as a set of randomly distributed point scattering centers with a scattering length $1/\alpha$. With this approximation we have obtained an exact solution of the Schrödinger equation for an arbitrary (but finite) number of centers.

As expected, in the classical limit $\lambda \ll l$ ($l = 1/n\sigma_{\alpha,k}$), which corresponds to low scatterer concentrations, the solutions of the classical and quantum scattering problems are essentially the same.

In our calculations we also observed a transition to another limit, $\lambda \gg l \gg 1/\alpha$, where the effective-potential approximation is valid. Here the transport mode is determined by the sign of α, which corresponds to a shift in the particle mobility edge by the value of the average scatterer potential $U = 4\pi n/\alpha$. For instance, for repulsive scatterers the transmittance of the layer is of tunnel origin. When the scatterers are unlike, the behavior of the transmittance is close to that in the classical diffusive-transport mode, and the characteristic mean free path can be much smaller than λ. Finally, for attractive centers the transmittance decreases nonmonotonically with increasing thickness, with the positions of the local maxima in transmittance coinciding with the resonant thicknesses of a potential well of depth U.

In calculations in the intermediate range ($\lambda \sim l$, $0 \leqslant |\alpha\lambda| \lesssim 1$), where the effective-potential approximation breaks down, we have verified the validity of the Ioffe–Regel hypothesis, according to which the particle's mean free path must remain on the order of the wavelength for arbitrarily high scatterer concentrations.

Our calculations have demonstrated that for centers whose absolute value of the scattering length is of order, or larger than the particle wavelength, the Ioffe–Regel hypothesis can be used for positive and negative scattering lengths. As the concentration of scatterers (attractive, repulsive, or unlike) increases, the transmittance of the layer indeed decreases to a certain constant value attained at scattering lengths $l \leqslant (0.15-0.2)\lambda$.

However, as the scattering length decreases, a plateau in the transmittance is not observed. As the value of the parameter $|\alpha\lambda|$ increases, the layer transmittance gradually approaches the value described by the effective-potential approximation.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 95-02-04704-A).

*)E-mail: schweig@site.itam.nsc.ru

[1] A. F. Ioffe and A. R. Regel, Prog. Semicond. **4**, 237 (1960).
[2] N. F. Mott and E. A. Davis, *Electronic Processes in Non-crystalline Materials*, Clarendon Press, Oxford (1979).
[3] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[4] K. B. Efetov, Zh. Éksp. Teor. Fiz. **94**(1), 357 (1988) [Sov. Phys. JETP **67**, 199 (1988)].

[5] B. L. Al'tshuler, V. E. Kravtsov, and I. V. Lerner, Zh. Éksp. Teor. Fiz. **91**, 2276 (1986) [Sov. Phys. JETP **64**, 1352 (1986)].
[6] B. Kramer, K. Broderix, A. Mackinon, and M. Schreiber, Physica A **167**, 163 (1990).
[7] Yu. N. Demkov, *Method of Zero-Range Potentials* [in Russian], Leningrad Univ. Press, Leningrad (1975).
[8] I. T. Yakubov, Zh. Éksp. Teor. Fiz. **58**, 2075 (1970) [Sov. Phys. JETP **31**, 1118 (1970)].
[9] B. V. Petukhov, V. L. Pokrovskiĭ, and A. V. Chaplik, Zh. Éksp. Teor. Fiz. **53**, 1150 (1967) [Sov. Phys. JETP **26**, 678 (1968)].

Translated by Eugene Yankovsky

The singlet state in the Hubbard model with $U=\infty$

E. V. Kuz'min*)

Krasnoyarsk State University, 660062 Krasnoyarsk, Russia

I. O. Baklanov

L. V. Kirenskiĭ Institute of Physics, Siberian Branch of the Russian Academy of Sciences, 660036 Krasnoyarsk, Russia

(Submitted 23 January 1997; resubmitted 9 June 1998)
Zh. Éksp. Teor. Fiz. **114**, 2130–2144 (December 1998)

We discuss, in connection with the problem of the ground state in the Hubbard model with $U=\infty$, the normal (nonmagnetic) N-state of a system over the entire range of electron concentrations $n\leqslant 1$. It is found that in a one-particle approximation, e.g., in the generalized Hartree–Fock approximation, the energy $\epsilon_0(n)$ of the N-state is lower than the energy $\epsilon_{\rm FM}(n)$ of a saturated ferromagnetic state for all values of n. Using the random phase approximation we calculate the dynamical magnetic susceptibility and show that the N-state is stable for all values of n. A formally exact representation is derived for the mass operator of the one-particle electron Green's function, and its expression in the self-consistent Born approximation is obtained. We discuss the first Born approximation and show that when correlations are taken into account, the attenuation vanishes on the Fermi surface and the electron distribution function at $T=0$ acquires a Migdal discontinuity, whose magnitude depends on n. The energy of the N-state in this approximation is still lower than $\epsilon_{\rm FM}(n)$ for $n<1$. We show that the spin correlation functions are isotropic, which is a characteristic feature of the singlet states of the system. We calculate the spin correlation function for the nearest neighbors in the zeroth approximation as a function of n. Finally, we conclude that the singlet state of the system in the thermodynamic limit is the ground state. © *1998 American Institute of Physics.* [S1063-7761(98)01712-0]

1. INTRODUCTION. STATEMENT OF THE PROBLEM

In this paper we discuss the problem of the ground state and the electron distribution function in the Hubbard model[1] when the single-site repulsive potential is infinite, $U=\infty$.

The system Hamiltonian

$$H_\infty=\sum_{f\sigma\Delta} t(\Delta)X_f^{\sigma 0}X_{f+\Delta}^{0\sigma}-\lambda\sum_{f\sigma} X_f^{\sigma\sigma}, \tag{1}$$

specified on a d-dimensional lattice with N sites with a coordination number z and periodic boundary conditions, describes N_e electrons that tunnel to the nearest vacant sites (f stands for the lattice sites, Δ is the vector connecting the nearest neighbors, $t(\Delta)$ is the tunneling integral, and λ is the chemical potential). Since repulsion is assumed infinite, each site is either vacant or contains a single electron with spin projection σ (pairs are forbidden). This fact is reflected in the use of Hubbard operators with well-known commutation relations, which differ from those of fermions (see below).

At this point it is convenient to normalize the Hamiltonian to the halfwidth of the ''bare'' electron band zt:

$$h_\infty=H_\infty/zt, \quad t(\Delta)/zt=-1/z.$$

In the thermodynamic limit ($N\to\infty$, $N_e\to\infty$, $N_e/N=n=\text{const}$), all the properties of our system depend solely on the topology of the lattice (the number d of the lattice dimensions and the number z of the nearest neighbors) and the electron concentration $n(0\leqslant n\leqslant 1)$.

Applying Fourier transformations to all the operators,

$$X_{k\sigma}=\frac{1}{\sqrt{N}}\sum_f e^{ikf}X_f^{0\sigma},$$

$$X^{\sigma\sigma'}(q)=\frac{1}{\sqrt{N}}\sum_f e^{iqf}X_f^{\sigma\sigma'}, \tag{2}$$

where the vectors k and q belong to the first Brillouin zone, we find that the Hamiltonian becomes

$$h_\infty=\sum_{k\sigma}(\omega_k-\mu)X_{k\sigma}^\dagger X_{k\sigma}, \tag{3}$$

where ω_k represents the dimensionless dispersion law in the nearest-neighbors approximation,

$$\omega_k=-\frac{1}{z}\sum_\Delta e^{ik\Delta}=-\gamma_k, \tag{4}$$

and $\mu=\lambda/zt$ is the dimensionless chemical potential.

The simplicity of the Hamiltonian (3) is an illusion, since the operators X obey the following commutation relations and completeness condition:

$$\{X_{k\sigma},X^{\dagger}_{p\sigma'}\}=\frac{1}{N}\sum_f e^{i(k-p)f}(\delta_{\sigma\sigma'}X^{00}_f+X^{\sigma'\sigma}_f),$$

$$X^{00}_f+X^{\sigma\sigma}_f+X^{\overline{\sigma}\overline{\sigma}}_f=1. \tag{5}$$

Using the commutation relations (5), we can write the equations of motion ($\hbar=1$) as follows:

$$i\dot{X}_{k\sigma}=[X_{k\sigma},h_\infty]=\xi^0_kX_{k\sigma}+R_{k\sigma},\quad \xi^0_k=\omega_k-\mu, \tag{6}$$

$$R_{k\sigma}=\frac{1}{\sqrt{N}}\sum_q \omega_{k-q}[X^{\overline{\sigma}\sigma}(q)X_{k-q\,\overline{\sigma}}-X^{\overline{\sigma}\overline{\sigma}}(q)X_{k-q\sigma}], \tag{7}$$

$$\overline{\sigma}=-\sigma.$$

The nonlinearity of Eq. (6) stems from the algebra of the operators X or the presence of ''kinematic'' correlations between electrons with opposite projections of spin.

What is important in the problem of strong electron correlations is the ground state of the system and the electron distribution function.

The energy of the system depends on the system's total spin S. In the case of a saturated ferromagnetic state ($S=N_e/2$), the solution of the problem is exact and trivial for any admissible number of electrons. In this case the system is an ideal Fermi gas of electrons with the projection of the spins in one direction (the state in the k-space is either vacant or occupied by a single electron). At $T=0$ the distribution function is a Fermi step function.

An alternative of the ground state of the system is the singlet state ($S=0$ for an even number of electrons), which we also call a normal (N) strongly correlated state (the numbers of electrons with spin ''up'' and ''down'' are equal and there is no long-range magnetic order). The energy advantage of the N-state is due to the possibility of double occupancy of states in the k-space ($X^{\dagger}_{k\alpha}X^{\dagger}_{k\beta}|\psi_0\rangle\neq 0$), which lowers the chemical potential in comparison with that for the ferromagnetic state. Here, of course, pairs are forbidden in the direct space and $N_e\leqslant N$.

Correlations between electrons with opposite spin projections are strongest in the N-state. By ''dispersing'' the electrons according to their momenta such correlations, on the one hand, enhance the energy of the system and, on the other, may modify the Fermi step function at $T=0$. Here it is important to establish whether the Migdal discontinuity[2] in the distribution function is retained on the Fermi surface or whether it disappears, as it does in a marginal[3] or Luttinger[4] electron liquid.

The goal of this work is to calculate the distribution function of electrons in a system in the N-state at $T=0$ by the method of two-time retarded Green's functions.[5]. In contrast to the previous previous work of one of the authors (E.V.K),[6,7] here we examine the approximation of the mass operator corresponding to the self-consistent Born approximation.

2. GREEN'S FUNCTIONS: GENERAL PROPERTIES

We consider the two-time retarded anticommutator Green's function

$$G_{k\sigma}(\tau)=-i\theta(\tau)\langle\{X_{k\sigma}(t),X^{\dagger}_{k\sigma}(t')\}\rangle,\quad \tau=t-t', \tag{8}$$

and its Fourier transform

$$\langle\langle X_{k\sigma}|X^{\dagger}_{k\sigma}\rangle\rangle_E\equiv G_{k\sigma}(E)=\int_{-\infty}^{\infty}d\tau\, e^{iE\tau}G_{k\sigma}(\tau), \tag{9}$$

where E is the spectral variable. It is convenient to write the Green's function as follows:

$$G_{k\sigma}(E)=c_\sigma F_{k\sigma}(E), \tag{10}$$

$$\langle\{X_{k\sigma},X^{\dagger}_{k\sigma}\}\rangle=1-n_{\overline{\sigma}}\equiv c_\sigma,\quad n_{\overline{\sigma}}=\frac{1}{N}\sum_f\langle X^{\overline{\sigma}\overline{\sigma}}_f\rangle.$$

We introduce the spectral intensity

$$I_{k\sigma}(E)=-\frac{1}{\pi}\,\mathrm{Im}\,F_{k\sigma}(E+i0). \tag{11}$$

Using this quantity and the spectral theorem, we can find the averages (here and in what follows we assume $T=0$):

$$\langle X^{\dagger}_{k\sigma}X_{k\sigma}\rangle=n_{k\sigma}=c_\sigma\int_{-\infty}^{0}I_{k\sigma}(E)\,dE\equiv c_\sigma f_{k\sigma}. \tag{12}$$

The spectral intensity obeys the sum rule

$$\int_{-\infty}^{\infty}I_{k\sigma}(E)\,dE=1. \tag{13}$$

The chemical potential can be found by solving the equation

$$n=\frac{1}{N}\sum_{k\sigma}n_{k\sigma}=\frac{1}{N}\sum_{k\sigma}c_\sigma\int_{-\infty}^{0}I_{k\sigma}(E)\,dE, \tag{14}$$

where n is the given electron concentration. The system energy (per lattice site) is

$$\epsilon=\frac{1}{N}\sum_{k\sigma}\omega_kn_{k\sigma}=\frac{1}{N}\sum_{k\sigma}\omega_kc_\sigma\int_{-\infty}^{0}I_{k\sigma}(E)\,dE. \tag{15}$$

Since all calculations are done in the thermodynamic limit, we can replace the sums by integrals:

$$\frac{1}{N}\sum_k A(\omega_k)=\int_{-1}^{1}A(\omega)\rho(\omega)\,d\omega,\quad \int_{-1}^{1}\rho(\omega)\,d\omega=1,$$

where $\rho(\omega)$ is the density of states corresponding to the dispersion law ω_k (for alternant lattices $\rho(\omega)=\rho(-\omega)$, $-1\leqslant\omega_k\leqslant 1$). After the chemical potential has been found, we can use formula (12) to find the one-particle distribution function.

A saturated ferromagnetic state in the Hubbard model can be described exactly. At $T=0$ the chemical potential and the system energy can be found by the formulas

$$n=\int_{-1}^{\mu}\rho(\omega)\,d\omega,\quad \epsilon_{\mathrm{FM}}(n)=\int_{-1}^{\mu}\omega\rho(\omega)\,d\omega, \tag{16}$$

which makes it possible to determine the explicit dependence of the energy of the ferromagnetic state on the electron concentration n.

3. THE MASS OPERATOR OF THE GREEN'S FUNCTION AND ITS APPROXIMATION IN THE SELF-CONSISTENT BORN APPROXIMATION

In the exact equation (6) we can isolate the linear part, which corresponds to the generalized Hartree–Fock approximation. This is achieved by introducing the irreducible operator[5–7]

$$\tilde{R}_{k\sigma}=R_{k\sigma}-\frac{\langle\{R_{k\sigma},X^{\dagger}_{k\sigma}\}\rangle}{\langle\{X_{k\sigma},X^{\dagger}_{k\sigma}\}\rangle}X_{k\sigma}\equiv R_{k\sigma}-\frac{a_{k\sigma}}{c_{\sigma}}X_{k\sigma}, \tag{17}$$

for which $\langle\{\tilde{R}_{k\sigma},X^{\dagger}_{k\sigma}\}\rangle=0$. Actually this means that the procedure allows for all ''internal'' pairings, which lead to linearization. Equation (6) takes the form

$$i\dot{X}_{k\sigma}=\xi_{k\sigma}X_{k\sigma}+\tilde{R}_{k\sigma}, \tag{18}$$

where

$$\xi_{k\sigma}=\xi^{0}_{k\sigma}+\frac{1}{c_{\sigma}}\langle\{R_{k\sigma},X^{\dagger}_{k\sigma}\}\rangle=\Delta_{\sigma}+\left(c_{\sigma}+\frac{\kappa_{\sigma}}{c_{\sigma}}\right)\omega_{k}-\mu, \tag{19}$$

$\Delta_{\sigma}=|\epsilon^{-}_{\sigma}|/c_{\sigma}$, $\kappa_{\sigma}=\langle X^{\overline{\sigma\sigma}}_{f}X^{\overline{\sigma\sigma}}_{f+\Delta}+\nu_{f\sigma}\bar{\nu}_{f+\Delta\sigma}\rangle$, and ϵ^{-}_{σ} is the energy (per lattice site) of the subsystem of electrons with spin projections $\overline{\sigma}$, with $\nu_{f\sigma}=X^{\sigma\sigma}_{f}-n^{-}_{\sigma}$.

The Green's function $G_{k\sigma}(E)$ (or $F_{k\sigma}(E)$) satisfies the Dyson equation. To set up the equation, we use the well-known method of first differentiating the Green's functions with respect to the ''first'' time t and then with respect to the ''second'' time t'. Using the equation of motion in the generalized Hartree–Fock approximation, we arrive[6,7] at

$$F_{k\sigma}(E)=\frac{1}{E-\xi_{k\sigma}-\Sigma_{k\sigma}(E)+i0}, \tag{20}$$

$$\Sigma_{k\sigma}(E)=\frac{1}{c_{\sigma}}\langle\langle\tilde{R}_{k\sigma}|\tilde{R}^{\dagger}_{k\sigma}\rangle\rangle^{(c)}_{E}.$$

The mass operator $\Sigma_{k\sigma}(E)$ (the self-energy part of the Green's function) is the connected (index (c)) part of the higher-order Green's function, which is not cut along the line of the graphical representation of the zeroth-order Green's function $F^{0}_{k\sigma}(E)=(E-\xi_{k\sigma})^{-1}$. Assuming

$$\Sigma_{k\sigma}(E+i0)\equiv M_{k\sigma}(E)-i\Gamma_{k\sigma}(E), \tag{21}$$

we have

$$I_{k\sigma}(E)=-\frac{1}{\pi}\mathrm{Im}\,F_{k\sigma}(E)$$

$$=\frac{1}{\pi}\frac{\Gamma_{k\sigma}(E)}{[E-\xi_{k\sigma}-M_{k\sigma}(E)]^{2}+\Gamma^{2}_{k\sigma}(E)}, \tag{22}$$

with $\Gamma_{k\sigma}(E)\geqslant 0$ (Ref. 5). Formally the representation (20) of the mass operator is exact. However, to perform calculations, we need the explicit form of the mass operator, which means we must use an approximation scheme of some sort.

The spectral representation of the higher-order retarded anticommutator Green's function is

$$\langle\langle\tilde{R}_{k\sigma}|\tilde{R}^{\dagger}_{k\sigma}\rangle\rangle_{E}=\int^{\infty}_{-\infty}\frac{J_{k\sigma}(E')}{E-E'+i0}dE'. \tag{23}$$

If we now employ the spectral theorem, we can express the spectral density of this function in terms of time averages:

$$J_{k\sigma}(E)=\int^{\infty}_{-\infty}\frac{d\tau}{2\pi}e^{iE\tau}(\langle\tilde{R}^{\dagger}_{k\sigma}(0)\tilde{R}_{k\sigma}(\tau)\rangle+\langle\tilde{R}_{k\sigma}(\tau)\tilde{R}^{\dagger}_{k\sigma}(0)\rangle) \tag{24}$$

with $\Gamma_{k\sigma}(E)=\pi J_{k\sigma}(E)/c_{\sigma}$.

Using the definition (7) of the operator $R_{k\sigma}$, we calculate the averages in (24), writing them as a product of quasi-Bose and quasi-Fermi averages (i.e., carrying out ''external'' couplings). Each of these averages can be found by the spectral theorem[5] in terms of the corresponding commutator and anticommutator Green's functions. As a result we obtain

$$J_{k\sigma}(E)\approx\frac{1}{N}\sum_{q}\omega^{2}_{k-q}\int^{\infty}_{-\infty}d\omega\,N(\omega)f(E-\omega)(1+e^{\beta E})$$
$$\times[S^{\overline{\sigma\sigma}}(q,\omega)c_{\sigma}\bar{I}_{k-q,\bar{\sigma}}(E-\omega)$$
$$+S^{\overline{\sigma\sigma}}(q,\omega)c_{\sigma}I_{k-q,\sigma}(E-\omega)], \tag{25}$$

where

$$S^{\overline{\sigma\sigma}}(q,\omega)=-\frac{1}{\pi}\mathrm{Im}\langle\langle X^{\overline{\sigma\sigma}}(q)|X^{\overline{\sigma\sigma}}(-q)\rangle\rangle_{\omega}, \tag{26}$$

$$S^{\overline{\sigma\sigma}}(q,\omega)=-\frac{1}{\pi}\mathrm{Im}\langle\langle X^{\overline{\sigma\sigma}}(q)|X^{\overline{\sigma\sigma}}(-q)\rangle\rangle_{\omega}, \tag{27}$$

and

$$F(E)=(e^{\beta E}+1)^{-1},\quad N(\omega)=(e^{\beta\omega}-1)^{-1}.$$

A similar approximation was done by Plakida *et al.*[8] (see also Ref. 9) for the t–J model. It corresponds to the self-consistent Born approximation.

In the N-state, all the main characteristics are independent of the spin projections ($c_{\sigma}=c=1-n/2$, $I_{p\sigma}(E)=I_{p}(E)$, etc.), and at $T=0$ we have an expression for the imaginary part of the mass operator:

$$\Gamma_{k}(E)=\pi\frac{1}{N}\sum_{q}\omega^{2}_{k-q}\int^{\infty}_{-\infty}d\omega\,[S(q,\omega)+\tilde{S}(q,\omega)]$$
$$\times[\theta(E)\theta(\omega)\theta(E-\omega)-\theta(-E)\theta(-\omega)$$
$$\times\theta(|E|-|\omega|)]I_{k-q}(E-\omega), \tag{28}$$

with the obvious redefinitions for $S(q,\omega)$ (Eq. (26)) and $\tilde{S}(q,\omega)$ (Eq. (27)). Thus, the imaginary part of the mass operator in the self-consistent Born approximation is represented by a convolution of the spectral intensities of quasi-Bose and quasi-Fermi excitations in the frequencies ω and momenta q. The real part of the mass operator is a Hilbert transform of $\Sigma_{k\sigma}(E)$, i.e.,

$$M_{k}(E)=\frac{1}{\pi}\mathscr{P}\int^{\infty}_{-\infty}\frac{\Gamma_{k}(E')\,dE'}{E-E'}. \tag{29}$$

If the functions (26) and (27), which describe the transverse and longitudinal components of spin density fluctua-

tions, have been calculated (at least approximately), the mass operator can be found self-consistently from (28) and (29), the representations (20) and (21), and the equation for the chemical potential

$$n=\frac{1}{N}\sum_{k\sigma}\langle X_{k\sigma}^{\dagger}X_{k\sigma}\rangle=(2-n)\frac{1}{N}\sum_{k}\int_{-\infty}^{0}I_k(E)\,dE. \tag{30}$$

Note the difference between a true singlet state and a normal (paramagnetic) state. In addition to $\hat{S}^i|0\rangle$ vanishing in the singlet state, this state has a nontrivial, isotropic spin correlation function

$$\langle S_f^x S_{f+r}^x\rangle=\langle S_f^y S_{f+r}^y\rangle=\langle S_f^z S_{f+r}^z\rangle\equiv C(r), \tag{31}$$
$$\langle S_f^+ S_{f+r}^-\rangle=2C(r),$$

which is independent of the direction of vector r. If there is no such correlation ($C(r)=0$), we have an ordinary paramagnetic state. Equations (31) are also true for the correlation functions $\langle S^i(q)S^i(-q)\rangle$ and the corresponding Green's functions.

4. THE ZEROTH APPROXIMATION AND THE RANDOM PHASE APPROXIMATION

Let us start with the zeroth approximation, i.e., the N-state and $T=0$, where in describing the electron states we ignore the mass operator. In this case the spectral intensity is a delta function and the distribution function, a Fermi step function:

$$I_p^0(E)=\delta(E-\xi_p),\quad n_p^0=cf_p^0=c\,\theta(-\xi_p),$$
$$\xi_p\approx\Delta+c\omega_p-\mu. \tag{32}$$

Here $c=1-n/2$, $\Delta=|\epsilon_0|/2c$, $\epsilon_0(n)$ is the system energy (per site), and we have used an approximate expression for the one-particle spectrum (the correlation function κ is dropped from (19)). At this point it proves convenient to introduce the effective chemical potential $m=(\mu-\Delta)/c$. Then m and the system energy as functions of the electron concentration n are, respectively,

$$\frac{n}{2-n}=\int_{-1}^{m}\rho(\omega)\,d\omega\equiv g(m), \tag{33a}$$

$$\frac{\epsilon_0(n)}{2-n}=\int_{-1}^{m}\omega\rho(\omega)\,d\omega\equiv v(m). \tag{33b}$$

By excluding the upper limit m we can obtain the explicit dependence of ϵ_0 on n. Analysis shows[6,7] that in the zeroth approximation $\epsilon_0(n)<\epsilon_{\rm FM}(n)$ for all concentrations, i.e., the singlet (nonmagnetic) state is the energy-advantageous one. Moreover, Eq. (33a) allows correctly for the excluded-volume effect in the Hubbard model with $U=\infty$: the Fermi surface is inflated compared to the case of free electrons ($n-2g(m)$ for such electrons) and occupies the entire Brillouin zone as $n\to1(m\to1)$.

The spectral densities of the quasi-Bose states in the mass operator can be found in the random phase approximation. In the N-state, the commutator Green's functions of transverse and longitudinal spin fluctuations are[7]

$$\langle\langle S^+(q)|S^-(-q)\rangle\rangle_\omega=\langle\langle S^-(q)|S^+(-q)\rangle\rangle_\omega=\frac{D_0(q,\omega)}{1+P(q,\omega)}\equiv D(q,\omega), \tag{34a}$$

$$\langle\langle S^z(q)|S^z(-q)\rangle\rangle_\omega=\frac{1}{2}D(q,\omega), \tag{34b}$$

and

$$\langle\langle X^{\overline{\sigma\sigma}}(q)|X^{\overline{\sigma\sigma}}(-q)\rangle\rangle=\langle\langle X^{\sigma\sigma}(q)|X^{\sigma\sigma}(-q)\rangle\rangle_\omega=\frac{D_0(q,\omega)}{1-P^2(q,\omega)}\equiv\tilde{D}(q,\omega), \tag{35}$$

where

$$D_0(q,\omega)=-\chi_0^{+-}(q,\omega)=\frac{1}{N}\sum_p\frac{n_p-n_{p+q}}{\omega-c\omega_{pq}},$$

$$P(q,\omega)=\frac{1}{N}\sum_p\frac{n_{p+q}\omega_{p+q}-n_p\omega_p}{\omega-c\omega_{pq}}, \tag{36}$$

$$\langle X_{p\sigma}^{\dagger}X_{p\sigma}\rangle=\langle X_{p\bar{\sigma}}^{\dagger}X_{p\bar{\sigma}}\rangle=n_p,\quad \omega_{pq}=\omega_{p+q}-\omega_p.$$

Equation (34b) points to spin isotropy (see Eq. (31)), which is characteristic of a singlet state. In this state the correlator

$$\langle X^{\sigma\bar{\sigma}}(q)X^{\bar{\sigma}\sigma}(-q)\rangle+\langle X^{\overline{\sigma\sigma}}(q)X^{\overline{\sigma\sigma}}(-q)\rangle=3\langle S^z(q)S^z(-q)\rangle+\tfrac{1}{4}\langle\rho(q)\rho(-q)\rangle, \tag{37}$$

where $\rho(q)$ is the operator of electron number density fluctuations. In the random phase approximation we have

$$\langle\langle\rho(q)|\rho(-q)\rangle\rangle_\omega=\frac{D_0(q,\omega)}{1-P(q,\omega)}\equiv\overline{D}(q,\omega), \tag{38}$$

so that in the singlet state the sum of the spectral intensities can be written as follows:

$$S(q,\omega)+\tilde{S}(q,\omega)=\tfrac{3}{2}S(q,\omega)+\tfrac{1}{4}\overline{S}(q,\omega), \tag{39}$$

where $\overline{S}(q,\omega)$ is the imaginary part of the function (38).

If we use the Fermi step function $f_p^0=\theta(-\xi_p)=\theta(m-\omega_p)$ to calculate the transverse susceptibility $\chi^{+-}(q,\omega)=-D(q,\omega)$, of the Hubbard electrons in the static case ($\omega=0$) for $q\to0$ we obtain[7]

$$\lim_{q\to0}\chi^{+-}(q,0)=\frac{\rho(m)}{1-g(m)+m\rho(m)}\equiv u(m)\rho(m), \tag{40}$$

where $u(m)$ is the paramagnetic gain factor. Equation (40) has no pole singularity, which is an indication that the N-state is stable against the development of ferromagnetism; this agrees with the results of Ref. 10.

Thus, the zeroth approximation and the random phase approximation are in full agreement (the absence of ferromagnetism in the Hubbard model with $U=\infty$ in the thermodynamic limit).

5. APPROXIMATIONS USED IN CALCULATING THE MASS OPERATOR AND THE DISTRIBUTION FUNCTION

Let us discuss the role of kinematic correlations in N-state of the system. Such correlations are described by the mass operator $\Sigma_k(E)$, and finding them self-consistently constitutes an extremely difficult problem, whose solution involves complicated numerical calculations. For this reason we introduce a number of simplifying assumptions, which, we believe, do not change the main conclusions and the essence of the problem.

1. The denominators of the functions $D(q,\omega)$, $\tilde{D}(q,\omega)$, and $\overline{D}(q,\omega)$ do not vanish, which is an indication that the model does not contain well-defined collective excitations (magnons or zero-point sound). This makes it possible to approximately replace the sum of spectral intensities (instead of using the random phase approximation) by

$$S(q,\omega)+\tilde{S}(q,\omega)\approx 2AS_0(q,\omega), \tag{41}$$

which is equivalent to considering the susceptibility of independent Hubbard electrons. In Eq. (41), A is a correction factor depending on the concentration (or on the chemical potential). With a distribution function of the general form $n_p=cf_p=cf(\omega_p)$ we have

$$\begin{aligned}-\frac{1}{\pi}\operatorname{Im}D_0(q,\omega)&=S_0(q,\omega)\\ &=\frac{c}{N}\sum_p f(\omega_p)[1-f(\omega_{p+q})]\\ &\quad\times[\delta(\omega-c\omega_{pq})-\delta(\omega+c\omega_{pq})]\\ &\equiv S_0^{(+)}(q,\omega)-S_0^{(-)}(q,\omega).\end{aligned} \tag{42}$$

This function, known as the dynamic form factor,[11] is defined for both positive ($\omega>0$) and negative ($\omega<0$) frequencies, since $\omega_{pq}>0$ (obviously, $S_0^{(-)}(q,-\omega)=S_0^{(+)}(q,\omega)$), and describes incoherent single-pair electron excitations with the spectrum ω_{pq} within the first Brillouin zone.

The correcting factor A can be found from the sum rule

$$\frac{1}{N}\sum_q\langle S_q^+S_{-q}^-\rangle=\frac{1}{N}\sum_q\langle X_q^{\sigma\sigma}X_{-q}^{\sigma\sigma}\rangle=\frac{n}{2}. \tag{43}$$

Using (41) and the spectral theorem, at $T=0$ we have

$$\langle S_q^+S_{-q}^-\rangle=\frac{Ac}{N}\sum_p f_p(1-f_{p+q}). \tag{44}$$

If we take the Fermi step function $f_p^0=\theta(m-\omega_p)$ as the zeroth distribution function with the effective chemical potential m and replace summation by integration with the electron state density $\rho(\omega)$, we find

$$\begin{aligned}\frac{n}{2}&=Ac\int_{-1}^{1}d\omega\,\rho(\omega)\theta(m-\omega)\int_{-1}^{1}d\omega'\,\rho(\omega')\theta(\omega'-m)\\ &=Acg(m)[1-g(m)].\end{aligned}$$

Since $c=1-n/2=[1+g(m)]^{-1}$ and $n/2=g(m)[1+g(m)]^{-1}$, we have

$$A(m)=[1-g(m)]^{-1}. \tag{45}$$

After this approximation, the imaginary part of the mass operator can be written as follows ($S_0^+(q,\omega)\equiv S_0(q,\omega)$ for $\omega\geqslant 0$):

$$\Gamma_k(E)\approx\begin{cases}\Gamma_k^+=\dfrac{2A\pi}{N}\displaystyle\sum_q\omega_{k-q}^2\int_0^E d\omega\,S_0(q,\omega)I_{k-q}(E-\omega),\\ E>0,\\ \Gamma_k^-=\dfrac{2A\pi}{N}\displaystyle\sum_q\omega_{k-q}^2\int_0^{|E|}d\omega\,S_0(q,\omega)I_{k-q}(-|E|+\omega),\\ E<0.\end{cases} \tag{46}$$

We see that the attenuation $\Gamma_k(E)$ vanishes as $E\to 0$. Note that the Fermi level corresponds to $E=0$, so that there is no attenuation on the Fermi surface proper.

2. Instead of the dynamic form factor $AS_0(q,\omega)=S(q,\omega)$ we consider its value averaged over all the momenta:

$$S(q,\omega)\to\frac{A}{N}\sum_q S_0(q,\omega)\equiv S(\omega,m), \tag{47}$$

which is the number density of single-pair excitations with an energy ω and a chemical potential m. Using the Fermi step function, introducing a new variable, $\Omega=\omega/c$, and employing Eqs. (42) and (47), we get

$$S(\Omega,m)=A(m)\int_{m-\Omega}^{m}\rho(x)\rho(x+\Omega)\,dx,\quad 0\leqslant\Omega\leqslant 2. \tag{48}$$

We see that $S(\Omega,m)\approx A(m)\rho^2(m)\Omega$ as $\Omega\to 0$. The shape of the function $S(\Omega,m)$ is depicted in Fig. 1 for an elliptical density of electron states, $\rho(x)=(2/\pi)\sqrt{1-x^2}$. Analysis of (48) shows that the behavior and numerical values of $S(\Omega,m)$ do not change significantly with other densities of states for two- and three-dimensional alternant lattices.

The essence of the approximation (47) is as follows. The general expressions (25) and (28) describe processes of the following type: an electron with an ''energy'' E and momentum k passes to the state $E-\omega$, $k-q$, exciting in the process an electron–hole pair with momentum q and energy $\omega=\omega_{pq}$. When the approximation (47) is used, we ignore the detailed description of states in the momenta and specify the transition of the electron from the constant-energy surface E to the constant-energy surface $E-\omega$; such transitions take place with different momenta q, which is reflected in (47). Since the description of states is done on a constant-energy

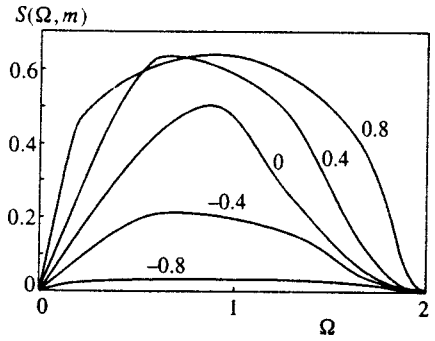

FIG. 1. The density of single-pair excitations for different values of the chemical potential m(Ω is the excitation energy). Here an elliptical density of electron states is used.

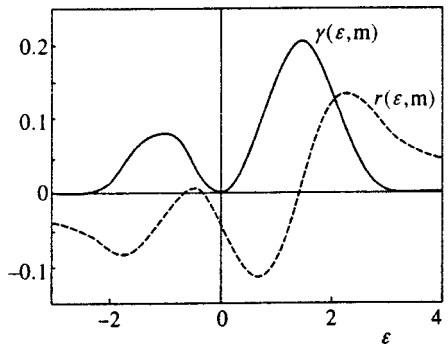

FIG. 2. The real ($r(\varepsilon,m)$) and imaginary ($\gamma(\varepsilon,m)$) parts of the mass operator in the first Born approximation as functions of the spectral variable ε at $m=-0.4$.

surfaces, the dependence of the attenuation and the entire mass operator on the quasimomentum k is lost, i.e., $\Sigma_k(E) \to \Sigma(E) \equiv M(E) - i\Gamma(E)$. Now, setting E at $c\varepsilon$, we obtain

$$\Gamma(\varepsilon) \approx \begin{cases} 2\pi \int_0^{\varepsilon} d\Omega\, S(\Omega,m) \dfrac{1}{N} \sum\limits_p \omega_p^2 I_p(\varepsilon-\Omega), & \varepsilon>0, \\ 2\pi \int_0^{|\varepsilon|} d\Omega\, S(\Omega,m) \dfrac{1}{N} \sum\limits_p \omega_p^2 I_p(-|\varepsilon|+\Omega), & \varepsilon<0, \end{cases} \tag{49}$$

and

$$M(\varepsilon) = \mathrm{Re}\,\Sigma(\varepsilon) \approx \frac{1}{\pi} \mathscr{P} \int_{-\infty}^{\infty} \frac{\Gamma(E')\,dE'}{\varepsilon - E'}. \tag{50}$$

3. Let us study the first Born approximation (the first iteration step in the self-consistent solution):

$$I_p(\varepsilon-\Omega) \to I_p^0(\varepsilon-\Omega) = \delta(\varepsilon-\Omega-(\omega_p-m)).$$

The sum in (49) is

$$\begin{aligned} \frac{1}{N} \sum_p \omega_p^2 I_p^0(\varepsilon-\Omega) &\approx \frac{1}{zN} \sum_p I_p^0(\varepsilon-\Omega) \\ &= \frac{1}{z} \int_{-1}^{1} dx\, \rho(x)\, \delta[\varepsilon-\Omega+m-x] \\ &= \frac{1}{z} \rho(\varepsilon-\Omega+m), \end{aligned}$$

where

$$\frac{1}{z} = \frac{1}{N} \sum_p \omega_p^2 = \int_{-1}^{1} \omega^2 \rho(\omega)\, d\omega \tag{51}$$

is the average value of the square of the bare spectrum over the entire Brillouin zone. In this case the attenuation is

$$\gamma(\varepsilon,m) = \Gamma^{(1)}(\varepsilon) \approx \begin{cases} \dfrac{2\pi}{z} \int_0^{\varepsilon} d\Omega\, S(\Omega,m) \rho(\varepsilon-\Omega+m), \\ 0 \leqslant \varepsilon \leqslant 3-m, \\ \dfrac{2\pi}{z} \int_0^{|\varepsilon|} d\Omega\, S(\Omega,m) \rho(-|\varepsilon|+\Omega+m), \\ -(3+m) \leqslant \varepsilon \leqslant 0, \end{cases} \tag{52}$$

with $-1 \leqslant m \leqslant 1$.

As usual, on the Fermi surface $\gamma(\varepsilon) \to 0$ as $\varepsilon \to 0$. Near the Fermi surface ($\varepsilon \to 0$, so that $\Omega \to 0$) we have

$$S(\Omega,m) \approx A(m) \rho^2(m) \Omega, \quad \rho(\varepsilon-\Omega+m) \approx \rho(m),$$

and

$$\gamma(\varepsilon,m) \approx \frac{\pi}{z} A(m) \rho^3(m) \varepsilon^2. \tag{53}$$

It should be recalled that in our calculations we deal with dimensionless units; actually, $1/z$ has the dimensions of energy squared and γ, of energy. The functional dependence of (53) on ε agrees with Landau's theory of a Fermi liquid[12] (indeed, if we put $\varepsilon = \varepsilon_p - \varepsilon_F = (p^2 - p_F^2)/2m \approx v_F(p-p_F)$, we find that $\gamma \propto (p-p_F)^2$) and with the results of quasiparticle lifetime calculations in the generalized random phase approximation (see Ref. 11). The results of calculations of $\gamma(\varepsilon,m)$ and of $r(\varepsilon,m)$, the real part of the mass operator (the Hilbert transform of $\gamma(\varepsilon,m)$), are depicted in Fig. 2.

Within these approximations, the spectral intensity can be written as follows:

$$I_p(\varepsilon) = I(\varepsilon,\omega_p,m) = \frac{1}{\pi} \frac{\gamma(\varepsilon,m)}{[\varepsilon-(\omega_p-m)-r(\varepsilon,m)]^2 + \gamma^2(\varepsilon,m)}. \tag{54}$$

It has been proved, both numerically and analytically, that the spectral intensity obeys the sum rule

$$\int_{-\infty}^{\infty} d\varepsilon\, I(\varepsilon,\omega_p,m) = 1$$

for arbitrary ω_p and m. The distribution function

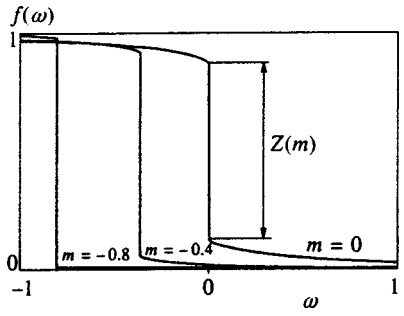

FIG. 3. Typical shape of the distribution function in the Hubbard model with $U=\infty$ at $T=0$.

$$n_k = cf_k = cf(\omega_k, m), \quad f(\omega_k, m) = \int_{-\infty}^{0} I(\varepsilon, \omega_k, m)\, d\varepsilon, \tag{55}$$

has a discontinuity at $\omega_k = m$ (Fig. 3). Equation (30) for the chemical potential with the distribution function (55) reduces to

$$\frac{n}{n-2} = \int_{-1}^{1} \rho(\omega) f(\omega, m)\, d\omega. \tag{56}$$

From (56) we see that to a high degree of accuracy the chemical potential calculated in the first Born approximation does not differ from m. Figure 4 depicts the dependence of the discontinuity $Z(n)$ on the electron concentration n. The energies of the saturated ferromagnetic state, $\epsilon_{\mathrm{FM}}(n)$, and of the N-state in the zeroth approximation, $\epsilon_0(n)$, and in the first Born approximation, $\epsilon_1(n)$, are depicted in Fig. 5. We see that although the kinematic correlations of the electrons in the N-state raise the energy of the system in comparison to the zeroth approximation, the N-state is still the most advantageous one energetically.

Let us discuss the spin structure of this state. Spin isotropy, which is a characteristic feature of the singlet state (34b), is retained if we use the approximation (41). The spin correlation function for the nearest neighbors (in the first coordination sphere) is

$$C_1 = \frac{1}{2N} \sum_q \gamma_q \langle S_q^+ S_{-q}^- \rangle. \tag{57}$$

In the approximation (41) we have

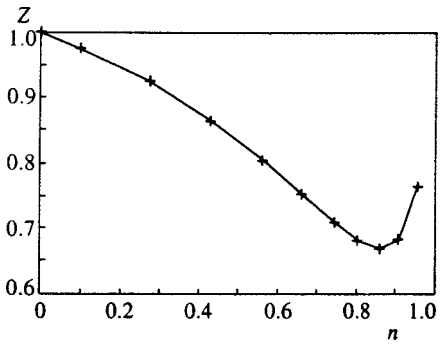

FIG. 4. Dependence of the discontinuity on the Fermi surface on the electron concentration n.

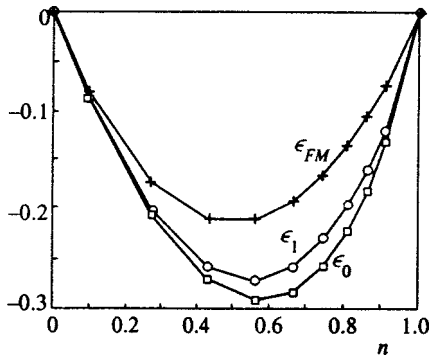

FIG. 5. The energies of the saturated ferromagnetic state, $\epsilon_{\mathrm{FM}}(n)$, the normal state in the zeroth approximation, $\epsilon_0(n)$, and the normal state in the first Born approximation, $\epsilon_1(n)$.

$$C_1 = \frac{cA}{2N} \sum_q \gamma_q \frac{1}{N} \sum_p f_p^0 (1 - f_{p+q}^0)$$

$$= \frac{cA}{2N} \sum_p \theta(m - \omega_p) \frac{1}{N} \sum_k \gamma_{k-p} \theta(\omega_k - m). \tag{58}$$

For the square and cubic lattices there is the well-known symmetry effect of ''splitting'' of γ_{k-p}, i.e.,

$$C_1 = \frac{cA}{2N} \sum_p [\theta(m - \omega_p) \gamma_p] \frac{1}{N} \sum_k \gamma_k \theta(\omega_k - m)$$

$$= \frac{cA}{2} \left(\int_{-1}^{m} \rho(\omega) \omega\, d\omega \right) \left(\int_{m}^{1} \rho(\omega') \omega'\, d\omega' \right)$$

$$= -\frac{1}{2} \frac{v^2(m)}{1 - g^2(m)}. \tag{59}$$

Using (33), we finally get

$$C_1 = -\frac{1}{8} \frac{\epsilon_0^2(n)}{1-n}. \tag{60}$$

If, e.g., we take the ''rectangular'' density of states $\rho(\omega) = 1/2$ we have $\epsilon_0(n) = -n(1-n)(1-n/2)^{-1}$, so that

$$C_1(n) = -\frac{1}{8} \frac{n^2(1-n)}{(1-n/2)^2}.$$

In the present model the correlations are due solely to electrons hopping to neighboring vacant sites and, naturally, $C_1(n) \to 0$ as $n \to 1$.

Thus, our N-state is a singlet state with a nontrivial isotropic spin correlation function.

6. CONCLUSION

Our analytical and numerical analysis of the Hubbard model for the limit of infinite repulsion at a single site ($U = \infty$) has shown that qualitatively the properties of the electron system are the same as those of a normal metal with electron–electron coupling. What is important here is that the kinematic electron correlations do not disrupt the Fermi surface, but leave a ''signature'' in the form of a Migdal discontinuity in the distribution function. Although this re-

sult has been obtained in the first Born approximation, we believe that it is not altered by further iterations, since the attenuation (the imaginary part of the mass operator) always vanishes on the Fermi surface, i.e., for a zero value of the spectral variables E.

Our numerical calculations have shown that the normal (nonmagnetic) singlet state is the ground state. The main reason for this is that the chemical potential is lower than that of the saturated ferromagnetic state, which is indicated already by the zeroth approximation. Our study was done in the thermodynamic limit, with the result that the subtle quantum mechanical effect described by the Nagaoka theorem,[13] which concerns the ground ferromagnetic state in the presence of one hole, does not appear in this limit. In this connection we would like to mention To'th's paper,[14] where it was stated that in the presence of two holes the ferromagnetic state is not the ground state.

In the model with $U=\infty$ we considered only the N-state as the main contender for the ground state of the system. Here the state with the electron concentration $n=1$ is the ''punctured'' point with zero energy and high degree of degeneracy. At finite but large values of Coulomb repulsion ($U \gg t$), the degeneracy is lifted by the antiferromagnetic exchange interactions $J \approx t^2/U$, and long-range antiferromagnetic order emerges at low hole concentrations.[9]. However, as the concentration grows, long-range antiferromagnetic order disappears and the system passes into its normal (metallic) state of the type described above.

The authors would like to express their gratitude to V. V. Val'kov, S. G. Ovchinnikov, and A. L. Pantikyurov for fruitful discussions. This work was supported by a grant from the Krasnoyarsk Regional Science Fund (Grant No. 7F0179).

*)E-mail: andrew@post.krasscience.rssi.ru

[1] J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 238 (1963); **277**, 237 (1964); **281**, 401 (1964).
[2] A. B. Migdal, *Qualitative Methods in Quantum Theory*, Addison-Wesley, Reading, Mass. (1977).
[3] P. W. Anderson, Phys. Rev. Lett. **64**, 1839 (1990).
[4] F. Haldane, J. Phys. C **14**, 2585 (1981).
[5] S. V. Tyablikov, *Methods in the Quantum Theory of Magnetism*, Plenum Press, New York (1967).
[6] E. V. Kuz'min, Fiz. Met. Metalloved. **81**(5), 33 (1996) [Phys. of Metals and Metallography **81**, 480 (1996)].
[7] E. V. Kuz'min, Fiz. Tverd. Tela (St. Petersburg) **39**(2), 193 (1997) [Phys. Solid State **39**, 169 (1997)].
[8] N. M. Plakida, V. S. Oudovenko, and V. Yu. Yushankhai, Phys. Rev. B **50**, 6431 (1994).
[9] Yu. A. Izyumov, Usp. Fiz. Nauk **167**, 465 (1997) [Phys. Usp. **40**, 445 (1997)].
[10] J. Hubbard and K. P. Pain, J. Phys. C, Ser. 2 **1**, 1650 (1968).
[11] D. Pines and F. Nozieres, *Theory of Quantum Liquids*, W. A. Benjamin, Reading, Mass. (1966).
[12] E. M. Lifshitz and L. P. Pitaevskiĭ, *Statistical Physics*, Part. 2, 3rd edn., Pergamon Press, Oxford (1981).
[13] Y. Nagaoka, Phys. Rev. **147**, 392 (1966).
[14] B. To'th, Lett. Math. Phys. **22**, 321 (1991).

Translated by Eugene Yankovsky

Electronic structure and local interactions on a Si(100)2×1 surface with submonolayer Ba overlayers

G. V. Benemanskaya*) and D. V. Daĭneka

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

G. É. Frank-Kamenetskaya

St. Petersburg State Technological Institute, 198013 St. Petersburg, Russia

(Submitted 14 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2145–2152 (December 1998)

The electronic structure and ionization energy for the system Ba/Si(100)2×1 have been studied as functions of the submonolayer coverage. It is found that there is an energy gap in the surface states spectrum and that the Ba/Si(100)2×1 interface is semiconducting up to 1.5 monolayers of Ba. Two surface bands induced by Ba adsorption have been detected. The evolution of the spectrum with increasing degree of Ba coverage points to the existence of two nonequivalent ''adsorption sites,'' which differ in binding energy by 0.11 eV. The development of the Ba-induced bands is found to terminate at a coverage corresponding to the minimum ionization energy and close to one monolayer. The adsorption bond is shown to have a primarily covalent character. © *1998 American Institute of Physics.* [S1063-7761(98)01812-5]

1. INTRODUCTION

A study of the structural and electronic properties of adsorbed metal coatings on silicon surfaces is of applied as well as fundamental importance, thanks to the potential for the use of such systems in nanotechnological applications. In this regard, it is of especial interest to study at the atomic level the interaction of the active dangling bonds of reconstructed silicon surfaces with metal atoms.

Among phenomena on semiconductor surfaces, the most studied have been changes in the spectrum of surface states during adsorption, the correlation between the structural and electronic properties, and the nature of the adsorption bond. Adsorption of metals can alter the electronic properties of the surface, inducing the appearance of new surface states of metallic or semiconductor type. As a result, the nascent interface either preserves the original conductivity type of the substrate or alters it. In this connection keen interest has arisen over the question of the relative importance of the effect of the adsorbate and the substrate on the electronic structure of ultrathin interfaces. Thus, studies in recent years have shown that adsorption of Cs onto a reconstructed Si(111)7×7 surface leads to the formation of ultrathin interfaces for which the surface conductivity has a metallic character.[1–3] A qualitatively different situation is observed for the adsorption of Ba on the same surface. It has been found that an insignificant concentration of Ba adatoms, less than 0.2 monolayers, on a Si(111)7×7 surface leads to the destruction of the metallic conductivity of the substrate, and the interface has a semiconducting character.[1] These studies show that in the case of adsorption of metals on a Si(111)7×7 surface the dominant role in the process of formation of the electronic properties of the metal–semiconductor interface belongs to the adsorbate.

A dimer-reconstructed Si(100)2×1 surface has two active dangling bonds per dimer, and their concentration $n=6.78\times10^{14}\,\mathrm{cm}^{-2}$ exceeds the concentration of dangling bonds on a Si(111)7×7 surface by more than a factor of two.[4] Therefore substantial differences may be expected in the adsorption properties of these surfaces. At present, most experimental and theoretical studies confirm the model of asymmetric dimers for a reconstructed Si(100)2×1 surface. In this model atoms of neighboring atomic rows on an ideal Si(100)1×1 surface form dimers during reconstruction, thereby lowering the number of dangling bonds by a factor of two. Each dimer atom has one dangling bond. The dimers so formed are asymmetric and tilted, and in connection with this there takes place a partial transfer of charge from the lower atom of the dimer to the upper. Such a relaxation is accompanied by a rehybridization of the surface bonds and a redistribution of surface charge. As a result, the electronic structure of a Si(100)2×1 surface has a semiconducting character whereas in the case of symmetric dimers the active dangling bond of each atom is partially filled with one electron, which should lead to an electronic structure of metallic type.[4–6] An atomically clean Si(100)2×1 surface has three surface states: S_1, S_2, and S_3, situated respectively 0.8, 1.0, and 3.5 eV below the Fermi level E_F (Refs. 5 and 6). The surface states S_1 and S_3 are ascribed respectively to the dangling bond and the valence bond of the upper atom of the dimer. The nature of the S_2 state is still unclear.

The high reactivity and the behavior of the geometrical and electronic structure have heightened interest in studies of local interactions of the metal adatoms on a Si(100)2×1 surface. For adsorption of alkali metals, the dimer surface structure, as a rule, is preserved, and the adatoms can occupy different sites on the surface, the so-called adsorption sites.

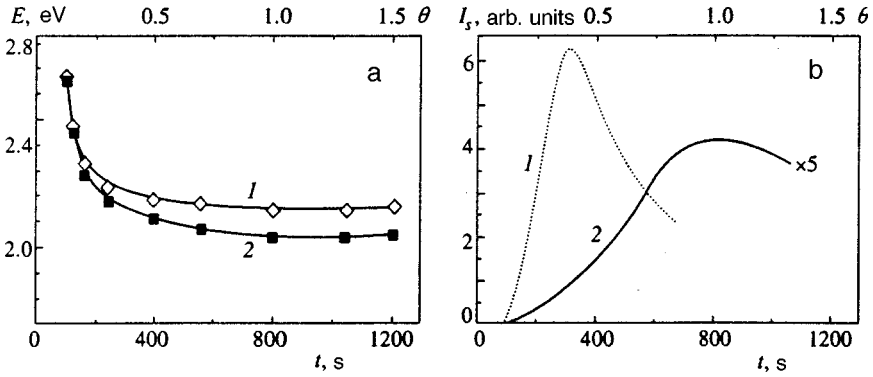

FIG. 1. Dependence on the Ba sputtering time of: a) the photoemission thresholds $h\nu_s=\phi$ (curve *1*) and $h\nu_p$ (curve *2*) for excitation of samples of Ba/Si(100)2×1 by *s*- and *p*-polarized light; b) the total photoemission current I_s for the systems Ba/W(110) (curve *1*) and Ba/Si(100)2×1 (curve *2*) for different degrees of Ba coverage θ (upper scale, in units of monolayers).

Despite the great variety of experimental studies on this subject, many questions touching on the number of adsorption sites, the type of conductivity of the nascent interface, values of the saturated coverage, and the nature of the adsorption bond for adsorption of alkali metals on a Si(100)2×1 surface remain open. Studies of the electronic structure of ultrathin Ba/Si(100)2×1 interfaces are still lacking. Reference 7 investigated adsorption of Ba on a Si(100)2×1 surface by methods of electron loss spectroscopy and thermal desorption, and also measured the work function.

We have carried out the first studies ever of the electronic structure of the Ba/Si(100)2×1 interface in the interesting energy region near the bulk valence band maximum and the Fermi level, together with a determination of the ionization energy for submonolayer coverages of 0.2–1.5 monolayers. We used the technique of threshold photoemission spectroscopy. Our results are adequately described by the double-layer model, proposed for adsorption of alkali metals on a Si(100)2×1 surface. We have compared Ba and Cs adsorption on a Si(100)2×1 surface and have shown that the substrate plays a decisive role in the formation of these interfaces.

2. SAMPLES AND EXPERIMENTAL TECHNIQUE

We investigated a sample (*p*-type, 7.5 Ω·cm) of a Si(100) single crystal doped with KDB-7 boron and subjected it to etching in air using the standard technique before carrying out the measurements.[8] A clean Si(100)2×1 surface was obtained after removing the thin oxide layer by annealing to 1230 K in a high vacuum followed by the well-known regime of stepped cooling. Atomically pure barium was deposited onto the surface of the sample from a standard source (a BaNi evaporator, Ref. 9). The intensity of the Ba source was held constant during throughout the experiment. All measurements were performed *in situ* in a superhard vacuum $P\leqslant 1\times 10^{-10}$ Torr, which enabled us to eliminate the effect of residual oxygen as long as the duration of the experiment did not exceed 1800 s. Under these vacuum conditions, experiments were carried out using the technique of threshold photoemission spectroscopy, based on selective *s*- or *p*-polarized excitation of bulk photoemission and photoemission from surface states, and also on the effect of threshold amplification of photoemission from the surface band during quasiresonant excitation. A detailed description of the technique is given in Refs. 1, 10, and 11. The technique possesses high optical resolution ($\Delta E\sim 0.01$ eV) and substantially exceeds other photoemission techniques in sensitivity. Advantages of the technique include its fundamentally nondestructive character. Threshold photoemission spectroscopy allows one to obtain detailed information about the structure of the surface bands near the Fermi level and the maximum of the bulk valence band from the surface photoemission spectra $I_p(h\nu)/I_s(h\nu)$. Simultaneously, we determined the ionization energy ϕ, which is the photoemission threshold for *s*-polarized emission ($h\nu_s=\phi$). In the case of *p*-polarized excitation the photoemission threshold $h\nu_p$ corresponds either to the position of the long-wavelength edge of the surface band (if it is located between the valence band maximum and E_F) or to the position of E_F (if the band has a metallic character and intersects the Fermi level).

3. RESULTS AND DISCUSSION

Ba adsorption on a Si(100)2×1 surface sharply lowers the photoemission threshold. Figure 1a displays curves of the variation of the bulk photoemission threshold $h\nu_s$, i.e., the ionization energy ($h\nu_s=\phi$), and the surface photoemission threshold $h\nu_p$ as functions of the barium sputtering time. We found that over the entire investigated range of Ba coverages the photoemission thresholds are different, $h\nu_s>h\nu_p$, which proves the existence of surface states in the band gap. As the coverage is increased, a substantial increase in the difference between the thresholds from ~0.02 to ~0.1 eV is observed. It can be seen that the segment of abrupt decrease of the ionization energy, which is observed for sputtering times $t<200$ s, is followed by a segment of slow variation of ϕ with a weakly expressed minimum $\phi_{\min}=2.15\pm 0.01$ eV. To accurately determine the Ba sputtering time corresponding to $\phi_{\min}$, we recorded the dependence of the photoemission current I_s for *s*-polarized excitation by a cadmium laser ($h\nu=2.81$ eV) on the Ba sputtering time. We took into account the fact that the maximum of the dependence $I_s(t)$ corresponds to $\phi_{\min}$ (Ref. 11). Figure 1b shows that the maximum of I_s is reached at $t=800$ s (barium sputtering time). The Ba coverage obtained for the given sputtering time corresponds to the so-called saturated coverage θ_{sat}, since by definition θ_{sat} is the coverage needed to achieve a minimum of the work function and of the ionization energy of the system.

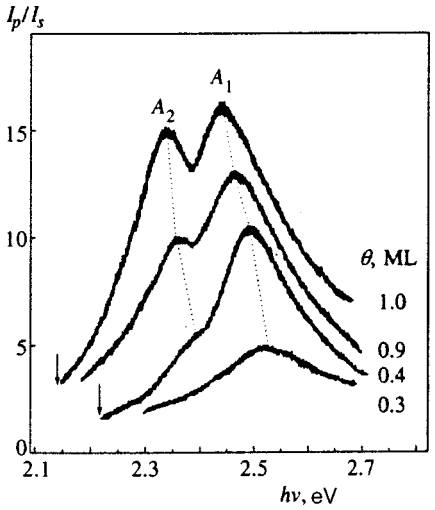

FIG. 2. Surface photoemission spectra I_p/I_s for the system Ba/Si(100)2 $\times$1 for different degrees of submonolayer Ba coverage θ (the degree of coverage is indicated alongside the curves in units of monolayers). Here A_1 and A_2 are the surface bands induced by Ba adsorption; the arrows indicate the position of the the valence band maximum.

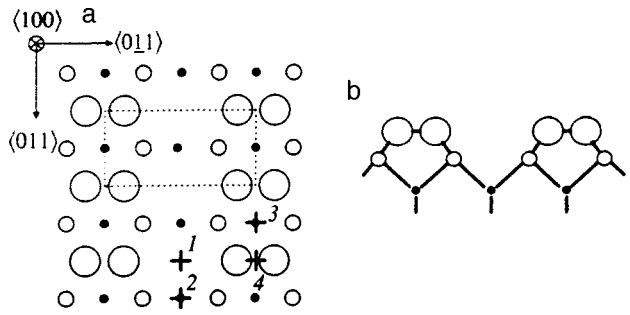

FIG. 3. Schematic depiction of a dimer-reconstructed Si(100)2$\times$1 surface [large empty circles — upper layer (dimer atoms), smaller empty circles — second layer, small filled circles — third layer] and possible adsorption sites (crosses): a — top view, b — side view.

To determine the Ba surface concentration for saturated coverage θ_{sat} we used the technique proposed in Ref. 11. The calculation was based on the experimentally obtained relation [see Fig. 1b] of the Ba sputtering times required to reach the photocurrent maxima for the investigated system Ba/Si(100)2$\times$1 ($t=800$ s) and a reference system Ba/W(110) ($t=300$ s). For the system Ba/W(110) the adatom surface concentration corresponding to the photocurrent maximum is well known and is equal to 2.5×10^{14} atoms/cm^{-2} (Ref. 9). It should also be pointed out that within the limits of one monolayer the Ba sticking coefficient is equal to unity for both surfaces. Consequently, the experimentally obtained value of the Ba adatom surface concentration on the Si(100) 2$\times$1 substrate at saturation is equal to $\theta_{\text{sat}}=(6.8\pm0.1)\times10^{14}$ atoms/cm^2. The concentration corresponds to one monolayer. Note that a coverage equal to one monolayer is identical for any reconstruction of the Si(100) surface and is defined as the density of silicon atoms on a Si(100)1$\times$1 surface, which is 6.78×10^{14} atoms/cm^2 (Ref. 4). Thus, on the basis of the above experimental data we can estimate the saturated Ba coverage on a Si(100)2$\times$1 surface as one monolayer.

The difference between thresholds for achieving monolayer coverage is 0.11 eV and does not vary further, which indicates that the edge of the surface band remains fixed relative to the maximum of the bulk valence band. Analysis of the surface photoemission $h\nu_p$ near the threshold shows that the density of surface states at the Fermi level E_F is equal to zero over the entire investigated range of coverages, i.e., the edge of the surface band lies between the maximum of the bulk valence band and E_F. Thus, we have established the presence of band gap in the spectrum of surface states of the system Ba/Si(100)2$\times$1 and the absence of metallization of the interface up to Ba coverages $\theta\simeq1.5$ monolayer.

Figure 2 displays surface photoemission spectra $I_p(h\nu)/I_s(h\nu)$ of the system Ba/Si(100)2$\times$1 for different Ba coverages. The spectra reflect the density of surface states located below the valence band maximum. As the Ba coverage is increased, two surface bands, A_1 and A_2, are observed to form. The band A_1 appears at low coverages $\theta\simeq0.2$ monolayer. A maximum at $h\nu=2.53$ eV is already distinctly visible for $\theta=0.3$ in Fig. 2. With growth of the Ba coverage the density of states in band A_1 increases, with abrupt growth being observed in the range of coverages from 0.3 to 0.5 monolayer. The band A_2 with a maximum at $h\nu=2.38$ eV, possessing a smaller binding energy than A_1, is observed in the spectrum for coverages close to 0.4 monolayer. A similar increase of the density of states in the band with growth of the Ba coverage is observed, and when a coverage of one monolayer is reached, the band A_2 is completely formed. The difference in binding energies for the two bands is 0.1 eV. Thus, Ba adsorption on a Si(100)2$\times$1 surface leads to the appearance of two induced bands, whose maxima are located below the maximum of the bulk valence band.

During adsorption, adatoms can occupy different sites on the surface known as adsorption sites (Fig. 3). Calculations reveal the existence of four possible adsorption sites on the Si(100)2$\times$1 surface (Refs. 12 and 13). Two of them are located between rows of dimers and two above the rows of dimers. Adsorption site *1* —a cave—is located between rows of dimers above a silicon atom of the third layer, site *2* —a valley—is located between rows of dimers above a silicon atom of the second layer, site *3*—a pedestal—is located between dimer atoms, and site *4*—a bridge— is located above the dimer atoms. At present, the experimental and theoretical results on adsorption of alkali metals is treated from the point of view of two alternative models. The model of one-dimensional metallic chains was proposed in Ref. 14 for the case of adsorption of Cs on Si(100)2$\times$1. According to this model, the Cs atoms occupy one adsorption site *3*—a pedestal. The interface has a metallic character, and saturated coverage is 0.5 monolayer, i.e., it is reached when the adatoms interact with only half of the active dangling bonds of the substrate. Below we will extend this model to adsorption of K and Na on a Si(100)2$\times$1 surface; however, it is assumed that the most probable adsorption site for the adatoms is site *1*—a cave.[15,16]

On the other hand, for Cs and K adsorption on Si(100) 2$\times$1, the model of a double layer has been proposed on the

basis of various experimental data.[17] In this model, adsorption takes place on two nonequivalent sites: a pedestal (*3*) and a valley (*2*); see Fig. 3. In this case, saturated coverage is one monolayer and corresponds to filling of all active dangling bonds of the substrate. This model was later extended to adsorption of other alkali metals and is supported by a large quantity of experimental[18–20] and also theoretical studies.[21,22] These studies show that the alkali metal/Si(100) 2×1 interface has a semiconducting character for saturated coverage. The difference in the adsorption processes and in the properties of the local interactions posited in the models prove to be fundamental for the formation of a metallic or semiconducting type of electronic structure of the interface.

The experimental data we have obtained for the Ba/Si(100)2×1 interface are adequately described in terms of the double-layer model. The existence of two induced bands indicates that Ba adsorption takes place on two nonequivalent adsorption sites. In the initial stage of formation of the interface, the local interaction of Ba adatoms with the active dangling bonds of the surface takes place primarily on one adsorption site and possesses a larger binding energy. When a coverage $\simeq 0.4$ monolayer is reached, the Ba adatoms begin to occupy a second adsorption site with lower binding energy. As theoretical studies have shown,[13,21] one of these adsorption sites should be in one row with dimers (site *3* or *4*, Fig. 3), and the other, between rows of dimers (site *1* or *2*, Fig. 3). In the model of asymmetric dimers the most probable eventuality is adsorption at site *3* (a pedestal) and the least probable is adsorption at site *4* (a bridge), whereas adsorption sites *1* and *2* are equally likely. Therefore, in the Ba/Si(100)2×1 system adsorption can take place either at sites *1* and *3* or at sites *2* and *3*. For saturated coverage, which is defined as a monolayer, both adsorption sites are occupied, as is confirmed by the complete development of the spectrum of induced surface bands. Ba adsorption leads to an insignificant increase in the density of surface states in the band gap, and the interface has a semiconducting character up to 1.5 monolayers. The adsorption bond has a primarily covalent character with partial charge transfer to the dangling bonds of the substrate.

Figure 4 schematically depicts the position of the electronic surface bands relative to the valence band maximum (II) for ultrathin Ba/Si(100)2×1 interfaces. The upper figure depicts a clean Si(100)2×1 surface, where only one surface state S_1 is depicted, having the lowest binding energy. Upon adsorption of alkali metals, the given surface eigenstate is shifted toward higher binding energies.[18] We assume the presence of a similar shift for Ba adsorption, which frees up an energy region ~ 1 eV below E_F for the formation of the barium-induced bands A_1 and A_2. For monolayer coverage, the low-energy edge of the band A_2 is found between E_F and the valence band maximum, which corresponds to different photoemission thresholds for bulk and surface photoemission [see Fig. 1(a)].

Comparison of the results of studies of Ba and Cs adsorption shows that in both cases the interface has a semiconducting character up to 1.5 monolayers, saturated coverage is one monolayer, and the electronic structure has two induced bands corresponding to two nonequivalent adsorption sites for the Si(100)2×1 surface. All of the results of the present study are in good agreement with the double-layer model, which we proposed previously for the system Cs/Si(100)2×1 (Ref. 23). The observed effects are evidence of the dominant influence of the properties of the Si(100)2×1 substrate on interface formation. Thus it can be asserted that the nature of the local interactions for two such different adsorbates as Cs and Ba is dictated by the Si(100)2×1 surface, in contrast to the Si(111)7×7 surface, where the influence of the adsorbates prevails.

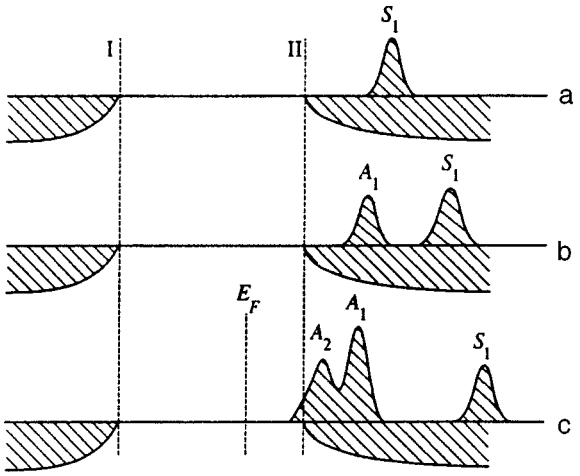

FIG. 4. Schematic diagram of the electronic structure of the surface bands for a clean Si(100)2×1 (a) surface and for the system Ba/Si(100)2×1 for different degrees of submonolayer Ba coverage: 0.3 monolayer (b) and 1 monolayer (c). S_1 — natural surface state of the Si(100)2×1 surface; I — conduction band minimum; II — valence band maximum.

This work was carried out with the support of Grant No. 96-2-28 of the ''Atomic Surface Structures'' program of the Russian Ministry of Science and Grant No. 98-02-18265 of the Russian Fund for Fundamental Research.

*)E-mail: benem@optun.ioffe.rssi.ru

[1] G. V. Benemanskaya, D. V. Daineka, and G. E. Frank-Kamenetskaya, Phys. Low-Dim. Struct. **10/11**, 233 (1995).
[2] K. O. Magnusson, S. Wiklund, R. Dudde, *et al.*, Phys. Rev. B **44**, 5657 (1991).
[3] G. V. Benemanskaya, G. E. Frank-Kamenetskaya, and M. N. Lapushkin, Surf. Sci. **331–333**, 552 (1995).
[4] W. Monch, *Semiconductor Surfaces and Interfaces*, Springer-Verlag, Berlin (1993).
[5] J. Pollmann, P. Kruger, and A. Mazur, J. Vac. Sci. Technol. B **5**, 945 (1987).
[6] L. S. O. Johansson, R. I. G. Uhrberg, P. Vartensson *et al.*, Phys. Rev. B **42**, 1305 (1990).
[7] D. Vlachos, M. Kamaratos, and C. Papageorgopoulos, Solid State Commun. **90**, 175 (1994).
[8] F. Ishizaka and Y. Shiraki, J. Electrochem. Soc. **133**, 666 (1986).
[9] A. G. Fedorus, A. G. Naumovets, and Yu. S. Vedula, Phys. Solid State **13**, 445 (1972).
[10] G. V. Benemanskaya, M. N. Lapushkin, and M. I. Urbakh, Zh. Éksp. Teor. Fiz. **102**, 1664 (1992) [Sov. Phys. JETP **75**, 899 (1992)].
[11] G. V. Benemanskaya, D. V. Daĭneka, and G. É. Frank-Kamenetskaya, JETP Lett. **65**, 729 (1997).
[12] Ye Ling, A. J. Freeman, and B. Delley, Phys. Rev. B **39**, 10144 (1989).
[13] R. Ramirez, Phys. Rev. B **40**, 3962 (1989).
[14] J. D. Levine, Surf. Sci. **34**, 90 (1973).

[15] C. M. Wei, H. Huang, S. Y. Tong *et al.*, Phys. Rev. B **42**, 11284 (1990).
[16] P. Soukiassian, L. Spiess, P. S. Mangat *et al.*, J. Vac. Sci. Technol. B **11**, 1431 (1993).
[17] T. Abukawa and S. Kono, Phys. Rev. B **37**, 9097 (1988); Surf. Sci. **214**, 141 (1989).
[18] A. J. Smith, W. R. Graham, and E. W. Plummer, Surf. Sci. **243**, L37 (1991).
[19] T. Abukawa, T. Kashiwakura, T. Okane *et al.*, Surf. Sci. **303**, 146 (1994).
[20] Y.-C. Chao, L. S. O. Johansson, C. J. Karlsson *et al.*, Phys. Rev. B **52**, 2579 (1995).
[21] I. P. Batra, Phys. Rev. B **39**, 3919 (1989); **43**, 12322 (1991).
[22] H. Ishida and K. Terakura, Phys. Rev. B **40**, 11519 (1989).
[23] G. V. Benemanskaya, D. V. Daineka, and G. E. Frank-Kamenetskaya, Surf. Rev. Lett. **5**, No. 1–2 (1998).

Translated by Paul F. Schippnick

Static displacements of atoms near isotopic impurities and residual resistivity

A. P. Zhernov*)

Kurchatov Institute Russian Scientific Center, 123182 Moscow, Russia
(Submitted 17 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2153–2165 (December 1998)

The question of the residual resistivity of chemically pure metals which are mixtures of atoms of different isotopes is discussed. The question of the static displacements appearing near isotopic impurities due to differences in the zero-point vibrations is analyzed in a microscopic approach. It is shown that such static displacements have a significant influence on the residual resistivity ρ_r. Their contribution to ρ_r is far greater than the contribution due to differences in the dynamic elastic electron scattering amplitudes. © *1998 American Institute of Physics.* [S1063-7761(98)01912-X]

1. INTRODUCTION

Most crystals are mixtures of atoms of isotopes with different atomic weights. However, the synthesis of chemically pure and highly isotopically enriched crystals is possible. In such a case the parameters of the natural and highly enriched compounds can be compared, and the contribution associated with the isotopic impurities can be isolated. We note that investigations of the thermal conductivity of germanium were recently performed with the use of samples of highly enriched ^{70}Ge.[1] In this context, the problem of the residual resistivity of metals, which has been investigated experimentally to an inadequate extent, would be of definite interest in our opinion.

Back in his time, Pomeranchuk[2] pointed out that chemically pure metals, whose crystal-lattice sites are occupied by isotopes of different kinds, exhibit a finite residual resistivity ρ_r at zero temperature due to persistent dynamic disorder (which is caused by fluctuations of the atomic mass from site to site). According to Ref. 2, this resistivity exists because the phonons appearing in virtual states upon electron scattering experience the influence of the isotopic disorder. The corresponding resistivity is proportional to the electron-ion interaction parameter raised to the fourth power. In other words, ρ_r appears in a higher order with respect to the electron-ion interaction than in the case of the Born approximation.

It was subsequently shown[3] that the finite resistivity ρ_r exists in an isotopically disordered lattice even in the Born approximation. Actually, the true elastic electron scattering amplitude on ion n (a_n) is the product of the static amplitude a_0 (which does not vary from site to site) and the dynamic Debye–Waller factor W_n, whose value exhibits a dependence on the mass of the vibrating atom, as a result of which the difference $a_n - a_{n'}$ is nonzero, and a residual resistivity appears in the standard approximation.

In this paper we focus our attention on the fact that a field of static displacements $\{\boldsymbol{\zeta}_n\}$ should exist around isotopic impurities. It is interesting that, generally speaking, the interatomic force parameters do not vary in the vicinity of an isotopic impurity. However, the zero-point vibrations of the atoms make a definite contribution to the binding energy and consequently influence the static configuration of the lattice ions, leading to displacements of the atoms from the equilibrium positions (which are characteristic of the ideal crystal) near an impurity. In addition, the scattering of electrons by static near-impurity displacements also makes a definite contribution to the resistivity.

We note that in the case of classical statistics the value of the mean square of the dynamic atomic displacements $\langle u^2 \rangle$ does not depend on the isotopic composition. The variation of $\langle u^2 \rangle$ and, therefore, of the resistivity ρ_r upon variation of the content of isotopes of different kinds in a crystal is a quantum effect.[3] In the temperature range where classical statistics hold, the factor $\langle u^2 \rangle$ ceases to depend on mass. As a result, the differences between the scattering amplitudes a_n are erased, and the field of specific static displacements $\{\boldsymbol{\zeta}_n\}$ vanishes.

Equations for the static displacements in the general case are obtained in a microscopic approach in Sec. 2. Some estimates are made in a simple lattice model. The residual resistivity ρ_r is discussed in Sec. 3. The contributions associated with differences in the dynamic electron-ion interaction amplitudes and with fields of near-impurity static displacements are analyzed simultaneously. For simplicity, we neglect the mutual influence of isotopic impurities and confine ourselves to the one-impurity approximation.

2. DYNAMIC DISORDER AND STATIC NEAR-IMPURITY DISPLACEMENTS

Let us consider a crystal lattice with an isotopic impurity. The total energy E depends on the coordinates of the ions $\mathbf{R}_n$. In a crystal with impurities the atoms are displaced from their original positions, so that

$$\mathbf{R}_n = \mathbf{R}_n^{(0)} + \boldsymbol{\zeta}_n + \mathbf{u}_n ,$$

where $\mathbf{R}_n^{(0)}$ is the equilibrium position of the respective ion in the ideal lattice, $\boldsymbol{\zeta}_n$ is the static displacement vector, and $\mathbf{u}_n$ denotes the dynamic displacement vector. Taking into account that the displacements are small in comparison to the

interatomic distances, we expand the structure-dependent part of the energy into a series in $\mathbf{u}_n$ and $\boldsymbol{\zeta}_n$ relative to $\mathbf{R}_n^{(0)}$.

The Hamiltonian of a harmonic crystal has the form

$$H=\sum_n \frac{(p_n^\alpha)^2}{2M_n}+\frac{1}{2}\sum_{n_1,n_2}\Phi^{\alpha\beta}_{2,n_1n_2}u^\alpha_{n_1}u^\beta_{n_2}. \tag{1}$$

The following notation has been adopted in (1): $\mathbf{p}_n$ is the momentum of the atom of mass M_n in site n, and $\Phi^{\alpha\beta}_{2,n_1n_2}$ is the second-order force parameter. It is assumed that

$$M_n=M_0+\Delta M\,\delta_{n,0},\quad \Delta M=M_1-M_0.$$

Here M_0 is the mass of an atom of the regular matrix, and M_1 is the mass of the isotope. The origin of coordinates coincides with the equilibrium position of the impurity atom. In addition, here and in the following the doubly repeated Cartesian indices α and β imply summation.

We next define the relaxation energy E_r associated with the near-impurity static displacements ζ_n of the atoms. We note that in the expression for E_r we have retained not only the standard terms that are first- and second-order in ζ_n, but also terms of the form

$$\frac{1}{2}\Phi^{\alpha\beta\gamma}_{3,nn_1n_2}\langle u^\alpha_{n_1}u^\beta_{n_2}\rangle\zeta^\gamma_{n_3}.$$

Here $\Phi^{\alpha\beta\gamma}_{3,nn_1n_2}$ is the third-order force parameter, and $\langle\ldots\rangle$ denotes averaging over the equilibrium thermodynamic distribution.

We note that the existence of the nonzero correlator

$$K^{\alpha\beta}_{n_1n_2}=\langle u^\alpha_{n_1}u^\beta_{n_2}\rangle, \tag{2}$$

which is associated with the zero-point vibrations, is a characteristic feature of the quantum motion. It reflects the fact that the concept of the complete rest of a particle is meaningless in quantum theory.

The energy E_r can be represented symbolically in the usual form as

$$E_r=F\zeta+\frac{1}{2}\Phi_2\zeta^2. \tag{3}$$

In this expression the effective force F which is exerted by an isotopic impurity on matrix atoms and leads to displacements can be represented in the following form:

$$F^\alpha_n=\frac{1}{2}\sum_{n_1n_2}\Phi^{\alpha\beta\gamma}_{3,nn_1n_2}\Delta K^{\beta\gamma}_{n_1n_2}. \tag{4}$$

Here the quantity ΔK is defined as the difference between the correlators K for the lattice with an isotopic impurity and the regular lattice. More specifically,

$$\Delta K^{\alpha\beta}_{n_1n_2}=K^{\alpha\beta}_{n_1n_2}(\Delta M\neq 0)-K^{\alpha\beta}_{n_1n_2}(\Delta M=0). \tag{5}$$

In (4) and (5) the force parameter Φ_3, as well as the correlator K, are assigned relative to the equilibrium positions of the ions in the ideal lattice.

Using (3), from the condition of equality of the effective force acting on an atom, i.e., $\partial E_r/\partial\zeta_n$, to zero we obtain the following system of equations for ζ:

$$F+\Phi_2\zeta=0. \tag{6}$$

On the basis of (6), in the coordinate representation we find that

$$\zeta^\alpha_n\approx-\bar{D}^{\alpha\beta}_{nn'}F^\beta_{n'},\quad \bar{D}=(\Phi_2)^{-1}. \tag{7}$$

In this case the relaxation energy equals $E_r=F\zeta/2$.

Let us obtain an explicit expression for the correlator K, which determines the effective force F in the problem under consideration. For this purpose we introduce the Green's function constructed from dynamic atomic displacement operators

$$D^{\alpha\beta}_{nn'}(t)=-i\,\theta(t)\langle[u^\alpha_n(t),u^\beta_{n'}(0)]\rangle. \tag{8}$$

As is generally known, the correlator K of interest to us can be expressed in terms of the Green's function D using the equality

$$\int_{-\infty}^{\infty}dt\,\exp(-i\omega t)K^{\alpha\beta}_{n_1n_2}(t)=\frac{2\mathrm{Im}\,D^{\alpha\beta}_{n_1n_2}(\omega-i\delta)}{1-\exp(-\omega/T)}. \tag{9}$$

In a situation where there are isolated isotopic impurities at the sites of an harmonic lattice, the Green's function (8) satisfies the equation (see, for example, Ref. 3)

$$D^{\alpha\beta}_{nn'}(\omega)=D^{\alpha\beta}_{0,nn'}(\omega)+\omega^2\times\sum_{n_1}D^{\alpha\gamma}_{0,nn_1}(\omega)(M_0-M_{n_1})D^{\gamma\beta}_{n_1n'}(\omega). \tag{10}$$

In this case the ''zeroth'' Green's function D_0 of the ideal lattice can be represented as

$$D^{\alpha\beta}_{0,nn'}(\omega)=\frac{1}{M_0N}\sum_{\mathbf{q}j}e^\alpha(\mathbf{q},j)e^\beta(\mathbf{q},j)\times\frac{\exp\{i\mathbf{q}(\mathbf{R}^{(0)}_n-\mathbf{R}^{(0)}_{n'})\}}{\omega^2-\omega^2(\mathbf{q},j)}. \tag{11}$$

In (11) $\omega(\mathbf{q},j)$ and $e^\alpha(\mathbf{q},j)$ denote the frequency and polarization vector of a phonon mode with the quasimomentum $\mathbf{q}$ and the polarization j.

Using an iteration method, we can write an approximate solution of Eq. (10) in the form

$$\begin{aligned}D^{\alpha\beta}_{nn'}(\omega)\approx{}&D^{\alpha\beta}_{0,nn'}(\omega)+\omega^2\times\sum_{n_1}D^{\alpha\gamma}_{0,nn_1}(\omega)(M_0-M_{n_1})D^{\gamma\beta}_{0,n_1n'}(\omega)+\omega^4\\&\times\sum_{n_1n_2}D^{\alpha\gamma}_{0,n_1n_2}(\omega)(M_0-M_{n_1})D^{\gamma\gamma_1}_{0,n_1n_2}(\omega)\\&\times(M_0-M_{n_2})D^{\gamma_1\beta}_{0,n_2n'}(\omega)+\ldots.\end{aligned} \tag{12}$$

As a result, in the single-impurity case of interest to us, we can use (12) and (9) to obtain

$$\Delta K^{\alpha\beta}_{n_1 n_2}(t=0) \approx -\frac{\Delta M}{\pi}\int_{-\infty}^{\infty} d\omega \frac{\omega^2}{1-\exp(-\omega/T)} \times \mathrm{Im}[D^{\alpha\gamma}_{0,n_1\mathbf{0}}(\omega) D^{\gamma\beta}_{0,\mathbf{0}n_2}(\omega)]. \quad (13)$$

Knowledge of ΔK, which directly characterizes the dynamic disorder, enables us to determine the field of static displacements on the basis of the relations (7) and (4). We obtain

$$\zeta^{\alpha}(\mathbf{q}) = -\frac{1}{2}\frac{\Delta M}{M_0}\sum_j \frac{e^{\alpha}(\mathbf{q},j)e^{\beta}(\mathbf{q},j)}{M_0\omega^2(\mathbf{q},j)} \times \sum_{\mathbf{q}_1,\mathbf{q}_2,j_1,j_2} \Phi^{\beta\gamma\delta}_{3,qq_1q_2} Z^{\gamma\delta}_{\mathbf{q}_1\mathbf{q}_2}. \quad (14)$$

In order to simplify the form of the expressions, we set

$$Z^{\gamma\delta}_{\mathbf{q}_1\mathbf{q}_2} = \sum_{j_1,j_2} e^{\gamma}(\mathbf{q}_1,j_1)e^{\gamma_1}(\mathbf{q}_1,j_1)e^{\gamma_1}(\mathbf{q}_2,j_2)e^{\delta}(\mathbf{q}_2,j_2)\frac{\omega(\mathbf{q}_1,j_1)[2n(\omega(\mathbf{q}_1,j_1))+1]-\omega(\mathbf{q}_2,j_2)[2n(\omega(\mathbf{q}_2,j_2))+1)]}{M_0[\omega^2(\mathbf{q}_1,j)-\omega^2(\mathbf{q}_2,j_2)]}, \quad (15)$$

where $n(\omega)=[\exp(\omega/T)-1]^{-1}$.

It follows directly from (15), first, that at absolute zero temperature

$$Z^{\beta\gamma}_{\mathbf{q}_1\mathbf{q}_2}(T=0) = \sum_{j_1,j_2} \times \frac{e^{\beta}(\mathbf{q}_1,j_1)e^{\gamma_1}(\mathbf{q}_1,j_1)e^{\gamma_1}(\mathbf{q}_2,j_2)e^{\gamma}(\mathbf{q}_2,j_2)}{M_0[\omega(\mathbf{q}_1,j)+\omega(\mathbf{q}_2,j_2)]}. \quad (16)$$

Second, in the high-temperature limit $Z=0$, and, therefore, the specific field of static displacements vanishes.

Let us obtain some estimates in the model of a linear chain with interactions between nearest neighbors. In such a case the spatial Fourier components for the second- (f_2) and third-order (g_3) force parameters, i.e., $\Phi_{2,q}$ and Φ_{3,qq_1q_2}, are represented as (see, for example, Refs. 4 and 5)

$$\Phi_{2,q} = M_0\omega_q^2, \quad M_0\omega_q^2 = 4f_2\sin^2(ql/2), \quad (17)$$

$$\Phi_{3,qq_1q_2} = -\frac{ig_3}{(f_2/M_0)^{3/2}}\tilde{\omega}_q\tilde{\omega}_{q_1}\tilde{\omega}_{q_2}\Delta(q+q_1+q_2),$$

$$\tilde{\omega}_q = 2\sqrt{\frac{f_2}{M_0}}\sin\frac{ql}{2}. \quad (18)$$

Here l is the lattice constant. The relation between f_2 and g_3 for an integral Grüneisen factor $\gamma_G \approx 2$ in the central-force model is

$$-g_3 l/f_2 \approx 10. \quad (19)$$

Based on (14)–(16) with consideration of (17) and (18), we can show that the displacements of atoms located in the first coordination sphere relative to an isotopic impurity are given by the relation

$$\zeta = -0.3\epsilon\frac{g_3 l^2}{f_2} \sim -\epsilon l, \quad (20)$$

where ϵ is a characteristic parameter of the theory and in the present case

$$\epsilon = \frac{\Delta M}{M_0}\frac{\langle u^2\rangle}{l^2}.$$

Here $\langle u^2\rangle$ denotes the mean square of the dynamic atomic displacements.

It follows from Eq. (20) that as a result of the differences in the zero-point vibrations the lattice undergoes ''compression'' near a heavy isotope and expansion near a light isotope.

In standard systems $\langle u^2\rangle/l^2 \sim 10^{-3}$, and $|\Delta M|/M_0 \leqslant 0.1$. In this case it turns out that $|\epsilon| \leqslant 10^{-4}$. Thus, the static near-impurity displacements are very small compared with the interatomic distances.

It would be interesting to perform estimates for quantum crystals, such as in the case of a mixture of ^{4}He and ^{3}He. We can utilize the data for the Debye temperature $\Theta = 26$ K and the lattice constant $l = 3.57$ Å in Ref. 6. We then have $\langle u^2\rangle/l^2 \approx 3\times 10^{-2}$. Since $|\Delta M|/M_0 = 0.25$ these displacements are appreciable. In the first coordination sphere $\zeta = 0.025$. These displacements are of the same scale as in the case of systems with standard (nonisotopic) impurities.

Let us turn our attention to the fact that distortion of a lattice around light isotopes with resultant renormalization of the force parameters was postulated in Ref. 7 in connection with the problem of the thermal conductivity of solid ^{4}He with ^{3}He impurities. It was found that the displacements of the atoms in the first coordination sphere are described by a formula of the type

$$\tilde{\zeta} = -\frac{1}{3}\frac{\Delta M}{M_0}\left(\frac{3E_0}{3E_0+8\pi\mu r^3}\right),$$

where $E_0 = (8Mr)^{-1}$ is the energy of the zero-point vibrations in a spherical potential well of radius r. The factor $8\pi\mu r^2$ is related to the relaxation energy due to inclusion of the radius r. More specifically, $\tilde{E}_r = 8\pi\mu r^2\tilde{\zeta}^2$, where μ is the compression modulus. According to Ref. 7, $\tilde{\zeta} \approx 0.02$. To a certain degree it is possible that the disparity with the estimate based on Eq. (20) is due to the use of values of the

parameters corresponding to a density of 0.208 g/cm^3 for the estimates in Ref. 7. In Ref. 6 the density was equal to 0.18 g/cm^3.

An explicit expression for $\boldsymbol{\zeta}_n$ is derived in Appendix A in the asymptotic limit where $|R_n^{(0)}|$ is much greater than the interatomic distance l.

3. INFLUENCE OF DYNAMIC LATTICE DISORDER ON THE RESIDUAL RESISTIVITY

Let us consider nontransition metals. Bearing in mind the qualitative aspect of the phenomenon, we assume that the electron Fermi surface is nearly spherical. We write the nonequilibrium electron distribution function $\varphi_\alpha(\mathbf{k})$ in the form $\varphi_\alpha(\mathbf{k}) \propto v_{\mathbf{k}}^\alpha$, where $\mathbf{v_k}$ is the electron group velocity. Then, the residual resistivity can be described by an expression of the form (see, for example, Refs. 8 and 3)

$$\rho_r = \frac{(2\pi)^3}{3m_*^2\Omega_0 J^2}\int_{\sigma_F}\int_{\sigma_F}\frac{d\sigma_k}{v_k}\frac{d\sigma_{k'}}{v_{k'}}(\mathbf{v_k}-\mathbf{v_{k'}})^2 a_0^2(\mathbf{q})S(\mathbf{q}), \tag{21}$$

where

$$\mathbf{q}=\mathbf{k}_F-\mathbf{k}'_F, \quad J=\frac{l}{4\pi^3}\int_{\sigma_F} d\sigma_{\mathbf{k}}\mathbf{v_k} \approx \frac{ev_F\sigma_F}{12\pi^3}.$$

Here we have adopted the following notation: a_0 is the value of the static electron-ion scattering amplitude; e and m_* are the electron charge and effective mass; and Ω_0 is the equilibrium unit-cell volume of the lattice. The integration is carried out over the Fermi surface (σ_F), an element of which is denoted by $d\sigma_k$. In addition, $\mathbf{q}$ is the scattering vector, and $\mathbf{k}_F$ is the value of the electron momentum on the Fermi surface.

The factor characterizing scattering in a lattice with dynamic and static disorder $S(\mathbf{q})$ is defined as

$$S(\mathbf{q}) = \frac{1}{N}\sum_{nn'}\exp\{i\mathbf{q}(\mathbf{R}_n-\mathbf{R}_{n'})\} \times \exp\left(-\frac{W_n(\mathbf{q})+W_{n'}(\mathbf{q})}{2}\right),$$

where $\mathbf{R}_n=\mathbf{R}_n^{(0)}+\boldsymbol{\zeta}_n$ and the Debye–Waller factor

$$W_n(\mathbf{q}) = \langle(\mathbf{q}\cdot\mathbf{u}_n(0))^2\rangle.$$

The dynamic and static displacements are small compared with the interatomic distances. Therefore, we expand the exponential function in a series, retaining the first three terms in W_n and the two terms in $\boldsymbol{\zeta}_n$. The terms containing the delta function in the momentum do not lead to resistivity. Thus, the elastic scattering of the conduction electrons is described by the expression

$$S(\mathbf{q}) \approx \frac{1}{N}\sum_{nn'}\exp\{i\mathbf{q}(\mathbf{R}_n^{(0)}-\mathbf{R}_{n'}^{(0)})\} \times\left\{\frac{1}{2}W_n(\mathbf{q})W_{n'}(\mathbf{q})+(\mathbf{q}\boldsymbol{\zeta}_n)(\mathbf{q}\boldsymbol{\zeta}_{n'}) + \frac{i}{2}[(\mathbf{q}\boldsymbol{\zeta}_n)W_{n'}(\mathbf{q})-(\mathbf{q}\boldsymbol{\zeta}_{n'})W_n(\mathbf{q})]\right\}. \tag{22}$$

We note that, by definition, W_n and $\boldsymbol{\zeta}_n$ are, respectively, even and odd functions of $\mathbf{R}_n^{(0)}$, and the value of $S(\mathbf{q})$ is consequently real. The first term in (22) describes scattering due to the differences between the dynamic amplitudes for elastic scattering by a matrix atom and an isotopic impurity (see the details in Ref. 3), the second term describes scattering by near-impurity static displacements, and the third term describes interference scattering.

The contribution to the resistivity due to the scattering mechanism proposed in Ref. 2 is not taken into account. The corresponding expression contains an additional small parameter $(v_0(2k_F)/\epsilon_F)^2$, which does not appear in (22). Here $v_0(2k_F)$ is the Fourier component of the atomic pseudopotential in the region of large momentum transfers, ϵ_F is the Fermi energy, and k_F is the Fermi momentum. In the case of standard metals $(v_0(2k_F)/\epsilon_F)^2 \ll 1$ (see, for example, Ref. 9).

In an ideal lattice the Debye–Waller factor does not depend on the site index n. Therefore, the quantity

$$\Delta W_n = W_n(\Delta M \neq 0) - W_0(\Delta M = 0)$$

actually appears in (22). Its spatial fourier component has the following form (see Ref. 3):

$$\Delta W(\mathbf{q}) = \frac{1}{N}\sum_n \exp(i\mathbf{q}\mathbf{R}_n^{(0)})\Delta W_n = \frac{1}{N}\sum_n \exp(i\mathbf{q}\mathbf{R}_n^{(0)})q^\alpha q^\beta \Delta K_{nn}^{\alpha\beta} = -\frac{1}{2}q^\alpha q^\beta \frac{\Delta M}{M_0}\sum_{\mathbf{q}_1 j_1 j_2} Z^{\alpha\beta}_{\mathbf{q}_1 j_1, \mathbf{q}+\mathbf{q}_1 j_2}. \tag{23}$$

Therefore, we can ultimately represent $S(\mathbf{q})$ in the form

$$S(\mathbf{q}) \approx \frac{1}{N}\left\{\frac{1}{2}\Delta W(\mathbf{q})\Delta W(\mathbf{q}) + (\mathbf{q}\boldsymbol{\zeta}(\mathbf{q}))(\mathbf{q}\boldsymbol{\zeta}(\mathbf{q})) + [(\mathbf{q}\boldsymbol{\zeta}(\mathbf{q}))\Delta W(\mathbf{q}) - (\mathbf{q}\boldsymbol{\zeta}^*(\mathbf{q}))\Delta W(\mathbf{q})]\right\}. \tag{24}$$

We recall that $\zeta(\mathbf{q})$ is defined by (14)–(16).

If impurities are regarded as isolated atoms, the quantity $1/N$ can be replaced in (24) by the defect concentration c when the resistivity is determined. In the case of isotopes, instead of c we use the factor

$$\delta M^2 = \frac{\langle M^2\rangle - \langle M\rangle^2}{\langle M\rangle^2}, \quad \langle M\rangle = \sum_s\sum_n c_n^{(s)}M_n^{(s)}, \tag{25}$$

where $c_n^{(s)}$ is the probability of finding an isotope of type s at lattice point n.

Then, using (23) we can show that in the model of a linear chain

$$\Delta W(q) \approx -\frac{1}{2}\epsilon q^2 l^2. \tag{26}$$

Simultaneously, according to (20), we have

$$q\zeta(q) \approx -\frac{1}{2}\lambda\epsilon\frac{\sqrt{f_2/M_0}}{\omega_q}ql, \quad \lambda = \frac{g_3 l}{2f_2} > 1. \tag{27}$$

We plug (26) and (27) into (24). We plug the resultant expression, in turn, into the expression (21) for ρ_r. Let us compare the values of the standard and nonstandard residual resistivities. For this purpose we consider the factor

$$\Upsilon = \frac{\int_0^{2k_F} q^3 a_0^2(q) NS(q)}{\epsilon^2 \int_0^{2k_F} dq\, q^3 a_0^2(q)}.$$

For the sake of fixing ideas and simplicity, for the amplitudes we assume that

$$\frac{a_0(q/2k_F)}{a_0(0)} = \frac{0.2}{(q/2k_F)^2 + 0.2}.$$

Then, with consideration of the relations (26) and (27) and under the assumption that the Debye wave vector is close to $2k_F$, we obtain

$$\Upsilon = 0.15 Z_0^{2/3}\pi^2\{Z_0^{2/3}\pi^2 + 2.66 Z_0^{1/3}\lambda\pi + 2.15\lambda^2\}, \tag{28}$$

where Z_0 is the number of valence electrons per atom.

It follows from the results of the numerical calculations, i.e., from (28), that in the model adopted the disorder due to the static displacements in (24) influences the magnitude of the residual resistivity to a greater degree than do the differences in the amplitudes due to the Debye–Waller factors alone. Consideration of the scattering by static displacements can alter the value of Υ by several fold when $Z_0 \geqslant 3$ and by an order of magnitude when $Z_0 = 1$.

Calculations were also performed for metallic lithium, in which there are two isotopes: ^{6}Li and ^{7}Li. In this case our treatment was based on the results of the Brovman–Kagan microscopic theory of nontransition metals (see the review in Ref. 10). The kinetic coefficients of regular metals were previously analyzed within this theory in Ref. 11. The corresponding formulas needed to determine the quantities appearing in the theory of the residual resistivity developed here are presented in Appendix B. We note that the two-wave approximation for the electronic wave function was used in the calculations of the residual resistivity. It was found with neglect of the lattice distortion that $\Upsilon \approx 21$. When the static displacements are taken into account, $\Upsilon \approx 123$. Thus, the estimates of Υ obtained above are confirmed.

We note that the dominant contribution to the residual resistivity in weak metal solutions with standard substitutional impurities is made by the scattering of electrons by defects due to the differences in the static amplitudes. The contribution to ρ_r caused by scattering on matrix atoms displaced near a defect is comparatively small. Isoelectronic weak solutions, in which ρ_r can vary by several tens of percent when the lattice distortion is taken into account, are exceptions (see, for example, Ref. 12). As was shown above, the scattering mechanism associated with static displacements is dominant only in the case of metals with isotopic impurities.

Let us consider the question of the mean free paths. As we know, the mean free path of conduction electrons in nontransition metals at room temperature is determined by the scattering of electrons by phonons. In this case (see, for example, Ref. 12)

$$\Lambda_{\rm ph} \sim 50\frac{T_m}{T}l,$$

where T_m is the melting point. The corresponding value of $\Lambda_{\rm ph}$ is of the order of several hundred angstroms. Taking into account the foregoing, we can show within the model used in Refs. 8 and 13 (which permits allowance for umklapp processes on a qualitative level) that the effective mean free path $\Lambda_{\rm is}$ associated with scattering due to differences in the dynamic amplitudes and static displacements in a metal with different isotopes at absolute zero temperature is equal in order of magnitude to

$$\Lambda_{\rm is} \sim \frac{1}{\delta M^2}\frac{l^2}{\langle u^2\rangle}\frac{1}{\Upsilon}\left[\frac{v_0(2k_F)}{v_0(q=0)}\right]^2\Lambda_{\rm ph}. \tag{29}$$

Here $\langle u^2\rangle/l^2 \sim 10^{-3}$. In the case of natural isotopic mixtures of such metals as Li, Zn, and Sn, the parameter defined by formula (25) $\delta M^2 \geqslant 10^{-4}$. For molybdenum $\delta M^2 \approx 6\times 10^{-4}$. Consequently, for such metals $\Lambda_{\rm is} \leqslant 1$ mm, while with neglect of the lattice distortion $\Lambda_{\rm is} \geqslant 1$ cm. If samples containing, for example, isotopes of two kinds in equal percentages are synthesized, the mean free paths $\Lambda_{\rm is}$ (29) can decrease by more than an order of magnitude due to the increase in the isotopic disorder factor δM^2. Nevertheless, the dimensions of the samples used in the experiment can be standard even for natural compositions.

We use $\Lambda_{\rm im}$ to denote the mean free path due to the elastic scattering of electrons by nonisotopic impurities with a concentration $c_{\rm im}$. Observation of effects caused by isotopic disorder is possible, if $\Lambda_{\rm im} \geqslant \Lambda_{\rm is}$. In this case the following condition for the concentration of point defects must be satisfied:

$$c_{\rm im} \leqslant 4\eta\delta M^2\Upsilon\left(\frac{\langle u^2\rangle}{l^2}\right)^2, \quad \eta = \frac{\langle\langle (q/2k_F)^2 a_0^2\rangle\rangle}{\langle\langle (a_{\rm im}-a_0)^2\rangle\rangle}, \tag{30}$$

where $a_{\rm im}$ is the scattering amplitude on impurities and

$$\langle\langle a^2\rangle\rangle = \int_{\sigma_F}\int_{\sigma_F}\frac{d\sigma_k}{v_k}\frac{d\sigma_k'}{v_k'}q^2 a^2(q).$$

For mixed-valence mixtures $\eta \leqslant 1$. If the condition (19) holds, then, according to (30), we have

$$c_{\rm im} \sim 10^{-3}\delta M^2,$$

i.e., the samples must be very perfect.

4. CONCLUSION

Chemically pure metals whose lattice sites are occupied by different isotopes have been considered. Equations describing the fields of static displacements appearing near isotopes due to dynamic disorder have been obtained in a microscopic approach in the first stage. The question of the residual resistivity has been analyzed within the Born approximation. The contributions associated with differences in the dynamic elastic electron-ion interaction amplitudes and with fields of near-impurity static displacements are taken into account simultaneously. It has been shown that specific static displacements have a significant influence on the residual resistivity ρ_r. This contribution to ρ_r can be far greater than the contribution due to differences in the dynamic electron scattering amplitudes.

As far as we know, the question of the residual resistivity of metals containing different isotopes was investigated experimentally by Sharvin in the case of natural tin.[14] For a concrete theoretical treatment it would be desirable to have data for samples with different isotopic compositions.

We are indebted to S. M. Stishov for a valuable comment and A. V. Inyushkin for a useful communication. We thank D. A. Zhernov for his assistance.

This work was performed with financial support from N. A. Chernoplekov.

APPENDIX A:

Let us consider the question of the asymptotic representation of the field of static displacements $\{\boldsymbol{\zeta}_n\}$. First, we must determine the value of $\zeta(q\to 0)$ under the condition $|R_n|\gg l$.

We take into account the recurrence relation for force parameters of different orders:

$$\sum_{n_1} R_{n_1}^{\alpha_1}\Phi_{p+1,n_1\ldots n_{p+1}}^{\alpha_1\ldots\alpha_{p+1}}=\Omega_0\frac{\partial}{\partial\Omega_0}\Phi_{p,n_2\ldots n_{p+1}}^{\alpha_2\ldots\alpha_{p+1}}\delta_{\alpha\alpha_1}, \tag{A1}$$

where $p=2,3$. This relation was derived in Ref. 15 for nontransition metals on the basis of the microscopic theory in Ref. 10.

On the basis of (14) and (A1) it can be shown that

$$\zeta^\alpha(q\to 0)\approx -i\frac{\Delta M}{M_0}\sum_j \overline{D}_{\mathbf{qj}}^{\alpha\beta}q^\beta \times\sum_{\mathbf{q}_1 j_1\mathbf{q}_2 j_2}\sum_{\mathbf{m}}\left\{\Omega_0\frac{\partial}{\partial\Omega_0}\Phi_{2,0\mathbf{m}}\right\} \times\exp(i\mathbf{q}_1R_m)Z_{\mathbf{q}_1 j_1\mathbf{q}_2 j_2}^{\beta\gamma}\Delta(\mathbf{q}_1+\mathbf{q}_2), \tag{A2}$$

where to simplify the form of the expressions we set

$$\overline{D}_{\mathbf{qj}}^{\alpha,\beta}=\frac{e^\alpha(\mathbf{q},j)e^\beta(\mathbf{q},j)}{M_0\omega^2(\mathbf{q},j)}.$$

Using the standard relation

$$\sum_{\alpha,\beta}e^\alpha(\mathbf{q},j)\Phi_{2,\mathbf{q}}^{\alpha,\beta}e^\beta(\mathbf{q},j)=M_0\omega^2(\mathbf{q},j)$$

and defining the partial Grüneisen factor in the form

$$\gamma_G(\mathbf{q},j)=-\frac{\Omega_0}{\omega(\mathbf{q},j)}\frac{\partial}{\partial\Omega_0}\omega(\mathbf{q},j),$$

instead of (A2) we obtain

$$\zeta^\alpha(q\to 0)\approx -i\frac{\Delta M}{M_0}\gamma_G E_{\rm vib}\sum_j \overline{D}_{\mathbf{qj}}^{\alpha,\beta}\frac{q^\alpha}{M_0}. \tag{A3}$$

Here

$$\gamma_G=\frac{\Sigma_{\mathbf{qj}}\gamma(\mathbf{q},j)\omega(\mathbf{q},j)}{\Sigma_{\mathbf{qj}}\omega(\mathbf{q},j)},\quad E_{\rm vib}=\frac{1}{2}\sum_{\mathbf{qj}}\omega(\mathbf{q},j),$$

where γ_G is the integral Grüneisen factor and $E_{\rm vib}$ is the energy of the zero-point vibrations.

The factor

$$\sum_{\mathbf{m}}\exp(i\mathbf{q}\cdot\mathbf{R}_{\mathbf{m}}^{(0)})\overline{D}_{\mathbf{q}}$$

was determined in the asymptotic limit for cubic crystals with bcc and fcc lattices in the general case in Ref. 16. Let us consider here the case of an isotopic elastic continuum. Then

$$\omega(q)=\sqrt{\frac{\lambda+\mu}{\rho}}q,\quad e^\alpha(\mathbf{q})=\frac{\mathbf{q}\cdot\mathbf{v}}{q^2}q^\alpha, \tag{A4}$$

where λ and μ are the Lamé constants and $\mathbf{v}$ is a unit vector. In addition,

$$M_0(\lambda+2\mu)/\rho=\Omega_0C_{11},$$

where ρ is the density and C_{11} is the elastic modulus.

Substituting (A4) into (A3), after some simple transformations, in the asymptotic limit we have

$$\boldsymbol{\zeta}_n=\frac{\Omega_0}{(2\pi)^3}\int d\mathbf{q}\,\exp(i\mathbf{q}\cdot\mathbf{R}_n^{(0)})\boldsymbol{\zeta}(\mathbf{q})\approx\frac{b}{4\pi}\frac{\mathbf{R}_n^{(0)}\Omega_0}{|\mathbf{R}_n^{(0)}|^3},$$

where b is the magnitude of the defect:

$$b=-\frac{\Delta M}{M_0}\frac{\gamma_G E_{\rm vib}}{\Omega_0C_{11}}.$$

APPENDIX B:

As has been noted, the calculations for lithium were performed within the Brovman–Kagan microscopic many-electron theory. In addition, the local pseudopotential of a ''naked'' ion was chosen in the form

$$v_0(q)=-\frac{4\pi}{q^2}\left[Z_0e^2\cos(qr_0)+v_0r_0\times\left(\frac{\sin(qr_0)}{qr_0}-\cos(qr_0)\right)\right]\exp(-\chi q^2),$$

where $\chi=0.03(2k_F)^{-2}$. The effective interelectronic interaction was considered in the Geldart–Taylor approximation.[17] The parameters $v_0(q)$ for Li were previously found: $v_0=-0.262$ and $r_0=0.825$ Å.[18]

The frequencies and polarization vectors of the phonon modes of the harmonic lithium lattice were determined using

a dynamic matrix in the standard form (see, for example, Ref. 9):

$$D_{\alpha\beta}(\mathbf{q})=\frac{4\pi Z_0^2e^2}{M_0\Omega_0}\sum_{\mathbf{B}}\left[\Psi_{2,\alpha\beta}(\mathbf{q}+\mathbf{B})-\sum_{\mathbf{B}\neq 0}\Psi_{2,\alpha\beta}(\mathbf{B})\right]$$
$$+Z_0^2e^2\sum_{n\neq 0}\left[\exp(\mathbf{q}\mathbf{R}_n^{(0)})-1\right]$$
$$\times\left[\frac{R_n^{(0),\alpha}R_n^{(0),\beta}}{R_n^{(0)2}}F_1(\mathbf{R}_n^{(0)})-\delta_{\alpha\beta}F_2(\mathbf{R}_n^{(0)})\right].$$

Here **B** is the reciprocal lattice vector. The factor $\Psi(q)$ characterizes the contributions of the Coulomb and indirect (by means of conduction electrons) ion-ion pair interactions. It has the form

$$\Psi_{2,\alpha\beta}(q)=q_\alpha q_\beta\Psi(q),$$
$$\Psi(q)=\frac{4\pi Z_0^2e^2}{\Omega_0 q^2}\exp\left(-\frac{q^2}{4\eta}\right)-\Omega_0 v_0^2(q)\frac{\Pi(q)}{\epsilon(q)},$$

where $\epsilon(q)=1+4\pi e^2\Pi(q)/q^2$ is the dielectric function with the polarization operator Π and η is the Ewald parameter. The expressions for $F_1(R)$ and $F_2(R)$, which are related to the Coulomb interaction can be written in the following manner:

$$F_1(R)=\frac{3\,\mathrm{erf}(\sqrt{\eta}R)}{R^3}+2\sqrt{\frac{\eta}{\pi}}\exp(-\eta R^2)\left(\frac{3}{R^2}+2\eta\right),$$

$$F_2(R)=\frac{\mathrm{erf}(\sqrt{\eta}R)}{R^3}+2\sqrt{\frac{\eta}{\pi}}\frac{\exp(-\eta R^2)}{R^2},$$

$$\mathrm{erf}(x)=1-\frac{2}{\sqrt{\pi}}\int_0^x dz\exp(-z^2).$$

The expression for the Fourier component of the third-order force parameter had the form

$$\Phi_{3,\mathbf{k},\mathbf{q},\mathbf{k}+\mathbf{q}}^{\alpha\beta\gamma}=\sum_{\mathbf{q}}\left[\Psi_{3,\alpha\beta\gamma}(\mathbf{k}+\mathbf{B})\right.$$
$$\left.+\Psi_{3,\alpha\beta\gamma}(\mathbf{q}+\mathbf{B})-\Psi_{3,\alpha\beta\gamma}(\mathbf{k}+\mathbf{q}+\mathbf{B})\right]$$
$$+Z_0^2e^2\sum_{n\neq 0}\left\{\left[R_\alpha R_\beta R_\gamma\overline{F}_3(R)+(R_\alpha\delta_{\beta\gamma}\right.\right.$$
$$\left.+R_\beta\delta_{\alpha\gamma}+R_\gamma\delta_{\alpha\beta})\overline{F}_2(R)\right]\left[\sin(\mathbf{k}\cdot\mathbf{R}+\mathbf{q}\cdot\mathbf{R})\right.$$
$$\left.\left.-\sin(\mathbf{k}\cdot\mathbf{R})-\sin(\mathbf{q}\cdot\mathbf{k})\right]\right\}_{R=R_n^{(0)}}.$$

To simplify the form of the expression, here we set

$$\Psi_{3,\alpha\beta\gamma}(\mathbf{q})=q_\alpha q_\beta q_\gamma\Psi(\mathbf{q}),$$
$$\overline{F}_n(R)=\left(\frac{1}{R}\frac{d}{dR}\right)^n\overline{F}(R),$$
$$\overline{F}(R)=\frac{2}{R}\sqrt{\frac{\eta}{\pi}}\int_R^\infty\exp(-\eta z^2)\,dz.$$

We note that explicit expressions for the force parameters, which permit allowance for the many-particle ion-ion interaction, were obtained in Ref. 15.

*)E-mail: zhernov@kurm.polyn.kiae.su

[1] M. A. Asen-Palmer, N. Bartcovsky, E. Gmelin, M. Cardona, A. P. Zhernov, A. V. Inushkin, A. N. Toldenkov, V. I. Ozhogin, K. M. Itoh, and E. E. Haller, Phys. Rev. B **56**, 9431 (1997).
[2] I. Ya. Pomeranchuk, Zh. Éksp. Teor. Fiz. **35**, 992 (1958) [Sov. Phys. JETP **8**, 693 (1959)].
[3] Yu. M. Kagan and A. P. Zhernov, Zh. Éksp. Teor. Fiz. **53**, 1744 (1967) [Sov. Phys. JETP **26**, 999 (1968)].
[4] G. Leibfried and W. Ludwig, ''Theory of anharmonic effects in crystals,'' in *Solid State Physics*, *Vol. 12*, F. Seitz and D. Turnbull (eds.), Academic Press, New York–London (1961), pp. 276–444 [Russ. transl., IL, Moscow (1963)].
[5] G. Leibfried and N. Breuer, *Point Defects in Metals*, *Vol. 1: Introduction to the Theory*, Springer-Verlag, Berlin–New York (1978) [Russ. transl., Mir, Moscow (1981)].
[6] N. W. Ashcroft and N. D. Mermin, *Solid State Physics*, Saunders, Philadelphia (1976) [Russ. transl., Mir, Moscow (1979)].
[7] P. G. Klemens and A. A. Maradudin, Phys. Rev. **123**, 804 (1961).
[8] J. M. Ziman, *Electrons and Phonons*, Oxford University Press (1960) [Russ. transl., IL, Moscow (1962)].
[9] V. Heine, M. L. Cohen, and D. Weaire, ''Pseudopotential theory of cohesion and structure,'' in *Solid State Physics*, *Vol. 24*, H. Ehrenreich, F. Seitz, and D. Turnbull (eds.), Academic Press, New York (1970), pp. 250–463 [Russ. transl., Mir, Moscow (1973)].
[10] E. G. Brovman and Yu. M. Kagan, Usp. Fiz. Nauk **112**, 369 (1974) [Sov. Phys. Solid State **17**, 125 (1974)].
[11] A. P. Zhernov and Yu. M. Kagan, Fiz. Tverd. Tela (Leningrad) **20**, 3306 (1978) [Sov. Phys. Solid State **20**, 1906 (1978)].
[12] Z. Popovic, J. P. Carbotte, and G. R. Piercy, J. Phys. F **3**, 1008 (1973).
[13] Yu. M. Kagan and A. P. Zhernov, Zh. Éksp. Teor. Fiz. **50**, 1109 (1966) [Sov. Phys. JETP **23**, 737 (1966)].
[14] Yu. V. Sharvin, V. E. Gantmakher, Zh. Éksp. Teor. Fiz. **39**, 1242 (1960) [Sov. Phys. JETP **12**, 866 (1961)].
[15] A. P. Zhernov and T. A. Mamedov, Phys. Status Solidi B **103**, 477 (1981).
[16] M. A. Krivoglaz, *Theory of X-Ray and Thermal Neutron Scattering by Real Crystals*, Plenum Press, New York (1969).
[17] D. J. W. Geldart and R. Taylor, J. Phys. **48**, 155 (1970).
[18] A. P. Zhernov and G. Solt, Fiz. Tverd. Tela (Leningrad) **19**, 3400 (1977) [Sov. Phys. Solid State **19**, 1985 (1977)].

Translated and edited by P. Shelnitz

Diffusion in a system of impurity centers with dipole-dipole hopping rates

F. S. Dzheparov, D. V. L'vov,*) and V. E. Shestopal

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia
(Submitted 30 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2166–2181 (December 1998)

Random walks in disordered media with dipole-dipole transition rates are considered. The long-time asymptotics of the process are investigated on the basis of a new numerical simulation method, which includes periodic continuation of the system without periodic continuation of the initial condition. It is shown that the long-time asymptotics have a diffusive character. The concentration dependence of the diffusion coefficient for simple cubic and face-centered cubic lattices is studied. © *1998 American Institute of Physics.* [S1063-7761(98)02012-5]

1. INTRODUCTION

Random walks in disordered media with dipole-dipole transition rates describe diverse migration processes of localized excitons, which can be investigated using fluorescence depolarization measurements,[1,2] time-resolved fluorescence line-narrowing spectroscopy,[3,4] and the results of four-wave mixing experiments,[5] as well as the transport of spin polarization.[6]

Various theoretical[4,7–18] and numerical[8,19,20] approaches have been employed to calculate the principal processes observed in these and similar systems. In particular, a method for investigating the autocorrelator at moderately large times was developed in Ref. 11. However, there has been no agreement in the literature in resolving the extremely important question of the form and parameters of the long-time asymptotics of the process. In particular, the values of the diffusion coefficient that have been proposed hitherto differ appreciably. The diffusive character of the asymptotics, which is not questioned by most investigators, has not yet been proved analytically.

This paper considers dipolar transport in a three-dimensional disordered system formed by impurity centers randomly dispersed in a crystal. Our investigation is based on a new method for analyzing the asymptotics of random walks in disordered media with a translationally invariant distribution of the randomness of the medium.[21] The procedure for substantiating the results includes comprehensive testing and repeated computer calculations for different values of the parameters, which provide convincing confirmation of the significant qualitative properties of the asymptotics, unequivocally (to within $<3\%$) revealing at least the two leading terms in the power-series expansion of the autocorrelator with respect to time. In addition, the method is successfully corroborated when it is applied to an extensive class of models of random walks in disordered media with exact asymptotic solutions.[22–24]

1.1. The dipole-dipole transport of excitations among centers of an disordered impurity can be described by the kinetic equation

$$\dot{p}_{im}=-\sum_j\ (\nu_{ji}p_{im}-\nu_{ij}p_{jm}),\quad p_{im}(t=0)=\delta_{im}, \tag{1}$$

where i, j, and m run through the numbers of the sites occupied by the impurity in the regular lattice L, and $p_{jm}(t)$ is the probability of finding a random walker at site j at the time t if it was initially at site m.

In the standard model of dipole-dipole transport

$$\nu_{i\neq j}=\frac{\nu_0 r_0^6}{|\mathbf{x}_i-\mathbf{x}_j|^6},\quad \nu_{ii}=0, \tag{2}$$

where $\mathbf{x}_j$ is the radius vector of site j and ν_0 is the hopping rate over the minimum distance r_0 between sites. The configuration of the impurity centers is uncorrelated, and the probability of finding an impurity in an arbitrary lattice site is $c\leqslant 1$. In the limit of low concentrations $c\ll 1$, where the disorder of the medium is manifested most strongly, the characteristic time scale is assigned by the Förster constant $\beta=(16/9)\pi^3(r_0^3/\Omega)^2c^2\nu_0$, where Ω is the unit-cell volume. Parametrically β coincides with the transport rate over the mean distance $\overline{r}=(\Omega/c)^{1/3}$ and is determined by averaging the decay process over all impurity configurations:

$$\langle\exp(-t\tilde{D}_i)\rangle=\exp[-\sqrt{\beta t}\,],\quad \tilde{D}_i=\sum_j\ \nu_{ji},$$

where $c\to 0$ and βt is a finite quantity.

A general method for investigating the behavior of one of the most important, experimentally observed characteristics of random walks in an disordered system, viz., the propagator

$$\mathscr{P}_{\mathbf{xy}}(t)=c^{-1}\sum_{ij}\ \langle p_{ij}(t)\,\delta_{\mathbf{x}\mathbf{x}_i}\delta_{\mathbf{y}\mathbf{x}_j}\rangle, \tag{3}$$

which is the probability, averaged over all the impurity configurations, of finding an excitation on lattice site $\mathbf{x}$ at the time t if it was initially on site $\mathbf{y}$, was developed in Ref. 11. It involves expanding the propagator in powers of $ct^{1/2}$. The cumulant version of this expansion with consideration of the first three members of the series provides a faithful description of the existing experimental data[1,3,4,6] for $\beta t<5$.

The principal unsolved problem is the form and parameters of $\mathscr{P}_{\mathbf{xy}}(t\to\infty)$.

1.2. In order to describe some difficulties in the theoretical analysis of the system (1) more explicitly and to more clearly see the place of the problem under discussion among other problems in modern physics, we present two expressions for the propagator in the form of a functional integral,[12] each of which is based on the most commonly used approaches to such representations. For this purpose it is convenient to introduce an occupation-number representation and to rewrite Eq. (1) as

$$\dot{P}_{\mathbf{xy}} = -\sum_{\mathbf{z}} (n_{\mathbf{z}}\nu_{\mathbf{zx}}n_{\mathbf{x}}P_{\mathbf{xy}} - n_{\mathbf{x}}\nu_{\mathbf{xz}}n_{\mathbf{z}}P_{\mathbf{zy}}),$$

$$P_{\mathbf{xy}}(t=0) = \frac{n_{\mathbf{y}}}{c}\delta_{\mathbf{xy}}. \tag{4}$$

Here $P_{\mathbf{xy}}(t)$ is the probability of finding an excitation on lattice site $\mathbf{x}$ if it was initially on site $\mathbf{y}$, and $n_{\mathbf{x}}$ is the donor occupation number of site $\mathbf{x}$, which takes values of 0 and 1. In this case $\langle n_{\mathbf{x}}\rangle = c$, and $\mathscr{P}_{\mathbf{xy}} = \langle P_{\mathbf{xy}}\rangle$. The equivalence of Eqs. (1) and (4) becomes clear if all the zero components corresponding to empty sites are eliminated in (4). After allowance for the fact that $n_{\mathbf{x}}P_{\mathbf{xy}} = P_{\mathbf{xy}}$, Eq. (4) takes the form

$$\dot{P}_{\mathbf{xy}} = -\sum_{\mathbf{z}} (n_{\mathbf{z}}\nu_{\mathbf{zx}}P_{\mathbf{xy}} - n_{\mathbf{x}}\nu_{\mathbf{xz}}P_{\mathbf{zy}}). \tag{5}$$

When the long-time asymptotics are analyzed, the initial condition imposed on this equation can be taken in the form $P_{\mathbf{xy}}(t=0) = \delta_{\mathbf{xy}}$. In fact, the excitation is initially placed in an arbitrary lattice site, and then it hops to the nearest donors and subsequently migrates among them. The long-time asymptotics of such a process clearly coincide with the exact asymptotics.[12,14] Using $P^1_{\mathbf{xy}}$ to denote the corresponding solution, we obtain

$$P^1 = \exp(-At), \quad A = \sum_{\mathbf{z}} n_{\mathbf{z}}A^{\mathbf{z}},$$

$$A^{\mathbf{z}}_{\mathbf{xq}} = \delta_{\mathbf{xq}}\nu_{\mathbf{zq}} - \delta_{\mathbf{zx}}\nu_{\mathbf{xq}}.$$

A ''quantum-mechanical'' representation is obtained on the basis of natural continuation of the kinetic equation from a lattice to a continuous medium and has the form

$$\mathscr{P}^1_{\mathbf{xy}}(t) = \langle P^1_{\mathbf{xy}}(t)\rangle = \int_{\mathbf{q}(0)=\mathbf{x}}^{\mathbf{q}(1)=\mathbf{y}} D\mathbf{p}(\tau)D\mathbf{q}(\tau) \times \exp\left[i\int_{\mathbf{x}}^{\mathbf{y}} \mathbf{p}\, d\mathbf{q} + c\int d^3z \times \left(\exp\left\{ -t\int_0^1 d\tau A^{\mathbf{z}}(\mathbf{q}(\tau),\mathbf{p}(\tau)) \right\} - 1 \right) \right], \tag{6}$$

where $c \ll 1$, the functional integral is construed as the limit of the finite-multiple integrals corresponding to the subdivisions of the time interval for unconstrained reduction of the length of these subdivisions, and the function $A^{\mathbf{z}}(\mathbf{q},\mathbf{p})$ is determined from a relation of the type

$$\int \frac{d\mathbf{p}}{(2\pi)^3} A^{\mathbf{z}}(\mathbf{q},\mathbf{p}) e^{i\mathbf{p}(\mathbf{x}-\mathbf{q})} = A^{\mathbf{z}}_{\mathbf{xq}} = \nu_{\mathbf{zq}}[\delta(\mathbf{x}-\mathbf{q}) - \delta(\mathbf{x}-\mathbf{z})]. \tag{7}$$

One of the solutions of Eq. (7) is $A^{\mathbf{z}}(\mathbf{q},\mathbf{p}) = \nu_{\mathbf{zq}}[1 - \exp\{-i\mathbf{p}(\mathbf{z}-\mathbf{q})\}]$. A discussion of the other solutions and a generalization of (6) and (7) can be found in Ref. 12.

The following representation of the Laplace transform of the propagator employs integration over two families of pairs of complex-conjugate variables, viz., the ''ordinary'' variables $a^+_{\mathbf{z}}$ and $a_{\mathbf{z}}$ and the Grassmann variables $\alpha^+_{\mathbf{z}}$ and $\alpha_{\mathbf{z}}$:

$$P^1_{\mathbf{xy}}(\lambda) = \left(\frac{1}{\lambda + A}\right)_{\mathbf{xy}} = \int_0^\infty dt\, \mathscr{P}^1_{\mathbf{xy}}(t) e^{-\lambda t} = \int \prod_{\mathbf{z}\in L} \frac{da^+_{\mathbf{z}}\, da_{\mathbf{z}}\, d\alpha_{\mathbf{z}}\, d\alpha^+_{\mathbf{z}}}{2\pi} a_{\mathbf{x}}a^+_{\mathbf{y}} \times \exp\left[-\lambda(a^+a + \alpha^+\alpha) + c\sum_{\mathbf{z}\in L} (\exp\{-(a^+A^{\mathbf{z}}a + \alpha^+A^{\mathbf{z}}\alpha)\} - 1) \right]. \tag{8}$$

Here $c \ll 1$, the path integral is construed as the limit of the finite-multiple integrals corresponding to the transition from a lattice to a finite discrete torus, and the expressions of the type a^+Ba are construed as $\Sigma_{xq}a^+_x B_{xq}a_q$.

We note that a rigorous proof of the convergence of the integrals in (6) and (8) was not found, but after formally expanding the integrands into power series in the concentration c and calculating the coefficients term-by-term, we obtain expressions which are correct in the continuous-medium limit.

As can be seen from the representation (6), our problem is more complicated than the polaron problem,[25] and the nonlocality and singularity of the action are far more pronounced. The representation (8) transforms our problem into the problem of a nonlinear nonlocal unrenormalized field theory with a strong interaction, since the actual nonlinearity parameter is $c\lambda^{-1/2}$ (Refs. 11 and 12) and large values of t correspond to $\lambda \propto 1/t \to 0$.

An investigation of the propagator asymptotics using similar representations in other random-walk problems in disordered media can be fully successful,[23] but such a technique has hitherto not yielded any appreciable results for the system (1) considered in this work. Therefore, it would be natural to expect that finding the long-time asymptotics of the averaged propagator will stimulate the development of a general field theory and functional integration and that new methods could be tested on the problem under consideration along with the Ising model and the polaron problem, which have traditionally been used for this purpose.

1.3. Numerical simulation of the asymptotic stage of the process is difficult for two reasons. The previous attempts at such simulation[8,19,20] were not based on a sufficiently detailed theoretical prediction of the outcome in the asymptotic stage.[10,4] First of all, the main features of the onset of the diffusion stage in such systems were disregarded in the simulation. In particular, it was not taken into account for a long time that diffusion is preceded by a slower process [see (23) below]. The second reason is that the replacement of an infinite random medium by a periodic medium with a periodically continued initial condition usually employed in such a

simulation[8,19,20] leads to an exponential type of approach of the autocorrelator $\mathscr{P}_{00}(t)$ to the stationary value corresponding to the uniformly filled state. When random walks in disordered media are simulated by such a method, to discover the asymptotic power-law $\mathscr{P}_{00}\sim 1/(\beta t)^{3/2}$ (which is characteristic of diffusion) it is necessary to ''strike a balance'' between long and ''excessively'' long times after preliminarily finding the spectral expansion of the large linear system to a very high accuracy. The sensitivity of the method toward errors is even higher for higher orders in the asymptotic expansion.

2. METHOD FOR INVESTIGATING RANDOM WALKS

Our investigation is based on a new method,[21] which calls for the treatment of a spatially infinite system. In the case of a simple cubic lattice the system is obtained from a pseudorandom configuration of an impurity in a cube with a center at the origin of coordinates and the edge Rr_0, which has cR^3 impurity sites, by periodically continuing it in all space with a period equal to the edge. Similarly, in the case of a face-centered cubic (fcc) lattice, a pseudorandom configuration is constructed in a parallelepiped with the edges $R\mathbf{a}_1$, $R\mathbf{a}_2$, and $R\mathbf{a}_3$, where $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ are the Bravais vectors of the fcc lattice.

We introduce the quantity

$$\rho_j(\mathbf{k},t)=\sum_m p_{jm}(t)\exp\{i(\mathbf{x}_m-\mathbf{x}_j)\mathbf{k}\}, \tag{9}$$

where the vector $\mathbf{k}$ runs through the Brillouin zone corresponding to the lattice L. The configurational averages

$$\mathscr{P}(\mathbf{k},t)=\langle\rho_i(\mathbf{k},t)\rangle=\sum_y \exp\{-i\mathbf{k}(\mathbf{x}-\mathbf{y})\}\mathscr{P}_{\mathbf{xy}}(t)$$

give the Fourier transform of the propagator and are observable in four-wave mixing experiments.[5] It follows from (1) that

$$\dot{\rho}_i=-\sum_j [\nu_{ji}\rho_i-\nu_{ij}\exp\{i(\mathbf{x}_j-\mathbf{x}_i)\mathbf{k}\}\rho_j],$$

$$\rho_i(t=0)=1. \tag{10}$$

When the impurity has a periodic distribution, $\rho_i=\rho_m$, if $\mathbf{x}_i-\mathbf{x}_m=R\mathbf{M}$, where $\mathbf{M}\in L$. In fact, according to (10), ρ_i and ρ_m satisfy identical equations with identical initial conditions and therefore coincide. Consequently, Eq. (10) can be represented as

$$\dot{\rho}_i=-(A(\mathbf{k})\rho)_i=-\sum_{j=1}^{N} [W_{ji}(0)\rho_i-W_{ij}(\mathbf{k})\rho_j],$$

$$\rho_i(t=0)=1. \tag{11}$$

Here i and j run through the numbers of the sites in the periodicity volume, and

$$W_{ij}(\mathbf{k})=\sum_{\mathbf{M}\in L}{}' \nu_0 r_0^6|\mathbf{x}_j-\mathbf{x}_i-\mathbf{M}R|^{-6}\times\exp\{i(\mathbf{x}_j-\mathbf{x}_i-\mathbf{M}R)\mathbf{k}\}, \tag{12}$$

where Σ' denotes summation over all $\mathbf{M}$ for which $|\mathbf{x}_j-\mathbf{x}_i-\mathbf{M}R|\neq 0$.

We note that the idea of treating a periodic random medium was proposed in Ref. 16 to solve a one-dimensional model with transfer to the nearest neighbor and was used in Ref. 17 to study diffusion in the case of weak disorder in a space of arbitrary dimensionality. In those studies quantities of the type

$$Q_{\mathbf{x}}(\mathbf{y})=\sum_{\mathbf{M}} P_{\mathbf{x}+R\mathbf{M},\mathbf{y}}, \quad \mathbf{S}_{\mathbf{x}}(\mathbf{y})=\sum_{\mathbf{M}} (\mathbf{x}+\mathbf{M}R)P_{\mathbf{x}+\mathbf{M}R,\mathbf{y}},$$

rather than ρ_i, were analyzed. Closed equations, which could be used to determined the diffusion tensor, were obtained for them.

To replace the propagator $\mathscr{P}(\mathbf{k},t)$ it is convenient to choose

$$P(\mathbf{k},t|N)=\sum_i \rho_i(\mathbf{k},t)\Big/\sum_i 1. \tag{13}$$

The quantity $P(\mathbf{k},t|N)$ can be regarded as 1) the Fourier transform of the propagator averaged over all the configurations obtained from a given configuration by displacing the initially excited site within the periodicity volume and 2) the projection of ρ_i onto the leading eigenvector of the operator $A(\mathbf{0})$. In order to elaborate on the first claim, we note that the generator (1) is symmetric ($p_{im}=p_{mi}$) and the subscript i in ρ_i can therefore be regarded as the label of the site in which the excitation walking among the impurity centers was initially located. We note that in the case of asymmetric transfer that is translationally invariant on the average (where $\nu_{ij}\neq\nu_{ji}$) we can investigate the equation which is conjugate to (1) and determine $\tilde{\rho}_i$ using the sum over the left-hand index, rather than over the right-hand index, as in (9). In this case the arithmetic mean of the components $\tilde{\rho}_i$ will be the exact mean over all configurations differing from one another with respect to the initial site.

2.1. It is convenient to investigate the matrix elements of the operator $A(\mathbf{k})$ (11) using the Poisson–Ewald method.[23] It gives a complete representation of the analytic properties of the generator at small values of $|\mathbf{k}|$ and permits rapid calculation of the elements to a high accuracy.

We bring (12) into the form

$$\sum_{\mathbf{M}\in L}{}' |\mathbf{x}-\mathbf{M}R|^{-6}\exp\{i(\mathbf{x}-\mathbf{M}R)\mathbf{k}\}$$

$$=\frac{1}{2}\sum_{\mathbf{M}\in L}{}' \int_0^\infty d\tau\,\tau^2\exp\{-\tau(\mathbf{x}-\mathbf{M}R)^2+i(\mathbf{x}-\mathbf{M}R)\mathbf{k}\}$$

$$=\frac{1}{2}\left(\int_0^{\tau_0}+\int_{\tau_0}^\infty\right)d\tau\sum_{\mathbf{M}\in L}{}' \tau^2\exp\{-\tau(\mathbf{x}-\mathbf{M}R)^2+i(\mathbf{x}-\mathbf{M}R)\mathbf{k}\}, \tag{14}$$

where $\tau_0>0$, and then we apply the Poisson transformation under the integral sign. As a result, we obtain

$$\sum_{\mathbf{M}\in L}\exp\{i\mathbf{k}(\mathbf{x}-\mathbf{M}R)-\tau(\mathbf{x}-\mathbf{M}R)^2\}$$
$$=\frac{1}{V_{\rm per}}\left(\frac{\pi}{\tau}\right)^{3/2}\sum_{\mathbf{g}\in G}\exp\left\{-\frac{(\mathbf{g}+\mathbf{k})^2}{4\tau}+i\mathbf{g}\cdot\mathbf{x}\right\},$$

where G is the reciprocal lattice of L, and $V_{\rm per}$ is the volume of the periodicity parallelepiped.

Since we shall henceforth be interested only in the ranges of values $R>10$ and $kr_0<2\pi/R$, after setting $\tau_0 r_0^2\sim 0.01$ in (14), we can retain only the term with $g=0$ in the sum obtained as a result of the Poisson transformation for our matrix element. This gives

$$W_{ij}(\mathbf{k})=\nu_0 r_0^6\left\{\sum_{\mathbf{M}\in L}{}'\ \exp\{-\tau_0(\mathbf{x}-\mathbf{M}R)^2+i(\mathbf{x}-\mathbf{M}R)\mathbf{k}\}\right.$$
$$\times\left[\frac{\tau_0^2}{2(\mathbf{x}-\mathbf{M}R)^2}+\frac{\tau_0}{[(\mathbf{x}-\mathbf{M}R)^2]^2}+\frac{1}{[(\mathbf{x}-\mathbf{M}R)^2]^3}\right]$$
$$\left.+\frac{\pi^{3/2}}{2R^3\Omega}\int_0^{\tau_0}d\tau\,\tau^{1/2}\exp\left(-\frac{\mathbf{k}^2}{4\tau}\right)\right\},\qquad(15)$$

where $\mathbf{x}=\mathbf{x}_i-\mathbf{x}_j$. When the value τ_0 is chosen, the upper bound for the modulus of the discarded sum

$$\sum_{\mathbf{g}\neq 0}\frac{\pi^{3/2}r_0^6}{V_{\rm per}}\int_0^{\tau_0}\tau^{1/2}\exp\left\{-\frac{(\mathbf{g}+\mathbf{k})^2}{4\tau}+i\mathbf{g}\cdot\mathbf{x}\right\}d\tau$$

can be estimated by the expression

$$\frac{\pi^{3/2}r_0^3}{1000V_{\rm per}}\sum_{\mathbf{g}\neq 0}\exp\left[-25\left(\mathbf{g}r_0-\frac{2\pi}{R}\right)^2\right]<10^{-16},$$

and it thus does not influence the results of the calculations. The nonanalytic contribution to the matrix element at small values of $|\mathbf{k}|$ is specified by the integral

$$\frac{\pi^{3/2}}{2}\int_0^{\tau_0}d\tau\,\tau^{1/2}\exp\left(-\frac{\mathbf{k}^2}{4\tau}\right)=\frac{\pi^2}{12}|\mathbf{k}|^3+\phi(\mathbf{k}),\qquad(16)$$

where

$$\phi(\mathbf{k})=\frac{\pi^{3/2}|\mathbf{k}|^3}{8}\left[\frac{\alpha^{-3/2}\exp(-\alpha)}{3}-\frac{2\alpha^{-1/2}\exp(-\alpha)}{3}\right.$$
$$\left.-\frac{4}{3}\int_0^{\sqrt{\alpha}}d\sigma\exp(-\sigma^2)\right],$$
$$\alpha=\frac{\mathbf{k}^2}{4\tau_0}$$

is a function which is analytic with respect to $\mathbf{k}$.

2.2. Let us perform a more detailed analysis of Eq. (11) at small values of $|\mathbf{k}|$ using perturbation theory.

The generator of Eq. (11) has the following easily verified properties:

1) $A_{ij}(\mathbf{k})^*=A_{ji}(\mathbf{k})=A_{ij}(-\mathbf{k})$;

2) $A_{ii}(\mathbf{0})>0$, $A_{i\neq j}(\mathbf{0})<0$, $\sum_i A_{ij}(\mathbf{0})=0$;

3) $\sum_{ij}\overline{\rho}_i A_{ij}(\mathbf{k})\rho_j$
$$=\frac{1}{2}\sum_{ij}\sum_{\mathbf{M}\in L}{}'\ \frac{|\rho_i\exp(i\mathbf{k}\cdot\mathbf{x}_i)-\rho_j\exp\{i\mathbf{k}(\mathbf{x}_j+R\mathbf{M})\}|^2}{|\mathbf{x}_i-\mathbf{x}_j-R\mathbf{M}|^6}.$$

Hence it follows that the generator $A_{ij}(\mathbf{k})$ is self-conjugate, the operator $A(\mathbf{0})$ is a generator of a connected Markov chain, and its smallest eigenvalue (which is equal to zero) is simple. Therefore, at small $|\mathbf{k}|$ the smallest eigenvalue of $A(\mathbf{k})$ is also simple. The eigenvector corresponding to the smallest eigenvalue of $A(\mathbf{0})$ appears in the initial conditions of Eq. (11), and at small $|\mathbf{k}|$ it differs slightly from the leading eigenvector of $A(\mathbf{k})$. When $\mathbf{k}=\mathbf{g}/R$, where $\mathbf{g}$ is a reciprocal lattice vector, the generator is nonnegative, and 0 is its simple eigenvalue. At all other (real) values of $\mathbf{k}$ the operator is positive.

For a more detailed analysis we represent (13) in the form

$$P(\mathbf{k},t|N)=\frac{1}{N}\sum_{i=1}^{N}\rho_i=\langle\overline{\mathbf{0}}|\exp[-A(\mathbf{k})t]|\overline{\mathbf{0}}\rangle,\qquad(17)$$

where $\langle i|\overline{\mathbf{0}}\rangle=1/\sqrt{N}$.

Using the spectral expansion of the Hermitian operator $A(\mathbf{k})$, we obtain

$$P(\mathbf{k},t|N)=\sum_{\mu=0}^{N-1}|\langle\mu|\overline{\mathbf{0}}\rangle|^2\exp[-a_\mu(\mathbf{k})t],$$
$$A(\mathbf{k})|\mu\rangle=a_\mu(\mathbf{k})|\mu\rangle.\qquad(18)$$

The lowest eigenvalue $a_0(\mathbf{k}\to\mathbf{0})$ is unique, and $|\langle 0|\overline{\mathbf{0}}\rangle|^2=1+O(\mathbf{k}^2)$. Introducing the projection operators $\pi=|\overline{\mathbf{0}}\rangle\langle\overline{\mathbf{0}}|$ and $\overline{\pi}=1-\pi$, we obtain

$$a_0(\mathbf{k})=\langle\overline{\mathbf{0}}|A(\mathbf{k})|\overline{\mathbf{0}}\rangle$$
$$+\langle\overline{\mathbf{0}}|A(\mathbf{k})\overline{\pi}\,\frac{1}{a_0(\mathbf{k})-\overline{\pi}A(\mathbf{k})\overline{\pi}}\,\overline{\pi}A(\mathbf{k})|\overline{\mathbf{0}}\rangle.\qquad(19)$$

The first nonanalytic term in the expansion of $A(\mathbf{k})$ is of the order of $|\mathbf{k}|^3$ [see (16)]. Since the resolvent is finite at small $|\mathbf{k}|$ and $a_0(\mathbf{k})$ is an even function of $\mathbf{k}$, the term $\sim|\mathbf{k}|^3$ in the expansion of $a_0(\mathbf{k})$ is completely determined by the first term in (19). Hence it follows that

$$a_0(\mathbf{k})=D_{\alpha\beta}k_\alpha k_\beta-\sigma k^3+O(\mathbf{k}^4)=\overline{a_0}(\mathbf{k})+O(\mathbf{k}^4),\qquad(20)$$

where $\sigma=(\pi^2/12)c(r_0^6/\Omega)\nu_0$.

It is convenient to represent the diffusion tensor in the form

$$D_{\alpha\beta}=(\kappa_{\alpha\beta}/6)\overline{r^2}\beta.\qquad(21)$$

As $N\to\infty$, because of the isotropy of the system $D_{\alpha\beta}\to D\delta_{\alpha\beta}$, and $\kappa_{\alpha\beta}\to\kappa\delta_{\alpha\beta}$.

We note that exact knowledge of the coefficient σ on the right-hand side of (20) ensures exact correspondence between the two leading terms of the asymptotic expansion of the autocorrelator $\mathscr{P}_{00}(t)$.

Taking into account that

$$A_{ij}\left(\mathbf{k}+\frac{\mathbf{g}}{R}\right)=\exp\left(-\frac{i\mathbf{g}\cdot\mathbf{x}_i}{R}\right)A_{ij}(\mathbf{k})\exp\left(\frac{i\mathbf{g}\cdot\mathbf{x}_j}{R}\right),$$

where $\mathbf{g}$ is an arbitrary reciprocal lattice vector, we can obtain the following representation for the propagator:

$$\begin{aligned}P_{\mathbf{x}0}&=\frac{1}{V_B}\int_B d^3k\exp(i\mathbf{k}\cdot\mathbf{x})P(\mathbf{k},t|N)\\&=\frac{1}{V_BN}\int_B d^3k\exp(i\mathbf{k}\cdot\mathbf{x})\sum_{ij}(\exp[-tA(\mathbf{k})])_{ij}\\&=\frac{1}{V_BN}\sum_{\mathbf{g}}\int_{B_\mathbf{g}} d^3k\exp(i\mathbf{k}\cdot\mathbf{x})\sum_{ij}(\exp[-tA(\mathbf{k})])_{ij}\\&=\frac{1}{V_BN}\sum_{\mathbf{g}}\int_{B_\mathbf{0}} d^3k\exp\left\{i\mathbf{x}\left(\mathbf{k}+\frac{\mathbf{g}}{R}\right)\right\}\\&\quad\times\sum_{ij}\exp\left\{\frac{i\mathbf{g}(\mathbf{x}_j-\mathbf{x}_i)}{R}\right\}(\exp[-tA(\mathbf{k})])_{ij}\\&=\frac{R^3}{V_BN}\int_{B_0} d^3k\exp(i\mathbf{k}\cdot\mathbf{x})\\&\quad\times\sum_{ij}(\exp[-tA(\mathbf{k})])_{ij}\delta_{\mathbf{x}-\mathbf{x}_i+\mathbf{x}_j,\mathbf{0}}.\end{aligned}\tag{22}$$

Here B denotes the Brillouin zone of volume $V_B=(2\pi)^3/\Omega$ corresponding to the original lattice, $B_\mathbf{g}$ is the Brillouin zone with a center at $\mathbf{g}/R$ corresponding to a cell equal to the periodicity parallelepiped, $\mathbf{g}/R$ runs through R^3 sites of the periodicity parallelepiped in the reciprocal lattice, and thus the shifts $B_\mathbf{g}$ of zone $B_\mathbf{0}$ by all the possible vectors $\mathbf{g}/R$ regularly fill the large zone B.

It can be seen from (22) that the formalism proposed here has a small integration region with respect to $\mathbf{k}$ in the propagator representation, which is an important property of the usual Fourier transformation of band theory (i.e., with respect to the left-hand coordinate in a lattice with the spacing R).

When the leading terms of the asymptotic expansion of the autocorrelator in powers of $1/t$ are calculated, it is sufficient to take into account only the main eigenvalue of $A(\mathbf{k})$ at small $\mathbf{k}$. As a result,

$$\begin{aligned}P_{00}(t\to+\infty)&=\frac{1}{V_Bc}\int_{B_0} d^3k\,\mathrm{Tr}(\exp[-tA(\mathbf{k})])\\&\approx\frac{1}{V_Bc}\int_{B_0} d^3k\exp[-ta_0(\mathbf{k})]\\&=\frac{\Omega}{c}[(4\pi Dt)^{-3/2}+\sigma(2\pi^2D^3t^2)^{-1}+O(t^{-5/2})].\end{aligned}\tag{23}$$

The relationship between two higher terms of the long-time expansion (see Refs. 10, 4, and 14), which is characteristic of the dipole-dipole interaction for very different distributions of the randomness of the medium, is thereby confirmed. This directly points out that the diffusion stage of the process is immediately preceded by a stage of slower decay of the autocorrelation functions and that in the final stage the plot of the autocorrelator approaches the asymptotic curve from above.

3. TESTING THE METHOD ON EXACTLY SOLVABLE MODELS

We tested all the stages of the application of the proposed approach to the investigation (and simulation) of random walks in disordered media on an extensive class of asymptotically, exactly solvable models with spatially unconstrained transport.[22–24]

To fix ideas we shall dwell on the isotropic random jump model, in which the propagator is defined by an equation of the form[23]

$$\dot{P}_{\mathbf{xy}}=-\sum_{\mathbf{z}}(\nu_{\mathbf{zx}}\xi_\mathbf{x}P_{\mathbf{xy}}-\nu_{\mathbf{xz}}\xi_\mathbf{z}P_{\mathbf{zy}}),\quad P_{\mathbf{xy}}(t=0)=\delta_{\mathbf{xy}},\tag{24}$$

where $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$ run through the entire simple cubic lattice L, $\nu_{\mathbf{xz}}=\nu_{\mathbf{zx}}=\nu_{\mathbf{z}-\mathbf{x},\mathbf{0}}$, $\nu_{\mathbf{00}}=0$, $\Sigma_\mathbf{x}\nu_{\mathbf{x0}}\mathbf{x}^2<+\infty$, and $\{\xi_\mathbf{x}\}$ is a set of independent, identically distributed positive random values with a sufficient number of finite inverse moments.

Let us consider a periodic system, leaving the $\xi_\mathbf{x}$ independent, where $\mathbf{x}$ runs through the sites of cube V with a center at zero and the edge Rr_0, and setting $\xi_{\mathbf{x}-R\mathbf{y}}=\xi_\mathbf{x}$ for any vector $\mathbf{y}\in L$.

In analogy to (9), we introduce

$$\rho_\mathbf{x}(\mathbf{k},t)=\sum_{\mathbf{y}}P_{\mathbf{xy}}(t)\exp\{i\mathbf{k}(\mathbf{y}-\mathbf{x})\}.\tag{25}$$

In this case $\rho_{\mathbf{x}-R\mathbf{y}}=\rho_\mathbf{x}$, and the following finite system of equations holds:

$$\begin{aligned}\dot{\rho}_\mathbf{x}&=-(A(\mathbf{k})\xi\rho)_\mathbf{x}=-\sum_{\mathbf{z}\in V}[W_{\mathbf{zx}}(\mathbf{0})\xi_\mathbf{x}\rho_\mathbf{x}-W_{\mathbf{xz}}(\mathbf{k})\xi_\mathbf{z}\rho_\mathbf{z}],\\\rho_\mathbf{x}(\mathbf{k},t=0)&=1,\end{aligned}\tag{26}$$

where $\mathbf{x}$ and $\mathbf{z}$ now run through all the sites of the periodicity cube and

$$W_{\mathbf{xz}}(\mathbf{k})=\sum_{\mathbf{M}\in L}\nu_{\mathbf{z},\mathbf{x}+\mathbf{M}R}\exp\{i\mathbf{k}(\mathbf{z}-\mathbf{x}-\mathbf{M}R)\}.\tag{27}$$

Instead of $P(\mathbf{k},t|N)$ [see (17)], it is now convenient to consider

$$\begin{aligned}\hat{P}(\mathbf{k},t)&\equiv\langle 0|\exp[-A\xi t]\frac{\kappa_0}{\xi}|0\rangle\\&=\langle 0|\sqrt{\frac{\kappa_0}{\xi}}\exp[-\sqrt{\xi}A\sqrt{\xi}t]\sqrt{\frac{\kappa_0}{\xi}}|0\rangle,\end{aligned}\tag{28}$$

where

$$\langle\mathbf{x}|0\rangle=\frac{1}{\sqrt{N}},\quad\frac{1}{\kappa_0}=\frac{1}{N}\sum_{\mathbf{x}\in V}\frac{1}{\xi_\mathbf{x}},\quad N=R^3.$$

The smallest eigenvalue of the operator $\sqrt{\xi}A\sqrt{\xi}$ is specified by an expression which coincides with (19) following

the replacements $\langle i|\overline{0}\rangle \to \langle \mathbf{x}|\overline{0}\rangle = \sqrt{\kappa_0/(\xi N)}$ and $A \to \sqrt{\xi} A \sqrt{\xi}$. We note that the second term in (19) is fourth-order with respect to $\mathbf{k}$ here and does not make a contribution to the diffusion coefficient. From the first term we find

$$D = \kappa_0 D_0, \quad D_0 = \frac{1}{6} \sum_{\mathbf{x}} \mathbf{x}^2 \nu_{\mathbf{x}}, \tag{29}$$

where D_0 is the diffusion coefficient in a regular lattice. When N is large, this result is in excellent agreement with the expression obtained in Ref. 23 for the diffusion coefficient

$$D = \kappa D_0, \quad 1/\kappa = \langle 1/\xi_0 \rangle,$$

where $\langle \ldots \rangle$ denotes averaging over an ensemble. This means that the limits $N \to \infty$ and $t \to \infty$ are interchangeable in the present model and that the relative accuracy of determining D from formulas (29) is of the order of $1/\sqrt{N}$. The interchangeability of these limits can be demonstrated in a similar manner for all the models considered in Ref. 24. The numerical algorithm for solving the system (26) was tested on the isotropic random jump model with $\nu_{\mathbf{xz}} \sim |\mathbf{x}-\mathbf{z}|^{-6}$ and uniformly distributed ξ with $\langle 1/\xi \rangle \approx 10/\langle \xi \rangle$. Just such a relation between $\langle \tilde{D}_i \rangle$ and $\langle 1/\tilde{D}_i \rangle$ is realized in the problem (1) when $c \approx 3\%$. The numerical solution of Eqs. (26) leads to a value of the diffusion coefficient coinciding with (29) to within 10^{-5}. This number characterizes the error in the calculations for a fixed realization of $\{\xi_{\mathbf{x}}\}$ and can serve as an estimate of the analogous error in the system (1) under investigation for an assigned impurity configuration, since below we shall use the same algorithm for the numerical calculation. The statistical spread of the diffusion coefficient is specified in the isotropic random jump model by the distribution of $\{\xi_{\mathbf{x}}\}$ [see Eq. (28)] and in the real system (1) by the number of sites N in the periodicity volume and the concentration c. It is considerably greater than the error of the numerical algorithm, and, therefore, the latter error will not be discussed further.

4. SIMULATION RESULTS

Transport cannot be investigated in our basic system (1) by purely analytic means with the same completeness as in the isotropic random jump model.

We carried out numerical simulations of migration on the basis of (11). Figures 1 and 2 present the results of a calculation of $P(\mathbf{k},t|N)$ in simple and face-centered cubic lattices, where each line corresponds to one impurity configuration. It can be seen from the curves that at large t the value of $-\ln P(\mathbf{k},t|N)/\beta t$ tends to a finite positive value. Precisely the same behavior of the Fourier transform of the propagator was observed for all the values of N considered from 200 to 2700 and at all concentrations regardless of the type of lattice. Thus, for $\beta t \gg 1$

$$P(\mathbf{k},t|N) = a \exp[-\omega(\mathbf{k})t + \sigma k^3 t], \tag{30}$$

where the coefficient σ is known exactly [see Eq. (20)].

Figure 3 presents a plot of $\omega(\mathbf{k})$ for one impurity configuration and $\beta t \approx 40$. Similar behavior of $\omega(\mathbf{k})$ was observed for different impurity configurations regardless of the concentration and lattice type. As $N \to \infty$, ω depends on $|\mathbf{k}|$

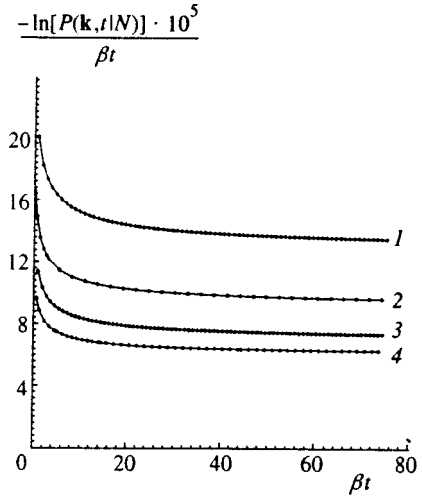

FIG. 1. Time dependence of the Fourier transform of the propagator for a simple cubic lattice: *1* — the number of sites in the periodicity volume $N = 240$, the concentration $c = 0.03$, and $\mathbf{k} = (\pi/200)(0,1,0)$; *2* — $N = 1350$, $c = 0.05$, and $\mathbf{k} = (\pi/200)(0,0,1)$; *3* — $N = 640$, $c = 0.08$, and $\mathbf{k} = (\pi/200) \times (1,0,0)$; *4* — $N = 1970$, $c = 0.1$, and $\mathbf{k} = (\pi/200)(0,0,1)$.

because of the isotropy of the system. The numerical calculations confirm the absence of a regular dependence of ω on the direction of the vector $\mathbf{k}$. Keeping the direction of $\mathbf{k}$ constant, we vary its length and approximate the values of $\omega(\mathbf{k})$ obtained according to the least-squares method by the expression $\omega(\mathbf{k}) = Dk^2 + \alpha k^4 = \beta[(\kappa/6)(k\overline{r})^2 + \alpha_1 (k\overline{r})^4]$. The terms of order $o(\mathbf{k}^4)$ are discarded, since small $\mathbf{k}$ are considered: $k\overline{r} \sim 0.05$. We obtain $\kappa = 0.3101(2)$ and $\alpha_1 = -0.0102(2)$. Thus, when $\beta t \sim 40$, it is seen that the behavior of the system is approximated by the expression $P(\mathbf{k},t|N) = a \exp[(-Dk^2 + \sigma k^3 + O(k^4))t]$, where the terms $\sim \mathbf{k}^4$ can be discarded to within the accuracy of the calculation of the diffusion coefficient. Then, up to $\beta t \sim 1000$ the

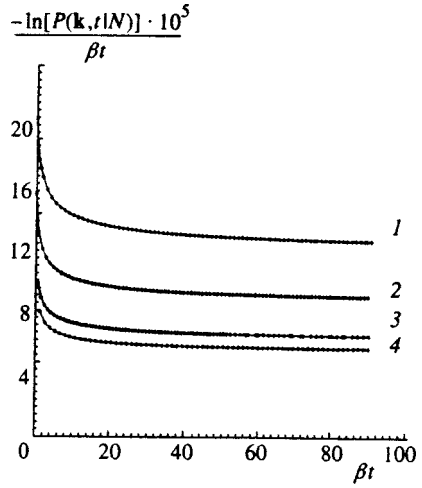

FIG. 2. Time dependence of the Fourier transform of the propagator for an fcc lattice: *1* — the number of sites in the periodicity volume $N = 240$, the concentration $c = 0.03$, and $\mathbf{k} = (\pi/200)(0,1,0)$; *2* — $N = 780$, $c = 0.05$, and $\mathbf{k} = (\pi/200)(0,0,1)$; *3* — $N = 640$, $c = 0.08$, and $\mathbf{k} = (\pi/200)(1,0,0)$; *4* — $N = 1760$, $c = 0.1$, and $\mathbf{k} = (\pi/200)(0,0,1)$.

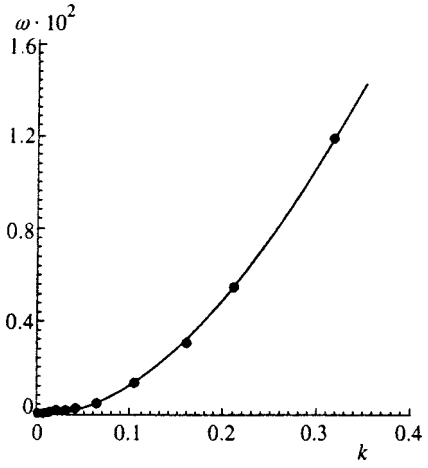

FIG. 3. Dependence of ω on $|\mathbf{k}|$ (points) and its least-squares approximation (continuous line); $N=1800$, $c=0.1$.

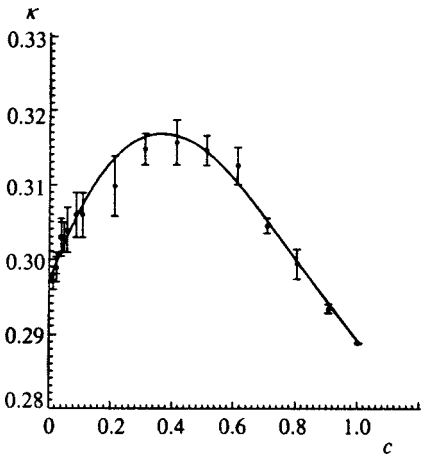

FIG. 5. Concentration dependence of κ and its approximation by a third-order polynomial. The lattice is face-centered cubic.

diffusion coefficient D decreases monotonically, remaining within 3% of the value for $\beta t=40$ and within 0.5% of the value for $\beta t=80$. The long-time asymptotics were investigated for various concentrations from $c=1\%$ to $c=90\%$. At each value of c we considered from three to five different pseudorandom configurations of the impurity sites and averaged the values of κ obtained over them. The results are presented in Fig. 4. A similar investigation of the concentration dependence of κ was performed for an fcc lattice (Fig. 5). As can be seen from Figs. 4 and 5, at large $c\to 1$ the value of κ tends to the known value of $\kappa(c=1)$: for a simple cubic lattice $\kappa(c=1)=0.2999$, and for an fcc lattice $\kappa(c=1)=0.2895$.

We constructed the approximation $\kappa=\kappa(c=1)+(\kappa_1+\kappa_2 c+\kappa_3 c^2)(1-c)$ by the least-squares method, and for a simple cubic lattice we obtained $\kappa_1=-0.0050(6)$, $\kappa_2=0.164(4)$, and $\kappa_3=-0.11(4)$. The value $\kappa_0=\kappa(c=1)+\kappa_1=0.2949(6)$ corresponds to the diffusion coefficient for $c\to 0$.

For an fcc lattice $\kappa_1=0.0066(7)$, $\kappa_2=0.132(4)$, and $\kappa_3=-0.09(4)$. In this case $\kappa_0=\kappa(c=1)+\kappa_1=0.2961(7)$, which coincides to within the error with the value of κ_0 in the simple lattice. The lack of a dependence of κ on the type of lattice at low concentrations was verified by investigating a random-walk process on a diamond lattice. The value $\kappa(c=1)=0.3299$ for this lattice differs significantly from κ_0. It was found at small concentrations that within the calculation error κ has the same value as in the case of simple cubic and fcc lattices. Such agreement should be observed, since at small concentrations the mean distance between impurities $\overline{r}$ is much greater than r_0, and the asymptotics of the migration process should not depend at all on the type of lattice.[11]

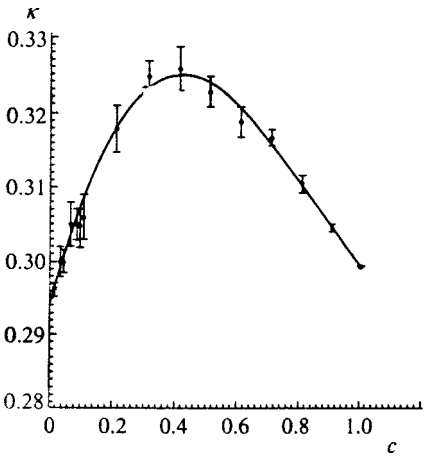

FIG. 4. Concentration dependence of κ and its approximation by a third-order polynomial. The lattice is simple cubic.

4.1. To speed up the calculations we employed the approximation

$$P(\mathbf{k},t|N)=\exp[-\overline{a_0}(\mathbf{k})(t+b/\beta)]+(\overline{rk})^2 f\exp(-a_1 t),$$
$$a_1=(2\pi/R)^2 D\kappa', \tag{31}$$

where b, f, and κ' are constants, which have finite limits for $c\to 0$ and $k\to 0$. This shortened the calculation time by roughly one third.

As a control, the diffusion coefficient was also determined by a third method with even faster convergence, which is associated with the relation

$$\overline{A}(\mathbf{k})=\sum_{ij}\rho_i^* A_{ij}(\mathbf{k})\rho_j \Big/ \sum_i |\rho_i|^2 \to \overline{a_0}(\mathbf{k}) \quad \text{for } t\to\infty. \tag{32}$$

In fact, using the spectral expansion, we have

$$\overline{A}(\mathbf{k})=\frac{\langle\rho|A(\mathbf{k})|\rho\rangle}{\langle\rho|\rho\rangle}$$
$$=\frac{\sum_\mu \langle\overline{\mathbf{0}}|\mu\rangle|^2 a_\mu(\mathbf{k})\exp[-2a_\mu(\mathbf{k})t]}{\sum_\mu |\langle\overline{\mathbf{0}}|\mu\rangle|^2\exp[-2a_\mu(\mathbf{k})t]}. \tag{33}$$

Thus, the high rate of convergence is due to the doubled exponents. More specifically, the value of the diffusion co-

efficient D calculated on the basis of (32) stabilizes to within an error less than 3% already at $\beta t=20$ and less than 0.5% at $\beta t=40$.

4.2. Extrapolation procedures for solving systems of ordinary differential equations like the classical Adams method[26] were used in the numerical simulation. By selecting quadrature formulas with consideration of the certain positiveness of the generator spectrum $A(\mathbf{k})$, we were able to markedly increase the integration step without the danger of computational instability, which would be virtually unavoidable if there are eigenvalues that are not positive.

We considered formulas of the type

$$\rho(t+4h)\approx(a+phA)\rho(t)+(b+qhA)\rho(t+h)$$
$$+(c+rhA)\rho(t+2h)+(d+shA)\rho(t+3h),$$
$$p=-3(b+d)/8-c/3,\quad q=(64-19b-32c-27d)/24,$$
$$r=(-32+5b-8c-27d)/24,\quad s=(64-b-9d)/24,$$

where

$$a=-f_1f_2f_3,\quad b=f_1f_2f_3+f_1f_2+f_2f_3+f_3f_1,$$
$$c=-(f_1+f_2+f_3+f_1f_2+f_2f_3+f_3f_1),$$
$$d=f_1+f_2+f_3+1$$

or

$$a=-f_1(f_2+f_3^2),\quad b=(f_1+1)(f_2+2f_3+f_3^2),$$
$$c=-[f_1+f_2+f_3^2+2(1+f_1)f_3],\quad d=f_1+1+2f_3.$$

In the first case $0<f_1,f_2,f_3<1$, in the second case the real quantities f_1, f_2, and f_3 are chosen so that the conditions $0<f_1,f_3<1$ and $f_1^2+f_2^2<1$ would hold.

As opposed to the standard four-point Adams procedure, which more correctly allows for all the possible signs of the eigenvalues of the generator of a linear (or linearized) system, the integration step allows a fourfold increase in the case of the system (11).

5. DISCUSSION OF RESULTS

The proposed method faithfully simulates the disordered system (1) provided the diffusion radius $r_D=\sqrt{6Dt}$ satisfies the condition

$$\varepsilon_D=(2r_D/R)^2=4\kappa\beta t/N^{2/3}<1. \tag{34}$$

In the opposite limit ($\varepsilon_D>1$) we have a model of transport in a crystal with a large complicated unit cell containing N atoms. The value of κ calculated for small $kr_0\leqslant 0.1\cdot 2\pi/R$ according to (30) stabilizes within 3% at all $\beta t\geqslant 40$ and $N\geqslant 150$ and varies monotonically as t increases. At the largest number of sites used ($N\approx 2700$) the criterion (34) holds up to $\beta t_{\max}\approx 160$. Therefore, in our basic problem (1) and (3), as in the case of the exactly solvable models in Refs. 23 and 24, the equality of limits

$$\lim_{t\to\infty}\lim_{N\to\infty}D(t,N)=\lim_{N\to\infty}\lim_{t\to\infty}D(t,N)\quad[=D]$$

is valid at least to an error of $<3\%$, and the long-time asymptotics in the system under consideration are diffusive.

We note that the asymptotic stage sets in at considerably shorter times for $\mathscr{P}(\mathbf{k},t)$ than for the autocorrelator $\mathscr{P}_{\mathbf{00}}(t)$. This property is realized both in all known methods of theoretical analysis[14,18] and in numerical simulation, allowing us to confine ourselves to comparatively small values of $N\approx 500$ and obtain results with sufficient accuracy.

A comparison of the value obtained above $\kappa_0=0.2954(5)$ with the literature values of the diffusion coefficient in the dipole-dipole disordered system (1) reveals that the value closest to ours was offered in the GAF theory:[9] $\kappa_{\mathrm{GAF}}=0.315$. The modification of the GAF theory in Refs. 13 and 14 (which was developed, in particular, to generalize the theory to one-dimensional and two-dimensional systems) yields $\kappa_0=\kappa_D=0.186$, although the heuristic, rather than mathematical, character of the proposed modification, as well as of the GAF theory itself, should be noted here. The semiphenomenological theory[10,14] uses the Scher–Lax[7] formula to calculate D and gives $\kappa_0=\kappa_{\mathrm{SL}}=0.373$. Goldman and Jacquinot[15] obtained $\kappa_0=\kappa_{\mathrm{GJ}}=0.49$.

Two experimental studies whose results have bearing on the investigation of the long-time behavior of the propagator have been performed. One of them[4] proved the occurrence of ''reoscillation,'' which we discussed above as a consequence of the relation (23), but no measurements were performed in the time interval needed for determining the diffusion coefficient D, and the value $\kappa_0=\kappa_{\mathrm{SL}}$ was used in the treatment of the results. Thus, on the basis of Ref. 4 it can only be concluded that this value of κ_0 does not contradict the results of that work. Conversely, the work described in Ref. 5 was directed at measuring D, and the result corresponds to $\kappa_0=\kappa_{\mathrm{GKM}}=0.147(23)$, which gives $\overline{\kappa}_{\mathrm{GKM}}=0.168(26)$ when the term σk^3 from (20) is correctly taken into account.[13,14] A correction for the dipolar anisotropy of the transition rates increases this number to $\overline{\kappa}_{\mathrm{GKM}}=0.187(29)$.[13,14] We stress, however, that the dependence of the results on the wave vector $\mathbf{k}$ was investigated inadequately in the experiment in Ref. 5 (as was noted by the authors themselves); therefore, the disparity between this value and the results of our calculations point out the need for new more exact and systematic measurements.

6. CONCLUSION

Let us briefly characterize the main features of the method which we used to investigate the asymptoticss of random-walk processes in disordered media.

At small values of $|\mathbf{k}|$ the finite systems (11) and (26) obtained for periodic configurations of random media provide the main information on the long-time asymptotic behavior, as follows from the following arguments.

1. The system (11), which is obtained from (1) by Fourier transformation in a periodic impurity configuration, describes the process in an infinite configuration with an excitation which is initially localized on one site. It eliminates the need to stop the process before it reaches a (nonexisting) boundary, which is encountered when the infinite configuration is replaced by a cube or when the correlator p_{ij} is continued periodically.[19,20] The exponential asymptotics expansion of ρ_i, i.e., the analog of the Fourier transform of the

correlator, following from the finite character of the linear system (11) is no longer undesirable and corresponds to the essence of the problem in a reliable estimate of the contribution of the remainder of the spectrum. Disclosing the asymptotic behavior of the process requires elucidation of the character of the dependence of the leading eigenvalue of the generator $A(\mathbf{k})$ on $\mathbf{k}$ at small values of $|\mathbf{k}|$.

2. We are able to confine ourselves to small values of $|\mathbf{k}|$ in obtaining the long-time asymptotics. Therefore, we can speed up the calculations significantly, since the initial vector of (11) for $\mathbf{k}=\mathbf{0}$ is the leading eigenvector of $A(\mathbf{k}=\mathbf{0})$. It is significant that $A(\mathbf{k}=\mathbf{0})$, being a symmetric generator of an irreducible Markov chain, has a single smallest eigenvalue equal to zero, while the operator $A(\mathbf{k})$ for $\mathbf{k}\neq\mathbf{0}$ is Hermitian and positive.

3. In all other respects unrestricted continuation of the calculations in time only refines the limiting form of this vector, i.e., the eigenvector of the generator, and the leading eigenvalue corresponding to it at $t\to+\infty$ even when calculation and rounding errors accumulate in the normalization of the vector ρ_i.

4. The calculation time of the parameters of the long-time asymptotics is shortened, if $P(\mathbf{k},t|N)$ is approximated by a sum of two exponential functions [see (31)] or if the tendency of the quantity $\rho^+A\rho/\rho^+\rho$ [see (32)] to approach the leading eigenvalue of the generator of (11) is taken into account, which is markedly more effective.

5. A highly important property, which should be verified in such calculations, is insensitivity toward the periodicity boundaries, i.e., establishment of the fact that the asymptotic state is clearly seen at times when the excitation is still fairly far from the boundaries of the cube.

This research was performed with support from the Russian Fund for Fundamental Research (Projects 96-15-96416 and 97-01-00714).

*)E-mail: stepanov@vitep5.itep.ru

[1] C. R. Gochanour and M. D. Fayer, J. Phys. Chem. **85**, 1989 (1981).
[2] D. E. Hart, P. A. Anfinrud, and W. S. Struve, J. Phys. Chem. **86**, 2689 (1987).
[3] M. Kh. Ashurov, T. T. Basiev, A. I. Burshteĭn *et al.*, JETP Lett. **40**, 841 (1984).
[4] V. P. Gapontsev, F. S. Dzheparov, N. S. Platonov, V. E. Shestopal, JETP Lett. **41**, 561 (1985).
[5] L. Gómez-Jahn, J. Kasinsky, and R. J. D. Miller, Chem. Phys. Lett. **125**, 500 (1986).
[6] Yu. G. Abov, M. I. Bulgakov, S. P. Borovlev *et al.*, Zh. Éksp. Teor. Fiz. **99**, 962 (1991) [Sov. Phys. JETP **72**, 534 (1991)].
[7] H. Scher and M. Lax, Phys. Rev. B **7**, 4491, 4502 (1973).
[8] W. Y. Ching, D. I. Huber, and B. Barnet, Phys. Rev. B **17**, 5025 (1978).
[9] C. R. Gochanour, H. C. Andersen, and M. D. Fayer, J. Phys. Chem. **70**, 4254 (1979).
[10] F. S. Dzheparov, Radiospektroskopiya (Perm') **13**, 135 (1980).
[11] F. S. Dzheparov, V. S. Smelov, and V. E. Shestopal, JETP Lett. **32**, 47 (1980).
[12] F. S. Dzheparov and V. E. Shestopal, Izv. Vyssh. Uchebn. Zaved. Fiz. **30**(6), 77 (1987).
[13] F. S. Dzheparov, JETP Lett. **52**, 269 (1990).
[14] F. S. Dzheparov, Zh. Éksp. Teor. Fiz. **99**, 982 (1991) [Sov. Phys. JETP **72**, 546 (1991)].
[15] M. Goldman and J. Jacquinot, J. Phys. (Paris) **43**, 1049 (1982).
[16] B. Derrida, J. Stat. Phys. **31**, 433 (1983).
[17] B. Derrida and J. M. Luck, Phys. Rev. B **28**, 7183 (1983).
[18] J.-P. Bouchaud and A. George, Phys. Rep. **195**, 127 (1990).
[19] S. K. Lyo, Phys. Rev. B **20**, 1297 (1979).
[20] O. K. Alimov, M. Kh. Ashurov, T. T. Basiev *et al.*, Tr. IOFAN **9**, 60 (1987).
[21] F. S. Dzheparov, D. V. L'vov, K. N. Nechaev, and V. E. Shestopal, JETP Lett. **62**, 662 (1995).
[22] P. J. H. Denteneer and M. H. Ernst, Phys. Rev. B **29**, 1755 (1984).
[23] F. S. Dzheparov and V. E. Shestopal, Teor. Mat. Fiz. **94**, 496 (1993).
[24] F. S. Dzheparov and V. E. Shestopal, JETP Lett. **60**, 182 (1994).
[25] R. P. Feynman, *Statistical Mechanics: a Set of Lectures*, Benjamin, Reading, Mass. (1972); Mir, Moscow (1978).
[26] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations, Vol. 1: Nonstiff Problems*, Springer-Verlag, Berlin–New York (1987); Mir, Moscow (1990).

Translated and edited by P. Shelnitz

Thermal expansion and equation for state of KCN of different isotopic composition

R. B. Von Dreele and J. G. Morgan

Lujan Center, MS H805, Los Alamos National Laboratory, Los Alamos, NM, 87545 USA

S. M. Stishov*)

Lujan Center, MS H805, Los Alamos National Laboratory, Los Alamos, NM, 87545 USA; Institute for High Pressure Physics, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

(Submitted 7 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 2182–2186 (December 1998)

Thermal expansion and the equation of state of isotopically different samples of KCN were measured in the temperature range 175–300 K and at pressures up to 2 GPa by the neutron powder diffraction technique. No evidence of the influence of isotopic composition on the thermal expansion or the compression isotherms was found. © *1998 American Institute of Physics.* [S1063-7761(98)02112-X]

1. INTRODUCTION

The influence of isotopic composition on the thermodynamic properties of substances is usually considered to be a manifestation of the quantum nature of matter. At high temperatures in the classical limit any mass effect on thermodynamic quantities disappears and hence any quantum isotopic effect vanishes. However, isotope effects of a nonquantum nature may exist under certain circumstances in condensed systems. For instance, variations of the isotopic composition in molecular systems can change the moments of inertia of the molecules, which can then influence the thermodynamic properties of the system. This is quite obvious in an ideal classical system of polyatomic molecules, where the rotational part of the free energy, entropy and thermal capacity directly depends on the molecular moments of inertia.[1] As we will see below, in the case of condensed systems one might also expect to find corresponding contributions to the compressibility and thermal expansion as well.

Below we report data on thermal expansion and equation of states (175–300 K and 0–2.0 GPa) of KCN, substituted with various isotopes of carbon and nitrogen so that the CN^- ions have the same mass number but with different moments of inertia. The corresponding experimental data were obtained by means of the neutron powder diffraction technique.

The phase diagram and various physical properties of normal KCN have been reported in a number of papers (see for example, Refs. 2–7). The general conclusion is that in the temperature and pressure range under study, KCN has NaCl crystalline structure in which the orientations of the CN^- ions are disordered over a number of directions. Furthermore, the orientations of individual ions are not fixed but change very rapidly. To simplify the description, one may say that CN^- ions in the KCN-l phase experience almost free rotation at near ambient temperatures. In this case, using the classical approach the rotational contribution to the free energy can be written in the form[1]

$$F_{\rm rot} = -NT\ln T - NT\ln(2J_{\rm eff}/\hbar^2). \tag{1}$$

For CN as a diatomic gas molecule the value of the rotational quantum, $\hbar^2/2J$, where J is the moment of inertia, is about 3 K. This looks exactly like the expression for an ideal gas of diatomic molecules, except that the true moment of inertia J has been replaced with an effective value $J_{\rm eff}$ to take into account, at least crudely, the interaction of a molecule with its environment. The corresponding contribution to the pressure can be derived from (1):

$$p_{\rm rot} = -\left(\frac{\partial F}{\partial V}\right)_T = \left(\frac{NT}{J_{\rm eff}}\right)\left(\frac{\partial J_{\rm eff}}{\partial V}\right)_T. \tag{2}$$

It is natural to assume that the effective moment of inertia $J_{\rm eff}$ increases with density, and hence from (2) it follows that molecular rotation gives a negative contribution to the total pressure. As can be seen from (2), the ''rotational'' pressure depends on the isotopic composition (through $J_{\rm eff}$), and hence the compression isotherm and the thermal expansion curves of KCN may be split due to the difference in effective moments of inertia of isotopically substituted CN^- ions.

2. EXPERIMENTAL

Isotopically enriched samples of KCN were obtained from the Cambridge Isotope Laboratories. They had composition KC ^{15}N (C—natural abundance; 99% ^{15}N) and K^{13}CN (99% ^{13}C: N—natural abundance). These materials were used without further purification. Both samples were subjected to ion beam analysis at the Los Alamos National Laboratory Ion Beam Materials Science Facility; both materials were found to have 180 ppm of Na contamination.

Variable temperature neutron powder diffraction studies were performed on the HIPD instrument at the Manuel Lujan, Jr., Neutron Scattering Center, Los Alamos National Laboratory. Samples of the two KCN materials were placed in 6.35×38.1-mm vanadium cans with a small amount of He gas and subsequently sealed. The thermal expansion was studied in the temperature range 175–300 K in 25 K steps, making use of a Displex (Air Products) system, which provides temperature control to within 0.1 K. The neutron dif-

TABLE I. Thermal expansion data for isotopically substituted KCN* at ambient pressure.

T, K	a_0 ($K^{13}CN$), Å	a_0 ($KC^{15}N$), Å
300	6.5278	6.5290
275	6.5196	6.5184
250	6.5119	6.5118
225	6.5040	6.5037
200	6.4960	6.4947
175	6.4876	6.4876

*Standard deviations for these measurements are $(1-2)\cdot 10^{-4}$ Å. Propagation of error from calibration data gives an estimated error of 4×10^{-4} Å.

TABLE II. Compressibility data for isotopically substituted KCN at 300 K*.

P, GPa	a_0 ($K^{13}CN$), Å	P, GPa	a_0 ($KC^{15}N$), Å
0.290(35)	6.4809(10)	0.194(19)	6.5126(9)
0.822(16)	6.4066(10)	0.583(22)	6.4556(9)
0.894(46)	6.4146(10)	1.153(26)	6.3882(10)
1.049(19)	6.3782(10)	1.844(33)	6.3142(12)
1.406(21)	6.3448(10)		
1.873(28)	6.2996(11)		
1.925(27)	6.2922(26)		

*Values in parentheses are estimated standard deviations obtained via propagation of error from the Rietveld refinements.

fraction patterns were obtained in 2–4 h each with the spallation neutron source operating at a proton current of ~70 μA: data taken in the backscattering detectors ($2\theta = 151°$) were used for subsequent data analysis. To obtain accurate values of the ambient lattice parameters for the KCN samples, neutron powder diffraction patterns were also obtained at 300 K from each KCN mixed with a small amount of CaF_2, which provided an internal lattice parameter standard ($a_0 = 5.46384$ Å at 300 K). The lattice parameter of this sample of calcium fluoride was previously determined in our laboratory by x-ray powder diffractometry with Cu K_α radiation.

The compression isotherms of the two KCN samples at 300 K up to 2 GPa were obtained with a toroidal high pressure cell.[8,9] To provide hydrostatic conditions, the samples were mixed with Fluorinert-70 (3M Co.) and placed in the cell along with a piece of lead, which was used as a pressure sensor.[10] Neutron diffraction patterns were obtained on the HIPD over 8–12 h each with the spallation neutron source operating at a proton current of ~70 μA; data taken at $2\theta = \pm 90°$ was used for subsequent data analysis.

3. DATA ANALYSIS

One problem with neutron time-of-flight measurements is that the measured lattice parameters are affected by the location of the sample scattering center in the diffractometer. The placement of the sample, particularly in the high-pressure cell, is not sufficiently reproducible to eliminate this source of systematic bias in measured lattice parameters. To mitigate this systematic effect, we have employed the data analysis procedures that follow. Both variable-pressure and variable-temperature diffraction data were processed using the General Structure Analysis System (GSAS) suite of computer programs[11] via Rietveld refinement to obtain the lattice parameters of each KCN sample, the Pb pressure standard, and the CaF_2 calibration material as appropriate. Refinement of the 300 K diffraction data of the two KCN/CaF_2 mixtures gave $a_0 = 6.5278(2)$ Å for $K^{13}CN$ and $a_0 = 6.5280(2)$ Å for $KC^{15}N$. The use of an internal standard (CaF_2) eliminates the sample position bias, so these KCN lattice parameters are at least self-consistent. These values were then used to calibrate both the variable-temperature and variable-pressure measurements, which are listed in Tables I and II. The corrections for Table I were determined from the intercept at 300 K of a linear fit to the measured lattice parameters with temperature. The corrections correspond to offsets of $\Delta a/a = -0.5\times 10^{-4}$ and 0.8×10^{-4} for $K^{13}CN$ and $KC^{15}N$, respectively. Sample position errors of less than 0.1 mm within the diffractometer would give these offsets, and they are smaller than the random errors in the individual lattice parameter values obtained from the Rietveld refinements. The variable-pressure data were corrected in a similar way. A simple second-order polynomial was fit to the lattice parameters for each KCN with pressure. The intercept at ambient conditions was then used to correct both the KCN and Pb lattice parameters. The pressure was then recalculated from the equation of state for lead to give the values in Table II. These corrections are much larger than for the variable temperature data and correspond to $\Delta a/a = 3.2\times 10^{-3}$ and 0.8×10^{-3} for $K^{13}CN$ and $KC^{15}N$, respectively. These arise from sample position errors of 2–3 mm in the location of the pressure cell relative to the standard sample position in HIPD.

4. DISCUSSION

The data obtained (Tables I and II) are illustrated by Figs. 1 and 2. The linear thermal expansions from the data shown in Fig. 1 are 0.320×10^{-3} Å/K for $K^{13}CN$ and 0.327×10^{-3} Å/K for $KC^{15}N$. A fit of the combined compression data for the two KCN materials in Fig. 2 to a Murnaghan equation[12] gives the parameters $V_0 = 279.0(5)$ Å^3, $B_0 = 13.3(9)$ GPa and $B_0' = (\partial B/\partial P)_{P=0} = 5.6(10)$ (to be compared with values $B_0 = 13.2$ GPa and $B_0' = 4.2$ previously obtained for ''natural abundance'' KCN[5]). As seen from Fig. 1, the thermal expansion experiments reveal that there is no clear difference in lattice parameters of isotopically substi-

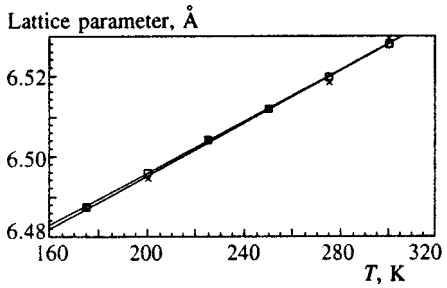

FIG. 1. Lattice parameter thermal expansion of KCN. The data for $K^{13}CN$ are plotted as squares and for $KC^{15}N$ as crosses. The lines are a guide to the eye.

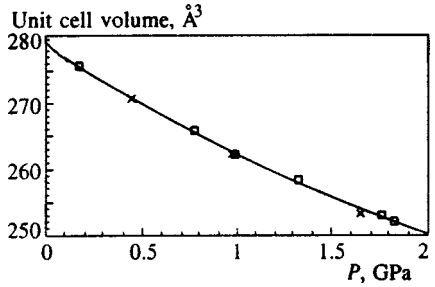

FIG. 2. Isothermal compression curve for KCN at 300 K. The data for $K^{13}CN$ are plotted as squares and for $KC^{15}N$ as crosses. The line is a guide to the eye.

tuted samples of KCN over a large range of temperatures. The high-pressure compression isotherms (Fig. 2) also reveal no difference, although the scatter in the data is much greater than in the temperature experiments.

Given that the reduced masses of the $^{13}C\,^{14}N$ and $^{12}C\,^{15}N$ ions differ by $\sim$1%, their true moments of inertia will differ by the same amount.[1)] Our accuracy in the lattice parameter determinations is $<0.01\%$ and $<0.1\%$ for the temperature and pressure runs, respectively. So we must see an isotopic effect unless the function $J_{\text{eff}}(V)$ has a special form. First, it could be equal to a constant, $J_{\text{eff}}=\text{const}$, in the density range studied; in this case there is no ''rotational'' contribution to the equation of state. However, this seems highly improbable because of the condensed nature of the substances under study. Second, it could depend on density in a linear way like $J_{\text{eff}}=J_0(1+a/V)$. In this case the ''rotational'' contribution to the equation of state could be large, though it would depend only slightly, if at all, on isotopic composition. Another possibility is that the ''free rotation'' approach can not be applied to the alkali cyanide family.[2)] In any case, further study is needed to resolve this problem and precise thermal capacity measurements would play a decisive role in the attempt to observe the classical isotopic effect.

This work was supported by the Manuel Lujan, Jr., Neutron Scattering Center, Los Alamos National Laboratory, funded in part by BES/DOE of USA under contract No. W-7405-ENG-36 and the Russian Fund for Fundamental Research (Grant No. 96-02-17220). We also thank M. Rubin and C. Maggiore at the Los Alamos Ion Beam Materials Science Facility for analytical work.

*)E-mail: sergei@hppi.troitsk.ru

1)The moment of inertia of a diatomic molecule $J=m_r l^2$, where $m_r=m_1m_2/(m_1+m_2)$ is a reduced mass and l is an interatomic distance. So, neglecting a possible slight difference in the length of the CN group due to the isotopic composition variations, we get for the moments of inertia of $^{13}C\,^{14}N$ and $^{12}C\,^{15}N$ values of 8.92 Å^2 and 8.82 Å^2, respectively at $l=1.15$ Å.

2)Obviously, the centers of mass of ions $^{13}C\,^{14}N$ and $^{12}C^{15}N$ do not coincide with their symmetry centers. That means that at ''free'' rotation the effective diameters of these ions would differ by $\sim$0.04 Å, and it would influence the lattice parameters.

[1] L. D. Landau and E. M. Lifshits, Statistical Physics, 3rd Edition, P. 1, Pergamon Press.
[2] D. L. Price, J. M. Rowe, J. J. Rush *et al.*, J. Chem. Phys. **56**, 3697 (1972).
[3] D. L. Decker, R. A. Beyerlein, G. Roult, and T. G. Worlton, Phys. Rev. **10**, 3584 (1974).
[4] J. M. Rowe and S. Susman, Phys. Rev. B **29**, 4727 (1984).
[5] K. Strössner, H. D. Hochheimer, W. Hönle, and A. Werner, J. Chem. Phys. **83**, 2435 (1985).
[6] H. T. Stokes, D. L. Decker, H. M. Nelson, and J. D. Jorgensen, Phys. Rev. B **47**, 11082 (1993).
[7] P. Bourson and D. Durand, J. Phys.: Condens. Matter **5**, 1723 (1993).
[8] L. G. Khvostantsev, L. F. Vereshchagin, and A. P. Novikov, High Temp.-High Press. **9**, 637 (1977).
[9] J. M. Besson, G. Hamel, R. J. Nelmes, J. S. Loveday, S. Hull, and D. Hausermann, High Press. Res. **8**, 625 (1992).
[10] R. A. Miller and D. E. Schuele, J. Phys. Chem. Solids **30**, 589 (1969).
[11] A. C. Larson and R. B. Von Dreele, Los Alamos National Laboratory Report LAUR 86-748 (1994).
[12] O. L. Anderson, J. Phys. Chem. Sol. **27**, 547 (1965).

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

Feasibility of covalent quasicrystals

V. A. Borodin and V. M. Manichev*)

Kurchatov Institute, 123182 Moscow, Russia
(Submitted 8 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 2187–2193 (December 1998)

An algorithm for constructing a quasicrystalline structure based on atoms with tetrahederal coordination of covalent atomic bonds (the typical coordination for group IV atoms) is proposed. The algorithm is used to construct a computer model of a cluster with icosahedral symmetry. The model is used to estimate the energy parameters of the structure and the distributions of bonds over interatomic bond lengths and interbond angles. The distributions obtained correlate well with the analogous results for silicon glasses and do not impose any fundamental constraints on the implementation of such structures in practice. © *1998 American Institute of Physics.* [S1063-7761(98)02212-4]

1. INTRODUCTION

Quasicrystals, which were discovered in 1984,[1] are an interesting example of a previously unknown type of atomic structure of solids. As a rule, the same atomic compositions can be obtained both in the amorphous state (metallic glass) and in the ordinary crystalline form. With respect to their atomic structure quasicrystals are, in a certain sense, intermediate between metallic glasses and true crystals. The local ordering of the atoms (short-range order) in all these structures is very similar, the basic distinction being the existence and type of long-range order. At the same time, there is no long-range order in the arrangement of atoms in metallic glasses, while crystals possess orientational and translational long-range order and quasicrystals possess, besides orientational order, nontranslational long-range positional order, which is manifested in the presence of sharp peaks in the Fourier transform of the structure. The similarity of the atomic structures is also indicated by the possibility of transformations between the indicated phases during annealing.[2]

All forms of quasicrystals with icosahedral group symmetry which have been discovered thus far are metal alloys where the atoms form close-packed structures. At the same time, we know of no published reports of the observation of quasicrystalline structures for materials with a covalent interatomic bond, though many such compounds are encountered in both limiting modifications (i.e., in the form of glasses and crystals). An example of a common compound of this kind is silicon dioxide.

On the one hand, the lack of information about covalent quasicrystals could be due to purely technical difficulties in producing them, since even metallic quasicrystals are observed only in exceptional alloys, optimally balanced with respect to the composition and the size of the constituent atoms. Moreover, even for optimal alloys the conditions for the solidification of the melt must be accurately matched, since experiments[1] and numerical simulations[3] both show that quasicrystals are formed over a narrow range of cooling rates, falling between the rates required to obtain glasses and crystals.

On the other hand, up to now it has been unclear whether covalent quasicrystalline atomic structures are possible in principle, or whether such structures do not occur because of the nature of the covalent bond, which imposes quite stringent geometrical constraints on the local atomic packing (specifically, because the extent and relative arrangement of the valence electronic orbitals forming interatomic bonds in covalent solids are quite strictly fixed).

In the present paper we wish to demonstrate that atomic packings with purely covalent bonds, exhibiting all topological features of quasicrystals with icosahedral orientational order, and at the same time satisfying reasonable constraints on local short-range order, are possible in principle.

2. ALGORITHM FOR CONSTRUCTING THE STRUCTURAL MODEL

We construct the model using as a guide the most common glass-forming group IV elements (carbon, silicon, and germanium). In crystals and glasses, the atoms of these elements often possess tetrahedral coordination of the electron clouds and they occupy sites in a network of bonds. These sites can be bonded to one another both directly (in pure materials) and through intermediate divalent atoms (for example, oxygen atoms). Thus, the algorithm for constructing a model of a covalent atomic structure with a prescribed global symmetry reduces to the purely topological problem of describing an appropriate bond network where exactly four edges meet at each site, the angles between neighboring bonds deviating very little from the ideal tetrahedral angle $\varphi_i = 109.5°$ (according to the experimental data and numerical simulation, a deviation of up to 10% is admissible for covalent glasses[4]).

To construct such a network of bonds we employ the principle of duality between the structures of close-packed metals and covalent networks that is observed for amorphous structures.[5] According to this principle, the bond network in a disordered covalent glass (Zachariasen network) is topologically equivalent to the network formed by the edges of Voronoĭ polyhedra for atoms of close-packed metallic

glasses. Similarly, an arbitrary close-packing of atoms that reproduces the model structure of a metallic quasicrystal can be taken as the basic packing, a Voronoĭ division can be constructed for it, and silicon-type atoms can be placed at the sites of the resulting network. If the obviously degenerate cases, which vanish with arbitrarily small deformations of the structure, are discarded, then four edges, which can be identified as interatomic bonds, meet at the sites of the Voronoĭ network. The symmetry of a network of Voronoĭ polyhedra is, by construction, identical to that of the original atomic packing. The only problem is the possibility of satisfying the constraints on the admissible bond lengths and on the interbond angles. The resulting structure can be optimized with respect to bond lengths and interbond angles by relaxing the structure using interatomic interaction potentials, taking account of the three-particle correlations in the positions of the particles. The distributions of bonds over bond lengths and interbond angles, along with the energy of the atomic bonds in an optimized structure, make it possible to assess the feasibility of a practical implementation.

In the present paper we employ as the original packing the atomic packing constructed in Ref. 6 and incorporating all the global properties of quasicrystalline structures (icosahedral orientational symmetry, lack of translational symmetry, self-similarity on different length scales) together with a specific indication of the atomic positions satisfying the requirements of local close-packing. It is well known that in close-packed structures, each atom can be characterized by a number of topological characteristics, specifically, the number Z of nearest neighbors (or the coordination number) and the type of polyhedron whose vertices coincide with the positions of these neighbors (so-called coordination polyhedra). The model employed contains three types of atoms with the coordination numbers 12, 14, and 16 and classical Frank–Kasper coordination polyhedra, which were first introduced to describe complex interatomic alloys.[7] Accordingly, for the Voronoĭ division only three types of Voronoĭ polyhedra arise in the model. We call them $V_{12}^{(1)}$ (dodecahedron), $V_{14}^{(1)}$, and $V_{16}^{(1)}$ cells. The corresponding atomic configurations are displayed in Fig. 1. For each atom three of four interatomic bonds connect the atom to its neighbors in the same cell, while the remaining bonds (shown as dangling bonds) serve to connect the atom to atoms in neighboring cells.

The topological structure of the resulting bond network is fixed by the original atomic structure, and can be described on the basis of an inflation procedure that is an obvious reformulation of the procedure used in Ref. 6. Nonetheless, we examine the first steps of this procedure in greater detail, starting with a 20-atom dodecahedron $V_{12}^{(1)}$. This atomic cluster is closed in the sense that each surface atom has no more than one dangling bond projecting out of the cluster. The attachment of atoms from the next shells to the cluster can be illustrated as the attachment of 12 similar dodecahedra to the faces of this dodecahedron and ''gluing'' the coincident faces. The resulting 110-atom cluster $V_{12}^{(2)}$ is shown in Fig. 2 and is also closed (for clarity, the external dangling bonds are not shown). The second step repeats the first step but starts with the cluster $V_{12}^{(2)}$. Cells of type $V_{16}^{(1)}$ are obtained automatically in 20 ''corners'' with four polyhedra $V_{12}^{(2)}$ meeting at each corner. The resulting structure $V_{12}^{(3)}$ now consists of 1220 atoms and contains several closed clusters, intermediate between it and $V_{12}^{(2)}$. At the third step, 13 $V_{12}^{(3)}$ clusters are joined; here, the $V_{16}^{(2)}$ clusters obtained by inflating a $V_{16}^{(1)}$ cell are positioned at corners where four $V_{12}^{(3)}$ clusters meet. At the same time, $V_{14}^{(1)}$-type cells arise on the glued edges of the clusters.

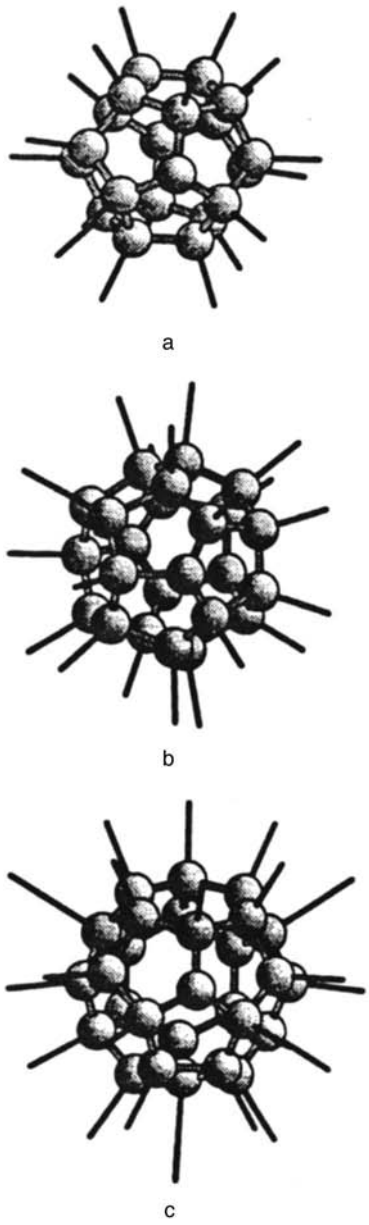

FIG. 1. Atomic configuration of the basic clusters $V_{12}^{(1)}$ (a), $V_{14}^{(1)}$ (b), and $V_{16}^{(1)}$ (c).

This procedure can be repeated an arbitrary number of times. This makes it possible formally to obtain a cluster of any size, consisting essentially of a covalent dodecahedral quasicrystal dual to the original icosahedral metallic quasicrystal.

It is interesting to examine the resulting structure from the standpoint of a dislocation model. As follows from the form of the faces of the $V_{12}^{(1)}$, $V_{14}^{(1)}$, and $V_{16}^{(1)}$ polyhedra, the

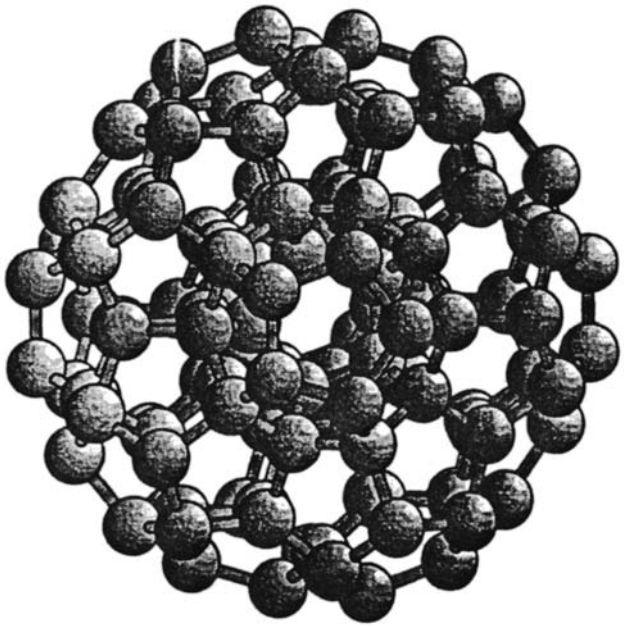

FIG. 2. Atomic configuration of the cluster $V_{12}^{(2)}$.

interatomic bonds in them form either five- or six-atom rings. Let us draw imaginary straight lines (disclinations) through the centers of the six-atom rings up to the point where the lines intersect at the centers of the polyhedra. The resulting linear multiple-bond structure can be characterized as a hierarchical disclination structure of dodecahedral networks. It is easy to check that the resulting disclination structure is topologically equivalent to the one proposed in Refs. 6 and 8. At the same time, our representation of disclinations differs from that proposed in the well-known Ref. 9, where the existence of disclinations (''odd number of lines'') was associated with the presence of rings containing an odd number of sides.

3. SIMULATION OF THE ATOMIC STRUCTURE OF A DODECAHEDRAL QUASICRYSTAL

The proposed algorithm for constructing the atomic structure of dodecahedral quasicrystals is purely topological and contains no information about the lengths and relative orientation of the atomic bonds. Such information is fundamental for assessing the feasibility of producing real covalent quasicrystals. To obtain such information we numerically simulated a dodecahedral cluster obtained from a 439-atom cluster by the duality procedure of Ref. 6. Only certain of the complete set of vertices of Voronoĭ polyhedra were selected, so as to obtain the largest centrosymmetric closed cluster. The resulting cluster contained 1140 atoms and 2120 interatomic bonds, and it was relaxed by the method of conjugate gradients using a three-particle Keating potential,[10] which is often employed for simulating structures with tetrahedral coordination. In the Keating approximation the energy E of an N-atom cluster is given by

$$E=\varepsilon\sum_{n=1}^{N}\left\{\frac{1}{2}\sum_{i=1}^{4}\left[\mathbf{r}_i(n)\cdot\mathbf{r}_i(n)-1\right]^2+\beta\sum_{i=1}^{3}\sum_{j=i+1}^{4}\left[\mathbf{r}_i(n)\cdot\mathbf{r}_j(n)+\frac{1}{3}\right]^2\right\},$$

where the vector $\mathbf{r}_i(n)$ connects the n-th atom with its i-th neighbor and is measured in units of the characteristic interatomic distance d (ordinarily, in the corresponding crystal compound), while ε and β are parameters determined by the specific material. In the present work we employed values typical of silicon: $\varepsilon=12.47$ eV, $d=0.234$ nm, and $\beta=0.29$.[10]

Figure 3a shows the relaxed structure of a model cluster, while Fig. 3b shows the relative intensity of the Fourier components of this structure in a plane perpendicular to one of the five-fold axes. Although the relaxation method employed did not constrain the symmetry to be icosahedral, the relaxation procedure had essentially no effect on the symmetry of a cluster, while the peaks in the Fourier transform indicate long-range positional order. The total energy of a cluster was 189.39 eV or approximately 0.09 eV per bond.

The distributions of the bonds over bond lengths and interbond angles obtained for this structure are presented in Fig. 4. The lengths of the interatomic bonds deviate from the optimal value by $\pm 5\%$, which falls within the admissible variance for disordered covalent structures.[4] The distribution over the angles between neighboring bonds has two distinct maxima—the main peak near 108° and a comparatively low peak near 120°. These peaks correspond to the angles be-

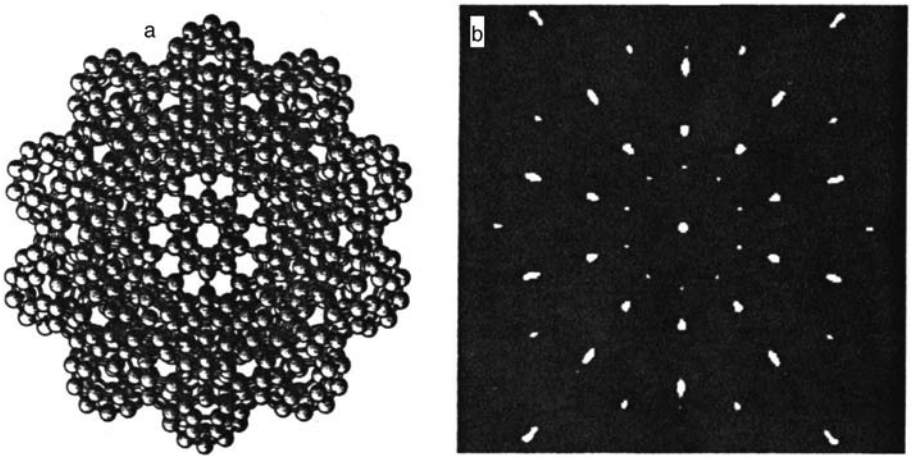

FIG. 3. An 1140-atom covalent cluster with icosahedral symmetry (viewed along the five-fold axis) and the relative intensity of the Fourier components of the structure of such a relaxed cluster in a plane perpendicular to the five-fold axis.

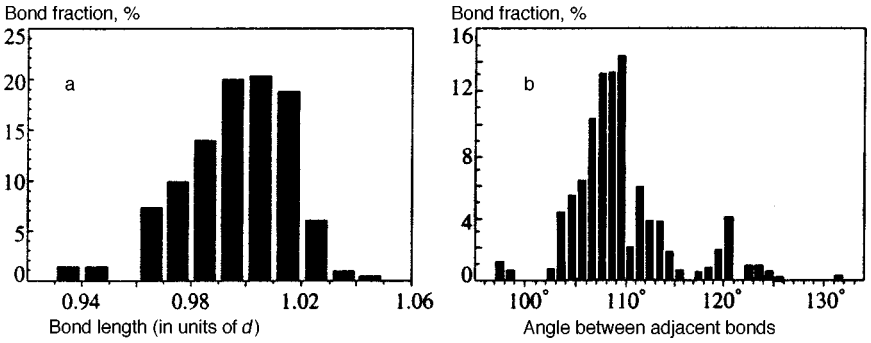

FIG. 4. Bond histograms over bond lengths (a) and interbond angles (b) for a closed 1140-atom cluster relaxed with a Keating potential.

tween the bonds forming five- and six-atom wings. Most interbond angles fall within a reasonable deviation of $\pm 10\%$ from the ideal tetrahedral angle. However, in contrast to the distribution over bond lengths, the distribution over bond angles contains an ''outlier'' at angles $\sim 130°$. Although the relative fraction of such angles is extremely small, they can impose definite constraints on the feasibility of quasicrystals consisting of pure group IV elements. However, in the most complex compounds with bridging atoms on the bonds (of the type SiO_2), where the limits on the angles between neighboring Si-type atoms are much weaker, the observed variance in the angles can hardly be viewed as a serious obstacle to the implementation of quasicrystalline structure.

4. CONCLUSIONS

An algorithm for constructing a quasicrystalline structure consisting of atoms with tetrahedral coordination of covalent atomic bonds was proposed. A computer model of such a cluster with 1140 atoms was constructed. The distributions of interatomic bonds over bond lengths and interbond angles obtained for this structure correlate well with the analogous results for silicon glasses, and do not preclude quasicrystalline structures in practice. This enables us predict the existence of covalent quasicrystals, which to date have not been observed.

This work was supported by Grants No. 97-02-16529 and 98-02-16216 from the Russian Fund for Fundamental Research.

*)E-mail: manichev@dni.polyn.kiae.su

[1] D. Shechtman, I. Blech, D. Gratias, and J. W. Kahn, Phys. Rev. Lett. **53**, 1951 (1984).
[2] D. A. Lilienfeld, H. Nastasi, H. H. Johnson *et al.*, Phys. Rev. Lett. **55**, 1587 (1985).
[3] J. W. Roth, R. Schilling, and H. R. Trebin, Phys. Rev. B **51**, 15833 (1995).
[4] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials*, Clarendon Press, Oxford (1971).
[5] R. Zallen, *The Physics of Amorphous Systems*, Wiley, New York (1983).
[6] V. A. Borodin and V. M. Manichev, Phys. Rev. B **54**, 15747 (1996).
[7] F. C. Frank and J. S. Kasper, Acta Crystallogr. **11**, 184 (1958).
[8] J. F. Sadoc and N. Rivier, Philos. Mag. B **55**, 537 (1987).
[9] N. Rivier, Philos. Mag. A **40**, 859 (1979).
[10] P. N. Keating, Phys. Rev. **145**, 637 (1966).

Translated by M. E. Alferieff

Asymptotic behavior of neutron small-angle multiple scattering spectra

Yu. G. Abov

Moscow State Engineering-Physics Institute, 115409 Moscow, Russia; Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia

D. S. Denisov, N. O. Elyutin, S. K. Matveev, Yu. I. Smirnov, and A. O. Éĭdlin

Moscow State Engineering Physics Institute, 115409 Moscow, Russia

F. S. Dzheparov and D. V. L'vov*)

Institute of Theoretical and Experimental Physics, 117259 Moscow, Russia
(Submitted 22 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 2194–2203 (December 1998)

The asymptotic behavior of the spectra for large values of the scattering vector for the case of elastic multiple small-angle neutron scattering (SANS) is investigated theoretically and experimentally. An expansion of the spectrum in terms of the reciprocal of the magnitude of the momentum transfer is obtained taking account of the influence of the instrumental line. It is shown that, to within some factor, the leading term of the expansion is identical to the differential single-scattering cross section averaged over a statistical ensemble of particles; several subsequent terms in the expansion are calculated and the range of applicability of the resulting expressions is determined. The asymptotic behavior of the multiple SANS spectrum is measured, using a two-crystal neutron spectrometer, for samples of an HTSC ceramic, the alloy Fe–Ni, and Al powder. The agreement between the experimental results and the theoretical predictions is analyzed.
[S1063-7761(98)02312-9]

1. INTRODUCTION

It is well known that the most detailed information can be obtained from elastic small-angle neutron scattering (SANS) data[1,2] if $z \ll L$, where z is the sample thickness and L is the mean free path in the sample. The intensity distribution (spectrum) $I(q)$ of the scattered radiation ($q=|\mathbf{k}-\mathbf{k}_0|$ is the scattering wave number and $\mathbf{k}$ and $\mathbf{k}_0$ are, respectively, the wave vectors of the scattered and incident radiation) is measured over the widest possible range. This makes it possible to determine not only the characteristic size R and density n but in some cases even the shape and details of the internal structure of an isolated scattering center.[1,2]

In studying scattering at very small angles, a two-crystal spectrometer gives the highest resolution because the instrumental line of such a spectrometer is narrow.[1] However, the wings of this line decrease as $1/q^2$ (see, for example, Ref. 3), which is appreciably slower than the Debye–Beuche–Porod $1/q^4$ asymptotic behavior typical of SANS ($1/q^3$ in the present experimental scheme; see below). This makes it necessary to use thick samples ($z>L$) and to take account of multiple-scattering effects. On the other hand, it is often impossible to prepare sufficiently thin samples, for example, when the overall integrity of a sample must be preserved. In this case the scattering is, as a rule, multiple scattering.

In the current practice[4–9] of SANS spectrum analysis with $z \gg L$, ordinarily the functions $\omega(z)$, where ω is the width at half-height of the SANS intensity distributions $I(q,z)$, measured for several samples with various thicknesses z, are used to obtain the parameters characterizing the scattering system. As a rule, the large-q asymptotic behavior of the distributions $I(q,z)$ is ignored, even though the importance of such investigations was pointed out some time ago[9] and the general theory of random walks[10] shows that the asymptotic behavior is similar in the single- and multiple-scattering cases.

Thus, a thorough analysis of large-q asymptotic behavior for multiple scattering has yet to be performed in SANS investigations. The purpose of the present work is to fill this lacuna. In the theoretical part we show on the basis of the results obtained in Refs. 9 and 10 that, to within some factor, the large-q behavior of $I(q,z)$ is identical to the behavior of the differential scattering cross section $\langle \sigma(q) \rangle$ for $q \gg R^{-1}$, where the symbol $\langle \ldots \rangle$ denotes averaging over all possible orientations and deviations from the average shapes and sizes in a statistical ensemble of scatterers. The case in which the wings of the scattering lines are proportional to q^{-4} is studied in greatest detail. As a result, the shapes of the boundaries of the inhomogeneities in samples of the ferromagnetic alloy Fe–Ni, the HTSC ceramic $YBa_2Cu_3O_{7-\delta}$, and Al powder were studied in the experimental part of this work.

2. THEORY

Elastic small-angle neutron scattering is forward scattering, since it results only in a small change in the direction of neutron motion. If the neutron mean free path is much

greater than the wavelength and the sizes of the scatterers, then each subsequent scattering of a neutron can be assumed to occur on a new random configuration of scattering centers. For this reason, neutron propagation is described by the equation

$$\frac{dP(\mathbf{k},z)}{dz}=-\int d^3k'[W(\mathbf{k}',\mathbf{k})P(\mathbf{k},z)-W(\mathbf{k},\mathbf{k}')P(\mathbf{k}',z)],$$

$$P(\mathbf{k},0)=\delta(\mathbf{k}-\mathbf{k}_0), \tag{1}$$

where $P(\mathbf{k},z)$ is the probability density for a neutron to possess momentum $\mathbf{k}$ at depth z from the surface of the sample. Equation (1) contains the transition rates (per unit length) $W(\mathbf{k}',\mathbf{k})$, averaged over the mean free path, from a state with momentum $\mathbf{k}$ into a state with momentum $\mathbf{k}'$. If the material is homogeneous on the average, then $W(\mathbf{k}',\mathbf{k})$ depends only on $\mathbf{k}^2=\mathbf{k}'^2$and the difference $\Delta\mathbf{k}=\mathbf{k}'-\mathbf{k}$. If neutrons are not rescattered, the function $P(\mathbf{k},z)$ decays exponentially

$$\frac{dP(\mathbf{k},z)}{dz}=-\kappa P(\mathbf{k},z),\quad P(\mathbf{k},z)=P(\mathbf{k},0)e^{-\kappa z} \tag{2}$$

at the rate

$$\kappa=\int d^3k'\,W(\mathbf{k}',\mathbf{k})=n\sigma_t,$$

where n is the density of scattering centers and σ_t is the total average scattering cross section of a single center. The neutron mean free path is $L=1/\kappa$.

For small-angle scattering, $\Delta k/k_0\ll 1$. Consequently, it can be assumed that $\Delta\mathbf{k}\perp\mathbf{k}_0$, i.e., the momentum varies only in a plane perpendicular to the direction of motion of the neutron beam. Equation (1) can therefore be written in the two-dimensional form

$$\frac{dP(\boldsymbol{\alpha},z)}{dz}=-\kappa\left(P(\boldsymbol{\alpha},z)-\int\frac{d^2\alpha'}{\sigma_t}\sigma(\boldsymbol{\alpha}-\boldsymbol{\alpha}')P(\boldsymbol{\alpha}',z)\right). \tag{3}$$

Here

$$\boldsymbol{\alpha}=\frac{\mathbf{k}-\mathbf{k}_0}{|\mathbf{k}_0|}\perp\mathbf{k}_0,\quad \sigma_t=\int d^2\alpha\,\sigma(\boldsymbol{\alpha}),$$

and $\sigma(\boldsymbol{\alpha})$ is the average cross section for scattering by the angle $\boldsymbol{\alpha}$ (it is convenient to call the two-dimensional vector $\boldsymbol{\alpha}$ the scattering angle).

Equation (3) is translationally invariant and can be diagonalized by a Fourier transform. Transforming to the Fourier representation, we obtain

$$\frac{dP(\mathbf{u},z)}{dz}=-\kappa\left(P(\mathbf{u},z)-\frac{\sigma(\mathbf{u})}{\sigma_t}P(\mathbf{u},z)\right),\quad P(\mathbf{u},0)=1, \tag{4}$$

where

$$\sigma(\mathbf{u})=\int d^2\alpha\exp(-i\mathbf{u}\cdot\boldsymbol{\alpha})\sigma(\boldsymbol{\alpha})$$

and $\sigma(\mathbf{u})=0=\sigma_t$. Solving Eq. (4), we find the Fourier transform of the propagator

$$P(\mathbf{u},z)=\exp(-\kappa z(\sigma(0)-\sigma(\mathbf{u}))/\sigma_t). \tag{5}$$

Finally, we obtain an expression for the intensity of multiple SANS by the angle $\boldsymbol{\alpha}$ for a sample of thickness z:

$$P(\boldsymbol{\alpha},z)=\frac{1}{(2\pi)^2}\int d^2u\exp(i\mathbf{u}\cdot\boldsymbol{\alpha})\times\exp\left(-\kappa z\frac{\sigma(0)-\sigma(\mathbf{u})}{\sigma_t}\right). \tag{6}$$

Since the system of scatterers is isotropic, the average cross section $\sigma(\boldsymbol{\alpha})$ depends only on $|\boldsymbol{\alpha}|$, and therefore in Eq. (6) the integration over the direction of the vector $\mathbf{u}$ can be performed. The result is the well-known integral representation

$$P(\alpha,z)=\frac{1}{2\pi}\int_0^\infty uJ_0(\alpha u)\exp\left[-\kappa z\left(1-\frac{\sigma(u)}{\sigma_t}\right)\right]du. \tag{7}$$

The two-crystal spectrometer has the feature that the deviations of the neutrons from the initial direction are detected only in one plane (horizontal). That is, in fact, in this method neutrons with different vectors $\mathbf{k}$ with equal projections on the horizontal plane are indistinguishable, and an experiment can only give an average distribution (over the vertical divergence of the beam). Therefore the expression (6) for the intensity $P(\alpha_x,\alpha_y)$ should be integrated over α_y. Since the vertical extent of the detector window is much greater than the characteristic beam width, the integration limits can be taken as infinite (a detailed discussion is given in the experimental part). In addition, to obtain the experimentally observed quantity, the intensity must be convolved with the experimental line. Thus, we find that the normalized experimentally observed intensity of neutrons scattered by an angle α_x in the horizontal plane is

$$P_{exp}(\alpha_x,z)=\int_{-\infty}^{\infty}d\beta\;I_0(\beta)\tilde{P}(\alpha_x-\beta,z), \tag{8}$$

where

$$\tilde{P}(\alpha_x,z)=\int_{-\infty}^{\infty}d\alpha_y P((\alpha_x^2+\alpha_y^2)^{1/2},z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}du_x\exp(iu_x\alpha_x)P(u_x,0,z). \tag{9}$$

Here $I_0(\beta)$ is the instrumental line, and the functions $P(\alpha,z)$ and $P(u_x,u_y,z)=P(\mathbf{u},z)$ are determined by Eqs. (7) and (5), respectively.

If $\alpha_x\gg\max(\alpha_s,\alpha_{ins})$, where the parameters α_s and α_{ins} characterize the widths of the functions $\tilde{P}(\alpha_x)$ and $I_0(\alpha_x)$, respectively, then to estimate $P_{exp}(\alpha_x,z)$ it is sufficient to take account of the fact that the expression in the integrand in Eq. (8) has two sharp peaks—at $\beta=0$ and $\beta=\alpha_x$—and that the functions $I_0(\alpha_x)$ and $\tilde{P}(\alpha_x,z)$ are normalized to 1. Therefore

$$P_{exp}(\alpha_x,z)=I_0(\alpha_x)+\tilde{P}(\alpha_x,z),\quad \alpha_x\gg\max(\alpha_s,\alpha_{ins}). \tag{10}$$

For a more detailed analysis we combine Eqs. (5), (8), and (9) to yield

$$P_{exp}(\alpha_x,z)=\int_{-\infty}^{\infty}\frac{du_x}{2\pi}\exp(iu_x\alpha_x)\times\exp\left(-\kappa z\left(1-\frac{\sigma(|u_x|)}{\sigma_t}\right)\right)g(\alpha_{ins}|u_x|), \tag{11}$$

where $g(\alpha_{ins}|u_x|)$ is the Fourier transform of the instrumental line. A general discussion of the instrumental line can be found in Ref. 3. In practice, it is sufficient to take account of the behavior $g(x)=1-x+o(x)$ in the limit $x\to 0$, which is determined by the asymptotic behavior of $I_0(\alpha_x)$ at large angles:

$$I_0(\alpha_x)=\frac{\alpha_{ins}}{\pi\alpha_x^2}+o\left(\frac{1}{\alpha_x^2}\right).$$

To calculate the asymptotic behavior of the intensity $P_{exp}(\alpha_x,z)$ at large angles it is sufficient to make the substitution $u_x\to u/\alpha_x$

$$P_{exp}(\alpha_x,z)=\frac{1}{\pi\alpha_x}\int_0^{\infty}du\cos u\times\exp\left(-\kappa z\left(1-\frac{\sigma(u/\alpha_x)}{\sigma_t}\right)\right)g\left(\alpha_{ins}\frac{u}{\alpha_x}\right), \tag{12}$$

expand the integrand in powers of the small parameter u/α_x, and carry out the integration.

For scattering by particles with a sharp step boundary (i.e., for $\alpha k_0 l_b\ll 1$, where l_b is the width of the boundary) $\sigma(\alpha)\propto 1/|\boldsymbol{\alpha}|^4$ for $|\boldsymbol{\alpha}|\gg\alpha_c$, where α_c is the characteristic scattering angle.[1,11,12] The cross section for single scattering by a sphere is studied in Ref. 9. However, there is no known expression for the scattering cross section of bodies with an arbitrary shape. It is therefore desirable to study a function of fairly general form that can be used to approximate the single-scattering cross section as

$$\sigma(\boldsymbol{\alpha})=\frac{\sigma_0\alpha_1^2\alpha_2^2}{(\alpha^2+\alpha_1^2)(\alpha^2+\alpha_2^2)}+\sigma_G\exp\left(-\frac{\alpha^2}{a^2}\right), \tag{13}$$

with five degrees of freedom: $\sigma_0,\alpha_1,\alpha_2,\sigma_G$, and a. In order that the function $\sigma(\boldsymbol{\alpha})$ determined by Eq. (13) corresponds to the real scattering cross section, the cross sections can be required to have the same value at $\alpha=0$, the same asymptotic behavior as $a\to\infty$, and the same integrals of $\sigma(\alpha)$, and two additional conditions can be imposed. After Fourier transforming we obtain

$$\sigma(\mathbf{u})=\frac{2\pi\sigma_0\alpha_1^2\alpha_2^2}{\alpha_2^2-\alpha_1^2}(K_0(\alpha_1 u)-K_0(\alpha_2 u))+\pi\sigma_G a^2\exp\left(-\frac{a^2u^2}{4}\right), \tag{14}$$

where $u=|\mathbf{u}|$, and $K_0(x)$ is the modified Bessel function of the second kind. It follows from Eq. (14) that for small u/α_x

$$\sigma_t-\sigma\left(\frac{u}{\alpha_x}\right)=\frac{\pi\sigma_0\alpha_1^2\alpha_2^2}{\alpha_2^2-\alpha_1^2}\left[\frac{(u\alpha_1)^2}{2\alpha_x^2}\ln\frac{u\alpha_1}{2\alpha_x}-\frac{(u\alpha_2)^2}{2\alpha_x^2}\ln\frac{u\alpha_2}{2\alpha_x}\right]+\frac{\pi u^2}{2\alpha_x^2}\left[(1-\gamma)\sigma_0\alpha_1^2\alpha_2^2+\frac{\sigma_G a^4}{2}\right]+\frac{\pi\sigma_0\alpha_1^2\alpha_2^2}{\alpha_2^2-\alpha_1^2}\left[\frac{(u\alpha_1)^4}{32\alpha_x^4}\ln\frac{u\alpha_1}{2\alpha_x}-\frac{(u\alpha_2)^4}{32\alpha_x^4}\ln\frac{u\alpha_2}{2\alpha_x}\right]+\frac{\pi u^4}{32\alpha_x^4}\left[\left(\frac{3}{2}-\gamma\right)\sigma_0\alpha_1^2\alpha_2^2(\alpha_1^2+\alpha_2^2)-\sigma_G a^6\right], \tag{15}$$

where $\gamma=0.577\ldots$ is Euler's constant. Introducing the parameterization $\alpha_0^2=\alpha_1\alpha_2$ and $\xi=\ln(\alpha_1/\alpha_2)$, which is convenient for representing the answer, taking account in Eq. (12) of the first three terms in the expansion of the exponential and two terms in the expansion of $g(u)$, and integrating the result, we obtain

$$P_{exp}(\alpha_x,z)=\frac{\kappa z\alpha_0^2}{8\alpha_x^3}\varphi(\xi)+I_0(\alpha_x)+\frac{\kappa z\alpha_{ins}\alpha_0^2}{4\pi\alpha_x^4}\varphi(\xi)\times\left(3\ln\frac{2\alpha_x}{\alpha_0}-\frac{5}{2}+\frac{\sigma_G a^4}{2\sigma_0\alpha_0^4}\right)-\frac{3\kappa z\alpha_{ins}\alpha_0^2}{4\pi\alpha_x^4}\psi(\xi)+\frac{\kappa^2z^2\alpha_0^4}{16\alpha_x^5}\varphi^2(\xi)\times\left(3\ln\frac{2\alpha_x}{\alpha_0}-\frac{13}{4}+\frac{3\sigma_G a^4}{2\sigma_0\alpha_0^4}\right)-\frac{3\kappa^2z^2\alpha_0^4}{16\alpha_x^5}\varphi(\xi)\psi(\xi)-\frac{3\kappa z\alpha_0^4}{32\alpha_x^5}\varphi(\xi)\psi(\xi), \tag{16}$$

where

$$\varphi(\xi)=\frac{4\sinh\xi}{\xi},\quad \psi(\xi)=2\cosh\xi.$$

Here, the exact expression for $g(u)$ was used to calculate the second term, and terms of order higher than $1/\alpha_x^5$ in α_x were dropped.

Let us consider the case in which the width of the instrumental line is small, i.e., $\alpha_{ins}\ll\alpha_0$. Then the first term in Eq. (16) is the leading term for $\alpha<\alpha_m$, where $\alpha_m=\pi\kappa z\alpha_0^2/2\alpha_{ins}\gg\alpha_0$, i.e., in this region $P_{exp}(\alpha_x,z)$ decays as $1/\alpha_x^3$. The leading term of the asymptotic expansion of the multiple SANS is therefore determined by the behavior of the single-scattering cross section at large angles, which was noted in previous work.[13] The preasymptotic terms in the expansion of $P_{exp}(\alpha_x,z)$ (the third through seventh terms in Eq. (16)) give an increase in the rate of decay of $P_{exp}(\alpha_x,z)$ with increasing sample thickness.

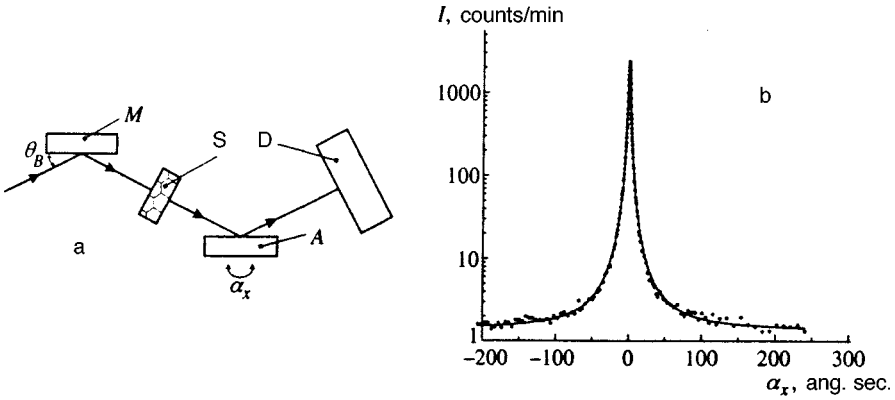

FIG. 1. Two-crystal method for performing small-angle neutron scattering experiments: a) Arrangement of a two-crystal spectrometer based on perfect crystals (M—monochromator crystal, A—analyzer crystal, D—detector, S—sample, θ_B—Bragg angle, α_x—angular deviation of the analyzer from a position parallel to the monochromator); b) instrumental curve of a two-crystal spectrometer (dots), measured without a sample at $\lambda_0 = 1.75$ Å and its least-squares fit by a Lorentzian function.

We shall now examine separately the single-scattering cross section of a sphere in the case of diffraction. In the Born approximation

$$\sigma_B(\alpha) = \sigma_0 \frac{(\sin(\alpha/\alpha_c) - (\alpha/\alpha_c)\cos(\alpha/\alpha_c))^2}{(\alpha/\alpha_c)^6}, \tag{17}$$

where $\alpha_c = 1/k_0 r_0$, k_0 is the initial neutron momentum and r_0 is the interaction radius. To find the behavior of the Fourier transform $\sigma_B(u)$ for small u, we add to and subtract from $\sigma_B(\alpha)$ the function $\sigma_L(\alpha) = \sigma_0 \alpha_c^4/2(\alpha^2+\alpha_c^2)^2$, whose asymptotic behavior is the same as that of $\sigma_B(\alpha)$ averaged over one period of the oscillations. The asymptotic representation of $\sigma_L(u)$ for small u follows from Eq. (15), and $\sigma_B(\alpha) - \sigma_L(\alpha)$ decreases more rapidly than $1/\alpha^4$. The first two terms in the expansion can be found by integrating the terms $\sim u^0$ and $\sim u^2$ in the integrand. Hence we find the behavior of the Fourier transform of the cross section in the Born approximation for small values of u:

$$\sigma_B(u) = \frac{\pi\sigma_0\alpha_c^2}{2} + \frac{\pi\sigma_0\alpha_c^4 u^2}{4} \ln\frac{\alpha_c u}{4}. \tag{18}$$

As in Eq. (16) we obtain

$$P_{exp}(\alpha_x, z) = \frac{\kappa z \alpha_c^2}{2\alpha_x^3} + I_0(\alpha_x) + \frac{\kappa z \alpha_{ins}\alpha_c^2}{2\pi\alpha_x^4}\left(6\ln\frac{4\alpha_x}{\alpha_c} + 6\gamma - 11\right) + \frac{3\kappa^2 z^2 \alpha_c^4}{\alpha_x^5}\left(\ln\frac{4\alpha_x}{\alpha_c} + \gamma - \frac{25}{12}\right). \tag{19}$$

All remarks made for Eq. (16) also hold for this asymptotic representation.

3. EXPERIMENTAL PROCEDURE

The experiments were performed on a universal neutron diffractometer with a doubly monochromated beam.[14] Perfect (111) Ge crystals arranged parallel to one another (see Fig. 1a) were set up as a spectrometric pair. The measured width of the instrumental curve $I_0(\alpha_x)$ (α_x—the rotation angle of the analyzer; $\alpha_x = 0$ for the exact Bragg position of the analyzer with respect to the incident beam) at the working neutron wavelength $\lambda_0 = 1.75$ Å was $\omega_0 = (3.1 \pm 0.2)''$ (see Fig. 1b), which to within the measurement error was identical to the value calculated from the dynamical theory of diffraction.[3] Detection was performed with a ^{3}He detector with a thermal neutron detection efficiency $\varepsilon \approx 96\%$; neutrons in the highest reflection orders with $\lambda_n = \lambda_0/n$ were filtered using a 100-mm thick, single-crystal quartz filter, which reduced their contribution to the beam to less than 1% of the main component. The samples S were set up in the space between the monochromator crystal M and the analyzer crystal A (see Fig. 1a). The intensities of the neutrons reflected from the crystal A were measured within the angular range outside of which the detector D detected only background counts. The minimum possible rotation step of the analyzer was $0.12''$.

In the expression (9) the integration over α_y is performed over infinite limits, i.e., it is assumed that the height H of the detector window is great enough that the entire scattered beam falls within the window. In real experiments the maximum angle is

$$\alpha_y^0 = \tan^{-1}(H/2s) \simeq H/2s, \tag{20}$$

where s is the sample–detector distance. In our diffractometer $H = 40$ mm and $s = 1200$ mm, so that $\alpha_y^0 = 0.017$ rad $= 57'$, which is typical of devices of this type. For all samples investigated, the widths of the multiple SANS spectra did not exceed $\omega_{\max} \approx 2' \ll \alpha_y^0$, i.e., we can set $\alpha_y^0 = \infty$. It should be noted that such a situation almost always occurs when perfect crystals are used in a two-crystal spectrometer, while for crystals with a mosaic structure the angular range of the small-angle spectra can exceed α_y^0, which introduces additional corrections in the asymptotic representations of the SANS intensity.

4. EXPERIMENTAL RESULTS AND DISCUSSION

The instrumental curve of the two-crystal spectrometer with perfect crystals is displayed in Fig. 1b. Approximation of this curve by a Lorentzian function shows that its asymptotic behavior is described by $1/\alpha_x^2$, as assumed in the theoretical part of this work. Figure 3 displays the angular dependence of the intensity in the wings of the instrumental curve. Clearly, in this range of angles $I_0(\alpha_x)$ is much smaller than $P_{exp}(\alpha_x)$ (by approximately an order of magnitude).

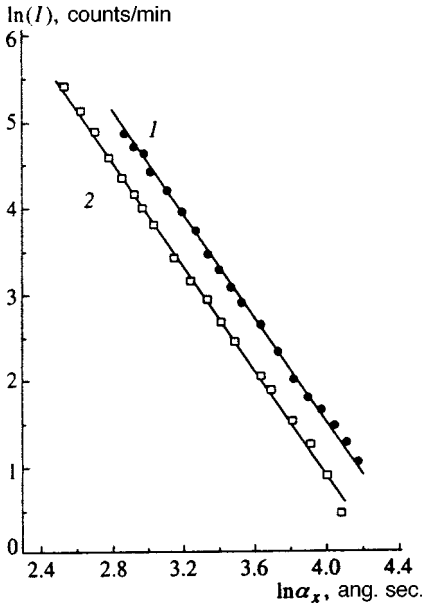

FIG. 2. Asymptotic behavior of the intensity I of scattered neutrons as a function of rotation angle α_x of the analyzer crystal (δ is the exponent): *1)* Al powder sample with average granule size $2R=9.5\,\mu$m and thickness $z=5.5$ mm, $\delta=3.02$ (0.03); *2)* HTSC ceramic sample, $z=1.69$ mm, average pore size 3 μm, $\delta=3.04$ (0.03). The solid line is a least-squares fit.

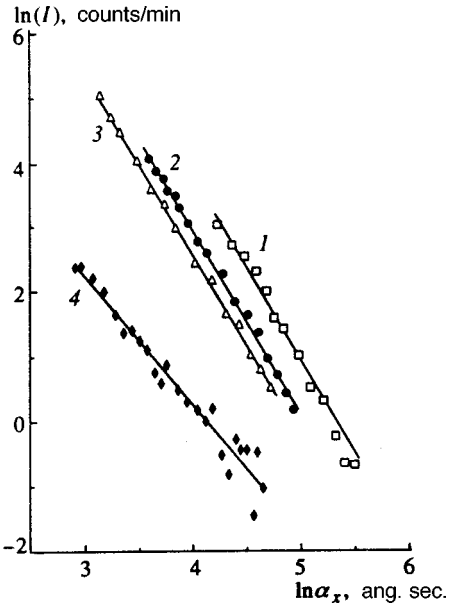

FIG. 3. Asymptotic behavior of the intensity I of scattered neutrons as a function of rotation angle α_x of the analyzer crystal for samples of the alloy Fe–Ni on a logarithmic scale (δ is the exponent): *1)* Sample thickness $z=7.0(0.1)$ mm, $\delta=2.93(0.03)$; *2)* sample thickness $z=2.0(0.1)$ mm, $\delta=2.89(0.03)$; *3)* sample thickness $z=1.0(0.1)$ mm, $\delta=2.85(0.03)$; *4)* instrumental line. The solid line is a least-squares fit.

If scattering occurs by solid compact particles with uniform density and a sharp boundary, the following relation, obtained by Debye and Bueche[11] and Porod[12] (see also Ref. 1), holds for the single-scattering intensity:

$$I^{(1)}(q)=\frac{2\pi}{q^4}\rho^2 S, \tag{21}$$

or $\sigma(q)\propto q^{-4}$, which holds for inhomogeneities of arbitrary shape for large values of q. Here ρ is the neutron-optical density, determined by the composition of the scatterers and their host medium, and S is the surface area of a particle. The asymptotic representation $I^{(1)}(q)\propto q^{-4}$ also materializes in some magnets near phase transitions.[15] If a sharp boundary does not exist between the medium and an inhomogeneity, the intensity decays more rapidly than q^{-4}. A slower dependence than Eq. (21) is possible near phase transitions[9] as well as for scatterers with a fractal surface.[16,17] On the basis of the theoretical analysis made above, investigations of the asymptotic behavior of the multiple SANS intensities should give similar information.

Figure 2 displays on a logarithmic scale the results of the measurements performed on a sample of the HTSC ceramic $YBa_2Cu_3O_{7-\delta}$, which is characterized by the presence of pores about 3 μm in size,[8] as well as the experimental scattering data for a sample of Al powder with thickness $z=5.5$ mm and average granule size $2R\approx 9.5\,\mu$m, determined in Ref. 7. The solid line in the figure shows the result of a simultaneous (least-squares) fit of the experimental data by the function $c_1/\alpha_x^2+c_2/\alpha_x^{\delta}$ and of the instrumental curve by the function c_1/α_x^2. The asymptotic behavior of the second term, corresponding to the sample, is clearly well-described by the law α_x^{-3}, which agrees with Eqs. (16) and (19). The small-angle neutron scattering by the same HTSC ceramic samples was investigated in Ref. 17, and $\sigma(q)\propto q^{-n}$, where $n=3.85\pm 0.1$, was obtained for the cross section. Our data, obtained with a larger statistical sample, suggest more convincingly that the scattering takes place at compact inhomogeneities with a sharp boundary.

Figure 3 shows (on a logarithmic scale, as in Fig. 2) the results of experiments performed on samples of the alloy Fe–Ni—a ferromagnet characterized by the presence of a disordered domain structure—covering a range of thickness. Using the procedure described in Ref. 6, the domains were determined to be about 8 μm in size. A simultaneous least-squares fit of the experimental data and the instrumental line using Eq. (16) (with $\xi=1$ and neglecting the Gaussian term) gives a large rms deviation ($\chi^2/(n-m-1)\approx 390/64$, with a total of $n=71$ experimental points and $m=6$ adjustable parameters). Here the errors for the experimental points were calculated taking account of the fact that, according to Ref. 18, the statistics of the readings in our apparatus deviates somewhat from Poisson statistics, and the variance is $\sigma^2=1.4\overline{I}$, where $\overline{I}$ is the average intensity (for fixed α_x). The known thickness ratios were included in the fit as two additional points with their own errors. A simultaneous fit of the data by functions of the form $c_1/\alpha_x^2+c_2/\alpha_x^{\delta}$, where the same value of δ is used for different samples and the ratio of the coefficients c_2 for different samples equals the thickness ratio, gives $\delta=2.91(0.04)$ with $\chi^2/(n-m-1)\approx 145/65$. A satisfactory result $\chi^2/(n-m-1)\approx 76/63$ is obtained only with a function $c_1/\alpha_x^2+c_2/\alpha_x^{\delta}$ of a more general form, where δ has different values for three samples, and for all samples $\delta<3$. The exponent δ increases with sample thickness. This agrees with Eqs. (16) and (19) and corresponds to the contribution of the preasymptotic terms in the large-angle

expansion. However, exponents $\delta<3$ are not obtained for independent scattering of neutrons by compact inhomogeneities with a sharp boundary. The observed difference likewise cannot be explained by the finite thickness of the domain walls, which causes δ to increase.

5. CONCLUSIONS

A method for analyzing inhomogeneities in a medium by measuring the large-q asymptotic behavior of the multiple SANS intensity $I(q,z)$ was studied. A two-crystal neutron spectrometer was used to investigate samples of a HTSC ceramic with the composition $YBa_2Cu_3O_{7-\delta}$, the ferromagnetic alloy Fe–Ni, and Al powder. The experimental data for the HTSC ceramic and Al are described well by the asymptotic approximation α_x^{-3}, which corresponds to the Debye–Beuche–Porod law ($I\propto q^{-4}$). Hence it follows that the inhomogeneities in these substances have nonfractal step boundaries. A similar investigation for samples of the ferromagnet Fe–Ni showed a slower than α_x^{-3} decrease of the SANS intensity.

In summary, the theoretical analysis performed above describes satisfactorily the experimentally observed behavior of the intensity of multiple scattering for the Al and HTSC ceramic samples, but it does not explain the slower than α_x^{-3} decrease of the intensity for the Fe–Ni samples. This is probably due to the fact that, as follows from the experimental data, the ratio of the neutron mean free path to the scatterer (domain) size is much smaller for Fe–Ni samples than for the Al and HTSC ceramic samples. A description of scattering by Fe–Ni samples should therefore be based on equations that take account of the correlation between successive neutron scattering events.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-17771-a) and the American Foundation ''Fundamental Nuclear Physics'' (Grant No. 1.3.5.-01).

*)E-mail: stepanov@vitep5.itep.ru

[1] D. I. Svergun and L. A. Feĭgin, *Small–Angle X-Ray and Neutron Scattering* [in Russian], Nauka, Moscow (1986).
[2] Yu. M. Ostanevich and I. N. Serdyuk, Usp. Fiz. Nauk **137**, 85 (1982) [Sov. Phys. Usp. **25**, 323 (1982)].
[3] Z. G. Pinsker, *X-Ray Crystal Optics* [in Russian], Nauka, Moscow (1982).
[4] D. L. Dexter and W. W. Beeman, Phys. Rev. **76**, 1782 (1949).
[5] R. J. Weiss, Phys. Rev. **83**, 380 (1951).
[6] S. Sh. Shil'shteĭn, V. A. Somenkov, M. Kalanov, and N. O. Elyutin, Fiz. Tverd. Tela (Leningrad) **18**, 3231 (1976) [Sov. Phys. Solid State **18**, 1886 (1976)].
[7] Yu. I. Smirnov, A. O. Pekshev, A. O. Éĭdlin *et al.*, Fiz. Tverd. Tela (Leningrad) **33**, 2273 (1991) [Sov. Phys. Solid State **33**, 1281 (1991)].
[8] Yu. G. Abov, Yu. I. Smirnov, D. S. Denisov *et al.*, Fiz. Tverd. Tela (Leningrad) **34**, 1408 (1992) [Sov. Phys. Solid State **34**, 748 (1992)].
[9] S. V. Maleev and B. P. Toperverg, Zh. Éksp. Teor. Fiz. **78**, 315 (1980) [Sov. Phys. JETP **51**, 158 (1980)].
[10] F. S. Dzheparov, Zh. Éksp. Teor. Fiz. **99**, 982 (1991) [Sov. Phys. JETP **72**, 546 (1991)].
[11] P. Debye and A. M. J. Bueche, J. Appl. Phys. **20**, 518 (1949).
[12] G. Porod, Kolloid-Z. **124**, 83 (1951).
[13] S. V. Maleyev, R. V. Pomortsev, and Yu. N. Skryabin, Phys. Rev. B **50**, 7133 (1994).
[14] Yu. G. Abov, N. O. Elyutin, D. S. Denisov *et al.*, Prib. Tekh. Éksp. No. 6, 67 (1994).
[15] H. Glaettly, in *Lectures at the 26th Winter School, Institute of Experimental and Theoretical Physics*, Moscow (1998).
[16] É. Z. Valiev, S. G. Bogdanov, Yu. A. Dorofeev *et al.*, Zh. Éksp. Teor. Fiz. **100**, 1000 (1991) [Sov. Phys. JETP **73**, 552 (1991)].
[17] A. I. Okorokov, V. V. Runov, A. D. Tret'yakov *et al.*, Zh. Éksp. Teor. Fiz. **100**, 257 (1991) [Sov. Phys. JETP **73**, 143 (1991)].
[18] A. V. Krasnikov, D. S. Denisov, N. O. Elyutin *et al.*, Prib. Tekh. Éksp. No. 2, 82 (1991).

Translated by M. E. Alferieff

Long-lived excited state of Te donors in GaP

Ya. E. Pokrovskiĭ,*) O. I. Smirnova, and N. A. Khval'kovskiĭ

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia
(Submitted 3 June 1998)
Zh. Éksp. Teor. Fiz. **114**, 2204–2210 (December 1998)

The kinetics of the photoresponses in constant and microwave electric fields and the variation of the absorption of background radiation in GaP doped with Te (2×10^{17} cm^{-3}) upon impurity excitation at 5–50 K are investigated. The lifetime of the excited state of the Te donors is determined ($\sim10^{-2}$ s). It is shown that the results presented are consistent with the model of carrier accumulation in long-lived impurity excited states in semiconductors. These results are compared with the results previously obtained for diamond-structure semiconductors. © *1998 American Institute of Physics.* [S1063-7761(98)02412-3]

1. INTRODUCTION

It has been established in a series of studies (see Ref. 1 and the references therein) that some donors and acceptors from groups V and III, respectively, in diamond, silicon, and germanium have excited states, in which the charge-carrier lifetime τ^* is many orders of magnitude greater than the free-carrier lifetime τ. These long lifetimes are due to the complex band structure of diamond-structure semiconductors, which causes valley-orbit or spin-orbit splitting of the ground states of the donors and acceptors. Because of the identical parity of these split-off lowest excited and ground states of the impurities, electric-dipole optical transitions between them are forbidden, and the probability of multiphonon transitions is low. The accumulation of carriers in long-lived excited states leads to several characteristic phenomena, particularly to the dominance of percolation hopping photoconduction at a constant bias voltage (diamond) or polarization hopping photoconduction in a microwave electric field (silicon and germanium). It should be expected that long-lived excited states also exist in other semiconductors with a complex band structure. In particular, in the indirect-gap III–V semiconductor gallium phosphide the ground state of Te donors is split by the valley-orbit interaction (see the inset in Fig. 1), and, therefore, as in silicon, the lowest excited state of Te should be long-lived.

In fact, the results obtained in Ref. 2 were recently explained by the existence of a long-lived excited state for Te donors in GaP. Pulsed impurity breakdown in the electric field of high-power electromagnetic waves of long-wavelength coherent radiation was used to excite the samples in Ref. 2. After a certain delay ($\sim0.5\mu$s) following the cessation of breakdown, a constant-voltage photoresponse appeared, and its relaxation time constant increased from 10^{-5} to 3×10^{-3} s as the temperature was lowered from 150 to 35 K. The phenomena observed were attributed to the accumulation of electrons in the long-lived ($\tau^*=7$ ms) $1S(\Gamma_3)$ excited state of Te. The temperature dependence of the slow relaxation of the photoresponse was attributed by Ganichev *et al.* to thermal electron exchange between the long-lived state and the higher-lying $2S(\Gamma_1)$ state, through which both transitions to the ground state and the thermal ejection of electrons into the conduction band take place. The kinetic equations proposed in Ref. 2 describe the experimental results quite well when a small number of adjustment parameters are used. However, an investigation of the constant-voltage photoresponse was possible only at relatively high temperatures, since the response appeared as a result of the thermal ionization of a long-lived state, and, therefore, the lifetime τ^* was determined in Ref. 2 as an adjustment parameter.

In the present work relaxation of the excitation of Te donors in GaP was investigated by measuring the photoconductivity in a microwave electric field, as well as the induced absorption of background radiation. We previously used these methods to prove the existence of long-lived excited states of donors and acceptors in diamond-structure semiconductors. This allowed us to extend the temperature range down to 5 K without resorting to extrapolation, to determine the lifetime τ^* of the excited state, and to study its properties in greater detail.

2. EXPERIMENTAL RESULTS

Samples of Te-doped GaP cut from wafers with a thickness of 0.25–0.35 mm were investigated. The contacts for the constant-voltage investigations were fabricated by the vacuum evaporation of indium followed by dissolution of the surface layer of GaP in the liquid metal,[3] and they were satisfactory over the entire temperature range. Figure 1 presents the temperature dependence of the resistance R of one of the samples. At 300 K, $R=18\Omega$, which, with consideration of the sample geometry and a mobility of 120 cm^2/V·s (Ref. 4), corresponds to a Te concentration $N_d=2\times10^{17}$ cm^{-3}. When the temperature was lowered, the value of R increased exponentially with an activation energy of 80 meV. This value agrees quite well with the energy determined in Ref. 4 from the temperature dependence of the Hall constant for a sample with a similar concentration, but a small ionization energy $E_i=90$ meV, which was determined from optical experiments[5] (see the inset in Fig. 1). The concentration of the compensating impurity in the samples in-

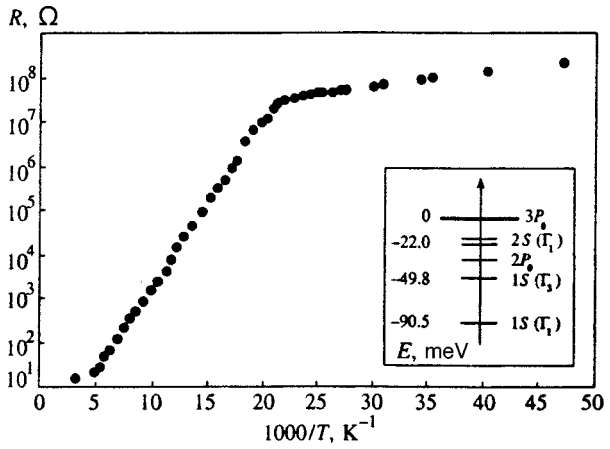

FIG. 1. Temperature dependence of the resistance R of a sample of Te-doped GaP. Inset — energies of the ground and several excited states of Te donors in GaP.

vestigated ($N_a \sim 10^{16}$ cm^{-3}) was determined from the high-temperature portion of the plot of R and a comparison of the latter with the dependence for a similar sample in Ref. 4. A segment exhibiting an exponential increase in the resistance with an activation energy of 7.5 meV, which does not depend on the intensity of the room-temperature background illumination, was observed at $T<35$ K. The resistance is determined along this segment by the hopping conductivity, and the value of the activation energy[6] corresponds to $N_d = 2\times10^{17}$ cm^{-3} when $N_a/N_d \ll 1$.

The photoconductivity of the GaP samples was investigated at constant (dc) and microwave (mcw) bias voltages. In the latter case samples measuring $1\times0.35\times0.03$ cm^3 filled the width of an 8-mm waveguide at an antinode of the electric field of the wave near its end face, which was closed off by a grid. The samples were excited through the grid by the output of a laser with a wavelength equal to 3.39 μm and a power of about 5 mW, which was modulated at the frequency f. The microwave radiation from a Gunn diode which passed through the sample and was reflected from the closed end was fed through a circulator into a detector, which isolated the component modulated at the frequency f. This signal is proportional to the photoresponse of the sample and was recorded by a phase-sensitive nanovoltmeter.

Figure 2 shows the dependence of photoresponse $U_{\rm mcw}$ on the modulation frequency f for one of the GaP samples at 5 K. It is seen that the response decreases by almost three orders of magnitude as f is raised from 8 Hz to 10 kHz. Unfortunately, because of the weak absorption in thin samples, the photoconductivity kinetics could not be investigated at large values of f. Therefore, we were unable to achieve conditions under which the response would no longer depend on f and would be determined by free photoelectrons, as in silicon.[1] The frequency dependence (Fig. 2) corresponds to relaxation of the photoresponse with a time constant $\tau^* = 8$ ms, which essentially coincides with the value obtained in Ref. 2 as an adjustment parameter. The constant-bias photoconductivity under the same conditions did not depend on f at least up to 10 kHz.

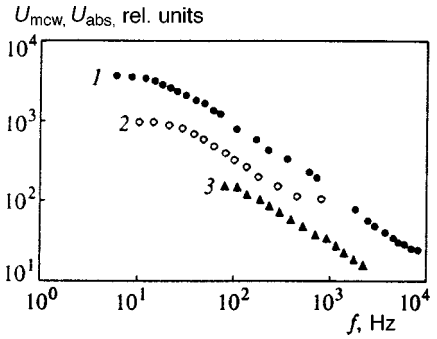

FIG. 2. Dependence of the response of a sample of Te-doped GaP on the modulation frequency of the exciting radiation f at 5 K: *1* — $U_{\rm mcw}$; *2* — $U_{\rm abs}$ with detection by Si⟨B⟩; *3* — $U_{\rm abs}$ with detection by Ge⟨Hg⟩.

Plots of the temperature dependence of the quasistationary photoresponses $U_{\rm mcw}$ and $U_{\rm dc}$ for the same sample are presented in Fig. 3. It is seen that at low temperatures $U_{\rm mcw}$ exceeds $U_{\rm dc}$ by almost three orders of magnitude. As the temperature is raised to 16 K, $U_{\rm mcw}$ slowly decreases and then drops sharply, and at $T=35$ K it is comparable to $U_{\rm dc}$. The temperature dependences of $U_{\rm mcw}$ and $U_{\rm dc}$ for Te in GaP are similar to the dependences obtained for doped silicon and germanium.[1] These dependences are attributed to the fact that $U_{\rm dc}$ is determined by free charge carriers with a short lifetime, while $U_{\rm mcw}$ is determined by the polarization hopping photoconductivity resulting from the accumulation of carriers in the long-lived excited states of the impurities.[1]

The accumulation of charge carriers in long-lived excited states in silicon leads to the appearance of absorption bands at photon energies which are smaller than the ionization energies of the impurity ground state and correspond to the photoionization of long-lived excited states.[1] However, in the case of GaP the measurements in the spectral region of photon energies smaller than 90 meV are made difficult by

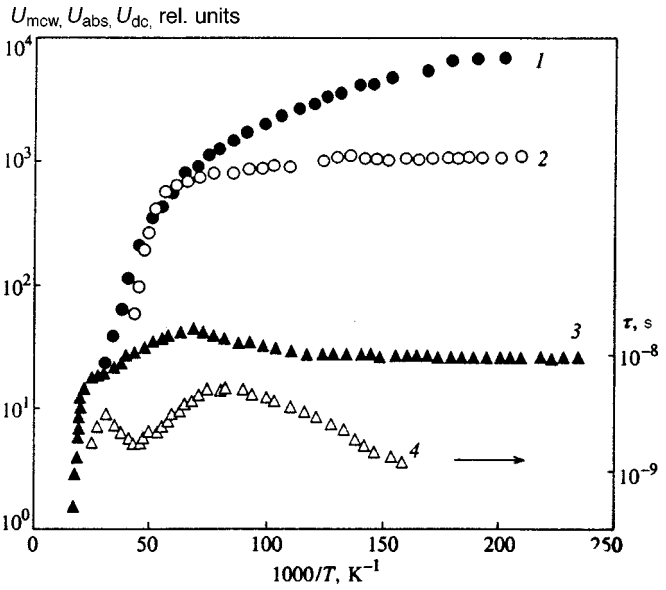

FIG. 3. Temperature dependences at $f=12$ Hz: *1* — $U_{\rm mcw}$; *2* — $U_{\rm abs}$ with detection by Si⟨B⟩; *3* — $U_{\rm dc}$; *4* — lifetime τ of holes in silicon doped with B ($N_a = 3.5\times10^{16}$ cm^{-3}) and P ($N_d = 2\times10^{13}$ cm^{-3}).

the absence of a series of strong bands of lattice absorption.[7] Therefore, we confined ourselves to an investigation of the kinetics of the absorption of the room-temperature background radiation induced by the modulated impurity photoexcitation of GaP samples. As in Ref. 8, the sample was placed in a cryostat. The radiation passed through its cold silicon window and then through the sample and impinged on an impurity photoresistor located beyond the sample. The sample and the photoresistor were separated by a polyethylene film, which virtually completely absorbed the modulated laser radiation with a wavelength of 3.39 μm used to excite the sample. The photoresistor detected the component of the background radiation modulated as a result of the change in absorption caused by redistribution of the charge carriers between the Te ground and excited states upon photoexcitation. The absorption associated with ionization of the Te excited state ($\sim$50 meV) was detected by a boron-doped silicon Si$\langle$B$\rangle$ photoresistor, which is sensitive in the region 45–100 meV. The absorption due to ionization of the ground state (90 meV) was detected by a mercury-doped germanium Ge$\langle$Hg$\rangle$ photoresistor with a long-wavelength cutoff at 90 meV.

Figure 2 presents plots of the dependence of the response U_{abs} of the photoresistors on the modulation frequency of the exciting radiation f. It is seen that the responses caused by the induced absorption of background radiation decrease with increasing f, i.e., are also determined by long-term processes. The response of the Ge$\langle$Hg$\rangle$ photoresistor was in phase, while the response of the Si$\langle$B$\rangle$ photoresistor was in antiphase to the exciting radiation. Hence it follows that photoexcitation leads to an increase in the population of the excited state and a decrease in the population of the ground state. Unfortunately, despite the very high sensitivity of the Ge$\langle$Hg$\rangle$ photoresistor, we were unable to reliably detect the photoresponse both at high values of f because of the small amplitude of the signal and at low values of f because of the appearance of $1/f$ noise. Therefore, we shall refrain from estimating the relaxation time of this response. Owing to the high intensity of the background in the long-wavelength portion of the spectrum, the response to the Si$\langle$B$\rangle$ photoresistor was recorded with sufficient confidence. The relaxation time estimated from Fig. 2 (5 ms) is close to the relaxation time of U_{mcw}.

Figure 3 presents the temperature dependence of the amplitude U_{abs} of the induced absorption at $f=12$ Hz normalized to the sensitivity of the Si$\langle$B$\rangle$ photoresistor. It is seen that the amplitude of this absorption scarcely depends on temperature up to 16 K. This is the temperature at which the rapid drop in U_{mcw} begins.

3. DISCUSSION OF RESULTS

Our interpretation is based on several results obtained in investigations of diamond-structure semiconductors,[1] as well as on the model of polarization hopping photoconduction in a microwave electric field.[9]

In doped and compensated semiconductors polarization conductivity appears in a microwave electric field as a consequence of the change in the dipole moment of pairs of ionized atoms of the principal and compensating impurities induced by the field. It can be many orders of magnitude greater than the dc conductivity.[10] In the case of impurity excitation, polarization photoconductivity appears as a result of the hopping of nonequilibrium charge carriers between excited atoms with large orbitals and ionized atoms of the principal impurity. The polarization photoconductivity relaxation time is determined by the lifetime of the impurity excited states τ^* (Ref. 9). It follows from Fig. 2 that the value of τ^* for Te in GaP at 5 K is 8 ms.

In our experiments the frequency of the microwave field (40 GHz) was so high that only complexes of ionized and excited atoms separated by minimal distances can contribute to the photoconductivity. It was shown in Ref. 9 that such active complexes, which consist of ions of the principal and compensating impurities and an excited impurity atom, form as a result of the trapping of free carriers on dipoles consisting of ions of the principal and compensating impurities and the hopping drift of ionized ions of the principal impurity toward the dipoles. Thermal dissociation of the dipoles leads to a decrease in the values of U_{mcw} in relatively weakly compensated semiconductors as the temperature is raised.

The temperature dependences of U_{mcw} and U_{dc} upon the steady-state photoexcitation of GaP in Fig. 3 are similar to those obtained for silicon and germanium.[1] The decrease in the response in a microwave field as the temperature is raised is naturally associated with thermal dissociation of the dipoles. At higher temperatures there is a rapid drop in U_{mcw} due to thermal emptying of the long-lived excited states.

The temperature dependence of U_{dc} is determined by the lifetime of the free electrons τ. At $T<10$ K, U_{dc} scarcely depends on temperature, since τ is determined in this range by the electron energy relaxation time. At $T>45$ K there is a rapid drop in U_{dc} due to thermal ionization of the Te ground state. In the intermediate temperature range there is a region with an increase in $U_{dc}\propto\tau$ as T is lowered. Populating of the long-lived Te state, whose thermal ionization leads to an increase in the lifetime τ of free electrons, occurs due to photoexcitation in just this temperature range. This conclusion is confirmed by the results which we previously obtained in a direct investigation of the temperature dependence of τ in silicon doped with B and compensated by P. The value of τ was determined from the phase shift between U_{mcw} and the exciting radiation.[11] The temperature dependence of τ for one of the silicon samples is presented in Fig. 3. It is seen that τ increases in the temperature range 22–14 K, in which populating of the long-lived excited state of boron impurities in silicon occurs.[1]

The kinetics and temperature dependence of the absorption of background radiation induced by modulated photoexcitation are consistent with the conclusion that nonequilibrium electrons accumulate in a long-lived excited state of Te in GaP. At $T<16$ K the absorption (Fig. 3) scarcely depends on temperature, since the transition rate of electrons from the excited state to the ground state exceeds the rate of thermal emptying of the excited state and the concentration of excited impurity atoms does not depend on T. A sharp drop in U_{abs} occurs in the same temperature range as the drop in U_{mcw} because of thermal emptying of the excited state. We

note that the amplitude of the induced absorption depends on the total concentration of excited impurity atoms regardless of whether these atoms are isolated or localized near the compensating ions.

It was established in the Ref. 1 in the case of silicon that the relaxation time of $U_{\rm mcw}$ is determined by the lifetime of the excited state and depends on the chemical nature not only of the principal impurity, but also of the compensating impurity, whose ions appear in the active complexes, and therefore can differ from the lifetime of isolated excited atoms. Therefore, the relaxation kinetics of $U_{\rm mcw}$ and $U_{\rm abs}$ can differ. In fact, $U_{\rm abs}$ has a somewhat shorter relaxation time than $U_{\rm mcw}$ (Fig. 2). This difference can be attributed to the fact that the influence of the proximity of compensating impurity atoms on the lifetime of excited impurity atoms is displayed in GaP, as well as in silicon.

As we have already noted, the signal $U_{\rm abs}$, which we associated with emptying of the Te ground state, was observed in the short-wavelength region for the absorption of background radiation (Fig. 2). The slow relaxation of this signal ($\tau \geqslant 2$ ms) cannot be associated with the luminescence reported in Ref. 2. The time dependences of the intensity of the luminescence signal and the exciting pulse in Ref. 2 were identical. Therefore, it is possible that the signal is caused not by luminescence, but by emptying of the Te ground state during the ionization of Te by the high-power short ($\sim 0.25 \mu$s) pulse and the resultant increase in the transmission of background radiation, which was detected by the photoresistor.

4. CONCLUSION

In conclusion, we note that the investigation of the highly long-lived excited state of Te in GaP confirms the conclusion drawn in Ref. 1 that such states exist in a broad class of semiconductors with a complex band structure. Manifestations of long-lived impurity states are possible for different semiconductors in various ranges of temperatures and concentrations of principal and compensating impurities and for different types of equilibrium conduction.

We thank S. D. Ganichev for supplying the GaP samples. This work was carried out with financial support from the Russian Foundation for Basic Research (Grant No. 96-02-16243) and the Program for Supporting Leading Scientific Schools in Russia (Grant No. 96-15-96701).

*)E-mail: yaep@mail.cplire.ru

[1] Ya. E. Pokrovskiĭ, O. I. Smirnova, and N. A. Khval'kovskiĭ, Zh. Éksp. Teor. Fiz. **112**, 221 (1997) [JETP **85**, 121 (1997)].
[2] S. D. Ganichev, E. Zepezauer, W. Raab, I. N. Yassievich, and W. Prettl, Phys. Rev. B **55**, 9243 (1997).
[3] Yu. A. Gol'berg, E. A. Posse, and B. V. Tsarenkov, Fiz. Tekh. Poluprovodn. **20**, 1510 (1986) [Sov. Phys. Semicond. **20**, 947 (1986)].
[4] H. C. Montgomery, J. Appl. Phys. **39**, 2002 (1968).
[5] W. Scott, J. Appl. Phys. **50**, 472 (1979).
[6] I. Shklovskiĭ and I. Ya. Yanchev, Fiz. Tekh. Poluprovodn. **6**, 1616 (1972) [Sov. Phys. Semicond. **6**, 1395 (1973)].
[7] A. Kleinman and W. G. Spitzer, Phys. Rev. **118**, 110 (1960).
[8] Ya. E. Pokrovskiĭ and O. I. Smirnova, JETP Lett. **54**, 97 (1991)].
[9] Ya. E. Pokrovskiĭ and O. I. Smirnova, Zh. Éksp. Teor. Fiz. **103**, 1411 (1993) [JETP **76**, 690 (1993)].
[10] M. Pollak and T. M. Geballe, Phys. Rev. **122**, 1742 (1961).
[11] I. V. Altukhov, Ya. E. Pokrovskiĭ, V. P. Sinis, and O. I. Smirnova, Fiz. Tekh. Poluprovodn. **24**, 1134 (1990) [Sov. Phys. Semicond. **24**, 717 (1990)].

Translated and edited by P. Shelnitz

Features of the optical absorption of crystals of the fullerene C_{60} in the region of the orientational phase transition

A. N. Izotov, V. V. Kveder,*) Yu. A. Osip'yan, É. A. Steinman,†) R. K. Nikolaev, and N. S. Sidorov

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

(Submitted 6 July 1998)
Zh. Éksp. Teor. Fiz. **114**, 2211–2224 (December 1998)

The absorption coefficient of perfect single crystals of the fullerene C_{60} is measured in the energy range 1.6–2.1 eV at temperatures from 4.2 to 300 K. An absorption fine structure is discovered in the and is assigned to electronic and vibronic transitions with the production of free excitons and excitons localized on structural defects. It is shown that in the region of the structural phase transition from a face-centered cubic structure to a simple cubic structure the absorption coefficient undergoes a jump, which is associated with an energy shift of the free exciton line toward lower energies. It is discovered that spatial inhomogeneity, which is associated with the growth of the new phase from a finite number of nuclei, appears in the crystal at the time of this transition. © *1998 American Institute of Physics.* [S1063-7761(98)02512-8]

1. INTRODUCTION

Molecules of C_{60} are bound to one another in a crystal mainly by van der Waals interactions, and in this sense crystals of C_{60} can be classified as typical molecular crystals. At the same time, the electronic properties of C_{60} crystals place them in the borderline region between ''classical'' molecular crystals (such as naphthalene) and ''classical'' semiconductors and have thus aroused great interest in them.

The high symmetry and nearly spherical shape of the C_{60} molecule, as well as the weak intermolecular interaction, lead to the interesting features of these crystals associated with orientational ordering of the molecules. According to numerous x-ray diffraction and neutron scattering data (see, for example, Ref. 1 and the references therein), at temperatures above 260 K the centers of the molecules form a face-centered cubic (fcc) lattice of $Fm3m$ symmetry. Since the symmetry of the C_{60} molecule is incompatible with the $Fm3m$ symmetry corresponding to the fcc structure of the crystal, this finding indicates the presence of static or dynamic disorder in the orientation of the molecules (but not in the positions of their centers), which provides for an averaged fcc symmetry. According to the presently available data, at temperatures above 260 K, C_{60} molecules undergo rapid [with a frequency of the order of 10^9 Hz (Ref. 2)] rotation about all of their axes, which also provides for fcc symmetry. X-ray measurements have shown that when the temperature is lowered to 250–260 K, the free rotation of the molecules ceases and definite correlations between the orientations of neighboring molecules are established.

This event takes place in the form of a phase transition with an enthalpy change ΔH of the order of 8–9.5 J/g (i.e., about 0.11 eV per molecule).[3–5] Although the latent heat of the transition is low, it indicates that it is a first-order transition. The positions of the centers of the molecules change only slightly in the process (the lattice constant decreases by 0.3% from $a=1.4154$ nm to $a=1.4111$ nm),[6,7] but the partial orientational ordering of the molecules leads to strong lowering of the symmetry of the crystal to $Pa3$, and the lattice becomes a simple cubic one with a basis consisting of four differently oriented molecules.[1,7,8]

The reason for this phase transition is the existence of the off-center part of the intermolecular interaction. The energy minimum corresponds to an orientation in which a region of increased electron density (a double bond joining pentagons) of one molecule is positioned opposite a region of reduced electron density [a pentagon center (*PF*) or a hexagon center (*HF*)] of a neighboring molecule.

At temperatures below the orientational phase transition ($T_c \approx 259$ K) neighboring molecules are oriented in one of these two configurations, i.e., *PF* or *HF*. The *PF* configuration is more advantageous, but the energy corresponding to the *HF* configuration is greater than that of the *PF* configuration by only $\Delta G=10$ meV.[8,9] Therefore, some orientational disorder remains in the crystal at temperatures below T_c: some of the molecules have *HF* configurations, and some of the molecules have *PF* configurations. The transition from one configuration to another requires overcoming an energy barrier of the order of $\Delta E_{PH}=0.25-0.3$ eV. In the range $T_c>T>0$ the molecules undergo thermally activated transitions from one configuration to the other. As a result, the population of the *PF* configuration p increases with decreasing temperature as

$$p=1/(1+\exp(-\Delta G/kT)), \tag{1}$$

where ΔG is of the order of 10 meV (see Refs. 8 and 9). However, since the barrier height ΔE_{PH} is fairly high, the frequency of $HF-PF$ transitions decreases rapidly with the temperature and becomes negligibly small at temperatures below 85–90 K. Therefore, at temperatures below 85–90 K the population of the PF configuration ceases to follow for-

mula (1) for kinetic reasons and is ''frozen'' at a level of the order of 85%. This phase has been termed an ''orientational glass.''

Despite the large number of x-ray, neutron diffraction, and calorimetric investigations, some details of the orientational ordering in C_{60} crystals have not yet been elucidated. In particular, the nature of some anomalies in the temperature dependences of the intensities of x-ray reflections[8] is unclear, the correlation radius in the ''orientational glass'' phase (i.e., whether it is, in fact, a glass) is unclear, and the reasons why the temperature of the orientational phase transitions varies fairly strongly (by several degrees) from sample to sample and the transition width sometimes reaches 3–5 degrees are not understood (see, for example Refs. 4 and 10–12). Many investigators attribute the broadening and shift of the transition to the presence of impurities. Of course, in some cases this is actually true. However, according to our data, appreciable displacement and ''broadening'' of the transition with the temperature can be observed even in single crystals grown in a vacuum from a high-purity raw material.

Since the orientational phase transition can be accompanied by appreciable changes in the overlap of the electronic wave functions of neighboring molecules, it should have a strong effect on the electronic properties of C_{60} molecules, especially on those which are associated with the presence of electron bands and the transport of electrons and excitons. In fact, according to the data in Ref. 13, the photoconductivity edge of C_{60} crystals corresponds to an energy of 1.7 eV at $150<T<250$ K, while at $T>260$ K (the temperature of the orientational phase transition) the photoconductivity peak begins to rise at 1.65 eV. However, this occurs smoothly without a threshold; therefore, it is unclear whether it is associated with a phase transition or is a chance coincidence. In Ref. 11 an anomalously sharp (10%) decrease in absorption was observed in crystalline C_{60} films in the energy range 2.4–2.8 eV as the temperature was increased in the range 230–260 K. The authors attribute this effect to the fact that the optical transitions in this temperature range are forbidden in an isolated molecule and become partially allowed as a result of the distortion of the symmetry of the molecules, which increases sharply below the phase-transition point. However, these measurements were performed on films whose structure is generally imperfect. In addition, only optical transitions at high energies were investigated. Therefore, it would be interesting to investigate the behavior of lower excited states on perfect single crystal samples.

The purpose of the present work is to investigate the behavior of the optical absorption edge in perfect single crystals of C_{60} in the region of the orientational phase transition.

2. SAMPLES AND EXPERIMENTAL METHOD

Single crystals of C_{60} were grown in the Institute of Solid-State Physics of the Russian Academy of Sciences by physical vapor transport (sublimation) in an evacuated ampul at temperatures from 600 to 640 °C. After chromatographic purification and removal of any traces of the solvent by repeated vacuum sublimation, the starting material C_{60} had a purity no poorer than 99.98%. Single crystals in the form of thin plane-parallel plates with the natural habit were selected for the measurements. The facet planes correspond to (111). The thickness of the samples ranged from 0.01 to 0.3 mm, and the other two dimensions were usually of the order of 1–2 mm.

Since the diffusion coefficient of oxygen in a C_{60} crystal at room temperature is of the order of 10^{-13} $cm^2 \cdot s^{-1}$,[14] holding a sample in air for 24 h can lead to saturation of a surface layer of the order of 10^{-3} mm with oxygen and can (in principle) influence the correctness of the measurements. Therefore, after the growth ampul was opened, the samples were placed fairly quickly in an optical cryostat for measurements and were subsequently kept in a vacuum for the bulk of the time.

The optical cryostat specially fabricated for these experiments has the form of a flat cylindrical vacuum chamber with quartz windows (on the top and the bottom), in which there was a copper heat exchanger cooled by a stream of cold helium (or nitrogen) passing through it. The sample investigated is placed on a horizontally oriented plane-parallel sapphire plate, which is in good thermal contact with the heat exchanger. The temperature of the sapphire substrate is measured by a thermocouple. To improve the thermal contact between the thermocouple and the substrate, and junctions are pressed down by a piece of aluminum foil glued to the substrate. The temperature scan rate during the measurements was usually 0.7–1.5 K/min. Since the thickness of the samples was small (of the order of 0.2 mm) and they were in close proximity to the substrate, according to our estimates, the difference between the sample and substrate temperatures did not exceed 0.5 K even without gluing the sample to the sapphire substrate. Nevertheless, in some measurements the sample was placed on a drop of GKZh silicone oil to improve the thermal contact between the sample and the substrate. Since this oil remains liquid at the measurement temperature, it does not create additional elastic stresses in the sample.

For the measurements of the local transmission spectra of a sample, light from a halogen lamp passed through an MDR-12 monochromator and was directed through the sample by means of a rotating mirror. The image of the sample in the transmitted monochromatic light was focused using an MBS-4 microscope onto a silicon CCD camera (Samsung CM120P or PIN756 LILIN), which has a fairly high sensitivity in the energy range 1.2–2.7 eV. The video signal from the CCD camera was fed through a suitable interface card into a computer, in which it was recorded and processed using specially written programs. The data from the thermocouple were also recorded by a computer. Such a system enabled us not only to obtain a series of photographs of the sample in transmitted monochromatic light, but also to calculate the dependence of the local absorption coefficient of the sample on photon energy and temperature by processing the images. Since the CCD camera had automatic sensitivity control, the absorption coefficient K was calculated by comparing the amplitude of the video signal V from a given region of the sample with the signal amplitude V_0 from a

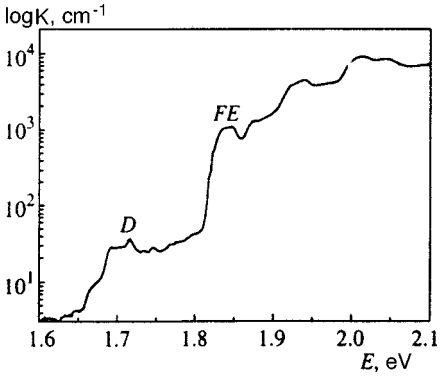

FIG. 1. Dependence of log$K$ on the photon energy E, where K is the absorption coefficient of the C_{60} crystal obtained at $T=10$ K.

region outside the sample on the same video card according to the simple formula

$$K=-\log(V/V_{0R})/d. \tag{2}$$

Here d is the sample thickness, and $V_{0R}=V_0/(1-R)^2$, where R is the reflection coefficient of the sample, which was estimated using the standard formula $R=(n-1)^2/(n+1)^2$ with a value of n of the order of 2.1 (Ref. 14) (n is the refractive index of the sample). This formula does not allow for the repeated reflection of light from the sample faces or the energy dependence of the reflection coefficient. Since we were working in the region of fairly high absorption (low transmission of the sample), the corrections associated with the reflection of light could not have a significant effect on the measured value of K, and there was no need to more carefully take into account the effects associated with reflection. Moreover, we were interested mainly in the variation of the absorption coefficient, rather than its absolute value. For measurements of the absorption spectra in the region of high absorption coefficients, a cooled FÉU-62 photomultiplier was used instead of the CCD camera to detect the light. In this case to protect the sample from scattered light, a large opaque mask with a small opening, significantly smaller than the dimensions of the sample in the plane of the mask, was clamped tightly against it. Samples of different thickness grown in a single batch were used to more accurately measure the absorption in the entire range investigated.

3. EXPERIMENTAL RESULTS AND DISCUSSION

We investigated nine samples from different batches grown at somewhat different temperatures. The results obtained on different samples qualitatively coincide. Before proceeding to a discussion of the influence of the temperature and the phase transition on the absorption spectra of C_{60} crystals, let us discuss the features of the absorption spectra of these crystals measured at low temperatures.

Figure 1 shows a typical absorption spectrum [$K(E)$ curve] of a C_{60} crystal obtained at $T>100$ K. Since the absorption coefficient varies by several orders of magnitude in the energy range 1.7–2.1 eV, the logarithm of the absorption coefficient is plotted in the figure. The spectrum shown in Fig. 1 is the result of several measurements on single-crystal plates of various thickness grown in a single ampul. The

TABLE I. Positions of the absorption lines (E), their FWHM (ΔE), and their amplitude (A) obtained from the absorption spectra as a result of decomposition into Gaussian lines. The substructure lines were obtained from a spectrum measured at 10 K, and the envelopes were obtained from spectra measured at $T>100$ K.

	Envelope			Susbstructure		
N	E, eV	ΔE, meV	A, cm^{-1}	E, eV	A, cm^{-1}	ΔE, meV
				1.671	7	6
$D1$	1.71	30	30	1.693	9	21
				1.716	10	32
				1.746	12	20
$D2$	1.78	30	25	1.778	14	17
				1.802	14	20
				1.825	428	7.5
$F1$	1.84	15	1005	1.836	799	7.3
				1.85	929	7.6
				1.872	1161	8.3
$F2$	1.88	16	1091	1.887	1198	8
				1.902	1269	8
				1.919	2923	9.5
$F3$	1.93	25	3870	1.931	2889	8
				1.943	3611	8
				1.956	2889	8.5
$F4$	1.97	19	2268	1.969	3119	8.5
				1.983	3340	8.5
				2.002	9028	10
$F5$	2.01	20	6494	–	–	–
				–	–	–
$F6$	2.05	22	5253	–	–	–

absorption spectrum of the single crystal has a complicated structure consisting of a large number of overlapping lines. Table I lists the parameters of the lines obtained by decomposing the low-temperature spectra into Gaussian lines, as well as the position and width of the envelope contours used to investigate the temperature dependence of unresolved groups of lines at $T>100$ K.

The absorption spectrum shown in Fig. 1 can be divided into two regions, which are labeled by D and FE in the figure.

In the region 1.65–1.8 eV (D) there is fairly weak absorption ($K<100$ cm^{-1}). We assume that this absorption is caused by intrinsic defects in the crystal corresponding to deep electron (and exciton) traps. The D absorption has a clearly expressed structure and consists of several overlapping lines, whose positions are reproduced very well in different samples. However, the intensities of these lines differ strongly in different samples, varying from 10 to 100 cm^{-1}. The nature of the defects discussed is still not clear. It has not been ruled out that these defects contain pairs of molecules joined by weak covalent bonds.

The FE absorption in the energy range from 1.82 to 2.1 eV is reproduced well from sample to sample both with respect to the absorption intensity and with respect to its structure. We assume that this absorption corresponds to the excitation of free singlet excitons.

The sharp increase in absorption[11] beginning at 2.2–2.25 eV (it is not shown in the figure) has a somewhat different nature and is associated, in our opinion, with the excitation

of electronic transitions from the valence band to the conduction band of the crystal (i.e., with charge-transfer electronic transitions).

This interpretation correlates well with the diverse data from other investigators, particularly with the results of measurements of luminescence photoexcitation spectra. It was shown in Refs. 15 and 16 that the numerous lines in the photoluminescence spectra can be separated into at least two classes: a family of lines with a photoexcitation edge at 1.65 eV and a family of lines with an excitation edge at 1.815 eV. The excitation spectra of photoluminescence lines from the first group contain several overlapping lines, whose position coincides with the position of the *D* absorption lines at 1.65–1.8 eV. Thus, the first group of photoluminescence lines corresponds to the excitons in deep defects producing the *D* absorption.

The edge of the excitation spectrum of the second group of photoluminescence lines coincides with the edge of ''strong'' absorption of the crystal beginning at 1.815 eV (the *FE* region). This group of photoluminescence lines includes a fairly weak line at 1.815 eV and a series of lines with smaller energies, among which the most intense are the lines at 1.724 eV and 1.636 eV. The fact that the position of the photoluminescence lines with an energy of 1.815 eV essentially coincides with the edge of the ''strong'' absorption allows us to assume that this line corresponds to pure 0–0 electronic transitions accompanying the recombination of free excitons (or excitons in very shallow traps). The next lines in this series are vibronic repeats of the first line, i.e., electronic transitions accompanied by the excitation of intramolecular vibrational modes. Under this assumption the most intense photoluminescence lines, with energies of 1.724 and 1.636 eV, correspond to the excitation of the well known h_u (90 meV) and t_{1u} (178 meV) molecular vibrational modes, in complete agreement with the calculations in Ref. 17.

It can be assumed on this basis that the absorption band with an edge specified by an energy of 1.815 eV (and a maximum with an energy 1.84 eV) also corresponds to pure 0–0 electronic transitions with the excitation of free singlet excitons. Then the two most intense *FE* absorption bands with maxima corresponding to energies of 1.93 and 2 eV can be interpreted as vibronic repeats of the 0–0 band, i.e., as the same electronic transitions accompanied by the excitation of the h_u (90 meV) and t_{1u} (178 meV) molecular modes.

It can clearly be seen on the low-temperature absorption spectrum that the band with a maximum corresponding to an energy of 1.84 eV has a weakly resolved structure, which is faithfully reproduced in different crystals. The decomposition of this band into Gaussian lines showed that it can be represented as the result of the superposition of three lines, whose positions and intensities are listed in Table I. It is noteworthy that the bands of the vibronic repeats are also described well by the superposition of three lines with the same intervals between them, which, however, are shifted in energy by the value of the corresponding molecular vibration. The following arguments can be advanced in regard to the possible interpretation of these lines. The first line with an energy of 1.825 eV can be assigned to the 0–0 transition, and the other two lines, which are displaced relative to the first line by 10 and 25 meV, respectively, are phonon repeats of the first line with the participation of crystal phonons. However, this interpretation is unlikely, since the energies of the corresponding crystal vibrations determined by inelastic neutron scattering measurements are significantly smaller.[18–20] Another possible explanation can be associated with a feature of the behavior of the density of states in the exciton band of molecular crystals.[21,22] In this case the structure of the bands with an energy of 1.84 eV can reflect the structure of the density of states in the exciton band.

Figure 1 shows that apart from the three *FE* absorption bands cited there is another series of three bands of lower intensity with energies equal to 1.88, 1.97, and 2.05 eV, which are displaced relative to the first series of bands by roughly 40–45 meV. The interpretation of this series is less obvious. Generally speaking, suitable vibrational modes, which can explain the presence of the absorption lines under discussion, can be found among the rich set of vibrational modes of the model. In particular, according to Ref. 17, the h_u vibration (430 cm^{-1}) can produce a vibronic band. However, the following alternative explanation is also possible.

The exciton band of the crystal is formed from the lowest unoccupied molecular orbitals (LUMO's) of C_{60}. According to theoretical calculations (see, for example, Refs. 17,19, and 23–25), the C_{60} molecule has three LUMO levels of singlet electronic excitations, whose symmetry correspond to the T_{1g}, T_{2g}, and G_g irreducible representations of the I_h symmetry point group. They all lie roughly 2 eV above the A_g ground state (the HOMO). The calculated energetic positions of these levels are very close, and they can produce a series of absorption lines shifted relative to one another. The results of calculations employing the ARGUS program (see Ref. 24) for the T_{1g} and T_{2g} levels correspond to energies of 1.86 and 1.90 eV, while the values according to Ref. 25 are 2.10 and 2.17 eV. Although the accuracy of the calculations is not very high, the order of magnitude of the gap between the levels (40–70 meV) correlates well with the observed difference of 48 meV between the series of absorption lines. Assuming that the first group of lines corresponds to transitions to T_{1g}, we can, in principle, assign the second series of lines to the corresponding transitions to T_{2g}.

The hypothesis that the sharp increase in absorption at 2.2–2.3 eV corresponds to electronic charge-transfer bands (i.e., the excitation of electrons from the valence band into the conduction band) is confirmed by the data in Refs. 13 and 26 from photoconductivity measurements at low temperatures ($T<100$ K) and measurements of the influence of an electric field on the photoluminescence excitation efficiency. It was shown that at low temperatures a sharp increase in photoconductivity begins at an energy of 2.3 eV and that at just that energy of the exciting light an electric field begins to lead to a decrease in the photoluminescence intensity. This can be attributed to separation of the photoexcited electrons and holes. Thus, the bottom of the zone of singlet excitons is roughly 0.4 eV below the bottom of the conduction band.

Let us now move on to the temperature features of the absorption spectra. Below we shall discuss, for the most part,

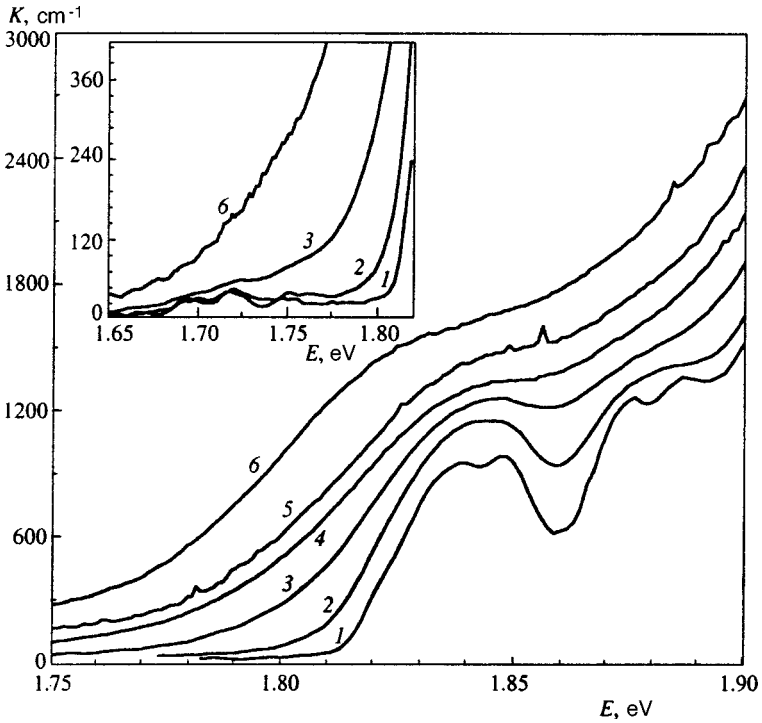

FIG. 2. Absorption spectra of a C_{60} single crystal obtained at various temperatures: *1* — 19, *2* — 108, *3* — 192, *4* — 245, *5* — 256, *6* — 260 K. The inset shows the low-energy part of the absorption spectra.

the behavior of the absorption band with a maximum near 1.84 eV, which can presumably be assigned to pure electronic transitions at the bottom of the band of singlet excitons. Figure 2 shows some absorption spectra of one of the single crystals measured as the sample temperature was raised from 19 to 300 K.

In order to analyze the temperature dependence obtained in greater detail, we used a computer expansion of the absorption spectra shown in Fig. 2 into three Gaussian lines (with energies of the order of 1.84, 1.88, and 1.93 eV). In fact, as was discussed above, it can be seen on curve *1* that each of these lines has a poorly resolved internal structure, which becomes less pronounced as the temperature is raised. In the range of interest to us, $T>70$ K, this structure becomes so diffuse that the spectra are described well by three Gaussian lines, and a more detailed analysis would be inappropriate.

The behavior of the full-width at half maximum (FWHM) and the energetic positions of the line corresponding to an energy of 1.84 eV, which specifies the exciton absorption edge, is shown in Fig. 3. It is seen that, for the most part, raising the temperature from 50 to 255 K leads to broadening of the absorption line without significant displacement, while the orientational phase transition leads to appreciable displacement of the line. The nature of the increased spread of the data in the region of the phase transitions is associated with the spatial inhomogeneity of the sample, which will be discussed below.

The width and temperature of the phase transition are conveniently analyzed by measuring the temperature dependence of the absorption of the sample at a fixed wavelength. Figure 4 shows the temperature dependence of the absorption of light with an energy of 1.81 eV obtained as the sample temperature is raised. The curve presented corresponds to the averaged absorption over an area of the sample of the order of 1 mm^2.

As the temperature is raised, a sharp increase in absorption is observed at 260 K, which corresponds to the temperature at which the orientational phase transition occurs according to the x-ray data. The observed width of the transition is of the order of 2 K.

Below 260 K the $K(T)$ curve has a smooth character, but it clearly exhibits a change in slope at 95 K and a hump at 150–180 K. The former feature is probably associated with the transition to an ''orientational glass'' caused by ''freezing'' of the orientational jumps of the molecules between the *PF* and *HF* orientations. The nature of the feature at 150–180 K is still unclear. We note that the x-ray data[7] also point to the existence of some peculiar anomalies in this region. In particular, the intensity of the (7 10 0) reflection, which should appear immediately after the orientational phase tran-

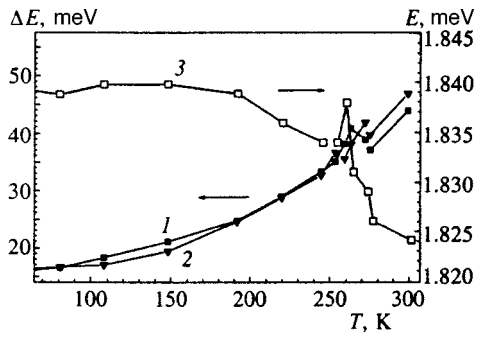

FIG. 3. Temperature dependence of the FWHM (curve *1*) and the position (curve *3*) of the line with an energy of 1.84 eV, which specifies the absorption band edge. Curve *2* — FWHM of the band specified by an energy of 1.88 eV.

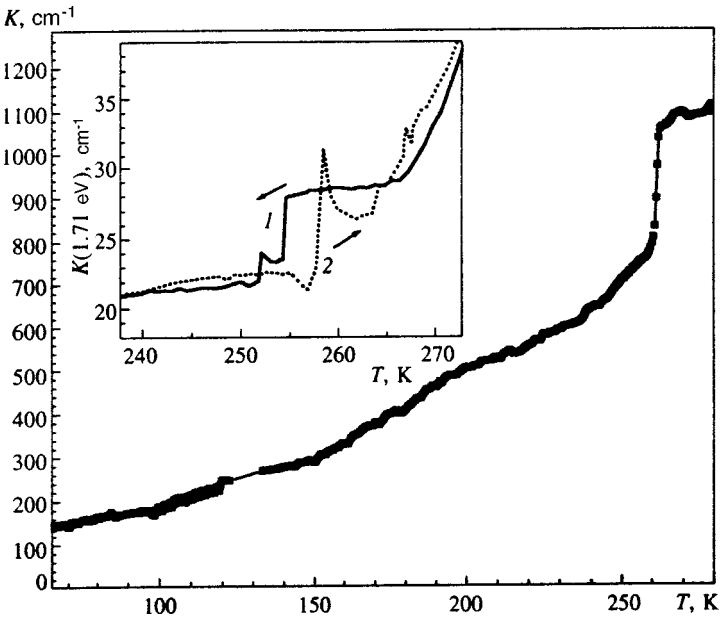

FIG. 4. Temperature dependence of the integrated absorption coefficient for light with an energy of 1.81 eV in a C_{60} single crystal recorded as the sample temperature is raised. The inset shows the temperature dependence of the absorption of light with an energy of 1.71 eV in a local region of the sample measuring 0.15×0.15 mm^2. Curve *1* was obtained as the temperature was lowered, and curve *2* was obtained as it was raised.

sition, is anomalously weak at $T > 180$ K for some reason, and it appears only at 180–170 K.

The inset in Fig. 4 shows the temperature dependence of the absorption of light with an energy of 1.71 eV in a local region of the sample measuring 0.15×0.15 mm^2. Curve *1* was observed as the temperature was lowered, and curve *2* was obtained as it was raised. It can be seen, first, that there is significant hysteresis and, second, that the absorption varies nonmonotonically and there is a series of features. The presence of hysteresis can be regarded as evidence that the orientational transition is a first-order phase transition. The nonmonotonic character of the variation of the absorption during the transition is due to the fact that the sample becomes inhomogeneous at the time of the transition.

These inhomogeneities are clearly visible in Figs. 5 and 6, which present photographs of C_{60} single crystals obtained in monochromatic transmitted light with an energy of 1.71 ± 0.01 eV as the sample temperature is lowered. Both figures correspond to the first cooling cycle of the samples after they were grown. An investigation of the samples in polarized light with crossed polarizers showed that there are no elastic stresses in the original samples and that they cannot be the cause of the observed inhomogeneity of the phase transition. We attribute the inhomogeneities observed during the phase transition to the fact that the new orientationally ordered phase probably grows from nuclei which are sparse in the original, fairly perfect crystal.

The observations in polarized light showed that the phase transition is accompanied by the appearance of elastic stresses in the sample, which subsequently remain even after the sample is warmed to room temperature. This, however, does not lead to an increase in the inhomogeneity of the phase transition. Conversely, the phase transition occurring during repeated cooling of the sample is more homogeneous. Already after two or three cooling-heating (290–150–290 K) cycles, the spatial dimensions of the inhomogeneities observed during the phase transition become less than the thickness of the crystal, and essentially uniform variation of the optical absorption is observed upon subsequent passages through the temperature of the orientational transition. Annealing one of these thermally cycled samples in a vacuum at 200 °C for 2 h greatly reduced the elastic stresses, but the variation of the optical absorption during the phase transition became even more uniform.

These results can be attributed to the fact that the elastic stresses appearing during an inhomogeneous phase transition lead to plastic deformation of the sample and the appearance of dislocations. On the one hand, these phenomena lead to the appearance of elastic stresses in the sample that persist when it is warmed to room temperature. On the other hand, dislocations apparently facilitate the formation of nuclei during the phase transition. As is shown in Fig. 6, when the cooling rate is sufficiently high, the elastic stresses appearing during the phase transition probably do not manage to relax by means of plastic deformation, and microcracks appear in the sample.

In conclusion, it should be noted that the value which we obtained for the absorption coefficient in the energy range 1.9–2 eV differs from the value presented in the literature.[1] This difference arises, in part, because the literature data generally refer to room temperature, at which the absorption is several times greater than the absorption at low temperatures due to the influence of the absorption tails from allowed transitions (see, for example, Fig. 2). However, this phenomenon does not fully account for the difference indicated. In our opinion, the main difference between our data and the literature data is due to the fact that our measurements were performed on highly perfect single crystals, while large absorption coefficients were obtained in measurements on thin films having a large density of grain boundaries. Local perturbations at these boundaries can lead to partial removal of the prohibition against HOMO-LUMO dipole-dipole transitions due to distortion of the symmetry. In fact, the value of 5×10^4 cm^{-1} (Ref. 27) for the absorption coefficient at 1.9 eV can correspond to allowed transitions, rather than forbidden transitions. Thus, we assume that

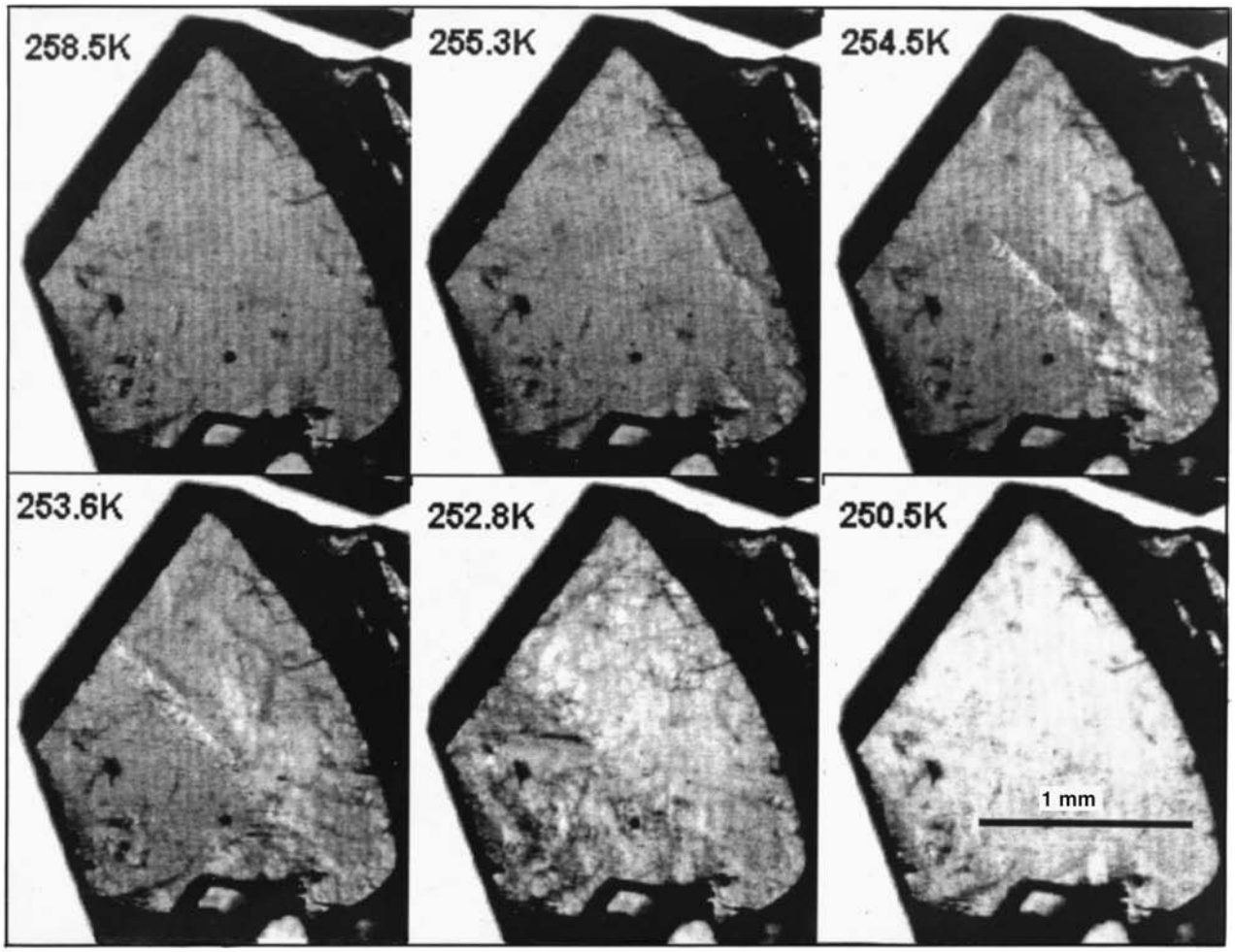

FIG. 5. Photographs of a freshly grown C_{60} single crystal in transmitted light with an energy of 1.71 eV taken as the temperature was lowered at the rate of 0.02 K/s. The sample thickness is approximately 0.2 mm. Six frames taken at various temperatures are shown.

the values which we obtained correspond to absorption in a perfect C_{60} crystal.

4. CONCLUSION

The behavior of the absorption band edge in the region for the excitation of singlet excitons in pure perfect single crystals of C_{60} has been in investigated in the present work. An absorption fine structure associated with the production of both free Frenkel excitons and excitons localized on structural defects has been discovered. It has been shown that the orientational phase transition in C_{60} crystals is accompanied by displacement of the exciton absorption edge, which corresponds to a change in the position of the bottom of the exciton band by roughly 15 meV. In all likelihood, this is due to alteration of the overlap of the electronic wave functions of neighboring molecules during the phase transition, which leads to reorganization of the band structure of the crystal.

It has been shown that the orientational phase transition in sufficiently perfect C_{60} single crystals is strongly inhomogeneous and is observed in the form of ''spreading'' of the new phase from a small number of nuclei. During this process, internal stresses appear in the sample, which lead to the appearance of defects in the crystal (most probably, dislocations). When the concentration of defects in the crystal increases, the phase transition in it becomes quasihomogeneous, probably due to the formation of nuclei on the defects.

At 85–95 K there is an abrupt change in the slope of the temperature dependence of the optical absorption coefficient in the region of the excitonic absorption edge of the crystal. This feature clearly corresponds to a transition of the sample to an ''orientational glass.'' It has been discovered that there is also a feature on the temperature dependence of the absorption coefficient in the temperature range 150–180 K, which is probably associated with a change in the relative orientation of the molecules, whose nature is still unclear.

It follows from the results obtained that even if the sample is a highly perfect single crystal in the sense of the positions of the centers of the molecules at $T>T_c$, it can become polycrystalline in the sense of orientational order at temperatures below the orientational transition. Moreover, it is known that orientational disorder occurs because of a thermodynamic nonequilibrium concentration of *HF* orientations below 80–90 K. All this has a strong influence on the electronic properties of the crystals investigated and must be

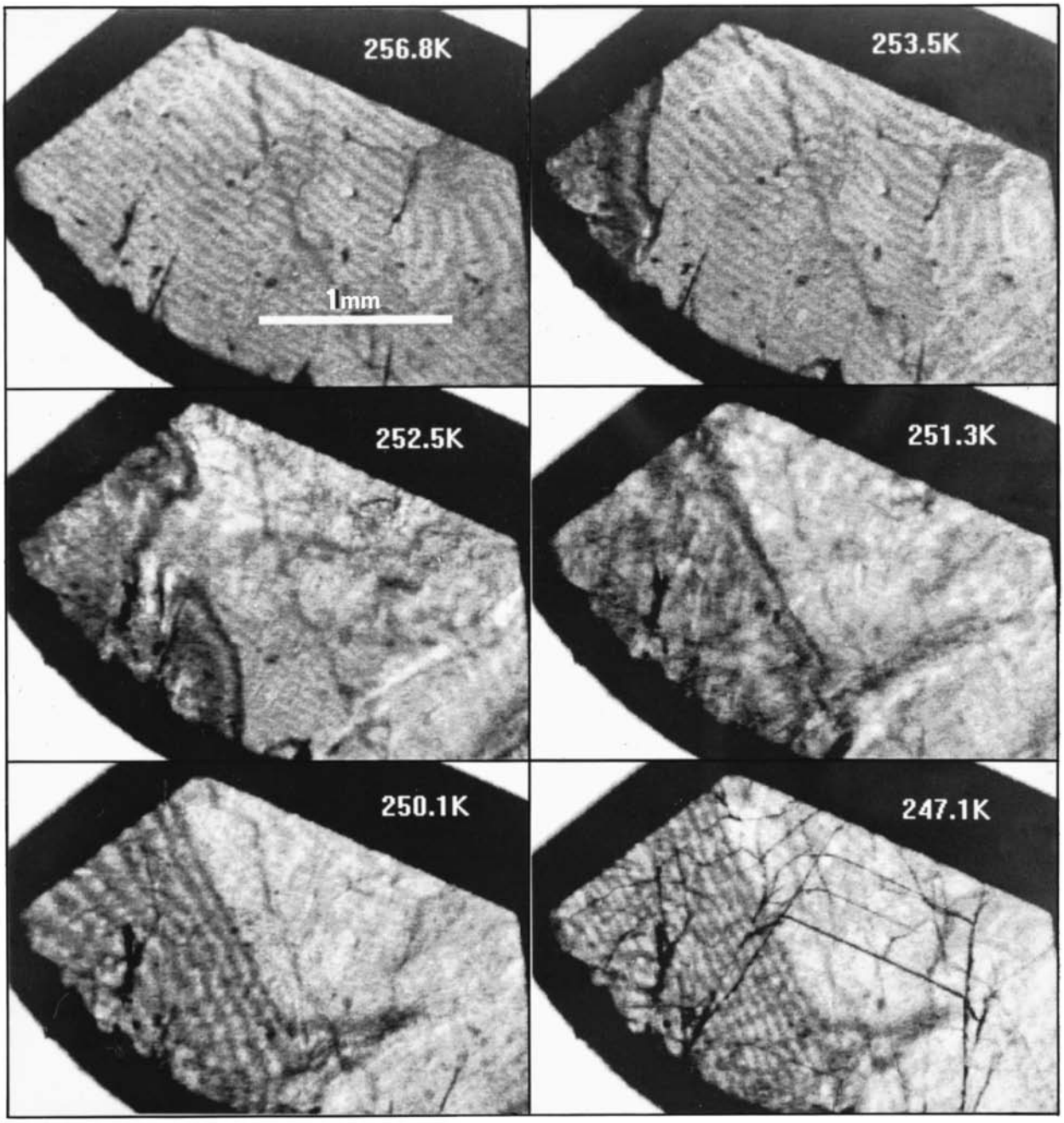

FIG. 6. Photographs of another freshly grown C_{60} single crystal in transmitted light with an energy of 1.71 eV taken as the temperature was lowered at the rate of 0.035 K/s. The thickness of the sample is of the order of 0.3 mm. Six frames taken at various temperatures are shown.

taken into account in interpreting experimental data.

This work was performed with financial support from the ''Current Directions in the Physics of Condensed Media'' Russian Scientific-Technical Program (the ''Fullerenes and Atomic Clusters'' Direction) (Projects 98067 and 98072).

*)E-mail: kveder@issp.ac.ru
†)E-mail: steinman@issp.ac.ru

[1] J. D. Axe, S. C. Moss, and D. A. Neumann, in *Solid State Physics*, Vol. 48, H. Ehrenreich and F. Spaepen (eds.), Academic, New York (1994), p. 150.
[2] C. S. Yannoni, R. D. Johnson, G. Meijer, D. S. Bethune, and J. R. Salem, J. Phys. Chem. **95**, 9 (1991).
[3] P. A. Heiney, Phys. Chem. Solids **53**, 1333 (1992).
[4] Min Gu, Tong B. Tang, J. Phys.: Condens. Matter **7**, 7475 (1995).
[5] T. Atake, T. Tanaka, H. Kawaji, K. Kikuchi, K. Saito, S. Suzuki, I. Ikemoto, and Y. Achiba, Physica C **185–189**, 427 (1991).
[6] P. A. Heiney, J. E. Fischer, A. R. McGhie, W. J. Romanow, A. M. Denenstein, J. P. McCauley, Jr., A. B. Smith III, and D. E. Cox, Phys. Rev. Lett. **66**, 2911 (1991).
[7] W. I. F. David, R. M. Ibberson, T. J. S. Dennis, J. P. Hare, and K. Prassides, Europhys. Lett. **18**, 219 (1992).
[8] R. Moret, P. Launois, and S. Ravy, Fullerene Sci. Technol. **4**(6), 1287 (1996).
[9] W. I. F. David, R. M. Ibberson, and T. Matsuo, Proc. R. Soc. London, Ser. A **442**, 129 (1993).
[10] S. V. Lubenets, V. D. Natsik, L. S. Fomenko, A. P. Isakina, A. I. Prokh-

vatilov, M. A. Strzhemechnyĭ, and N. A. Aksenova, Fiz. Nizk. Temp. **23**, 338 (1997) [Low Temp. Phys. **23**, 251 (1997)].
[11] N. Minami, S. Kazaoui, and R. Ross, Synth. Met. **70**, 1397 (1995).
[12] B. M. Vaughan, Y. Chabre, and D. Dubois, Europhys. Lett. **31**(9), 525 (1995).
[13] S. Matsuura, T. Ishiguro, K. Kikuchi, and Y. Achiba, Phys. Rev. B **51**, 10 217 (1995).
[14] B. Pevzner, A. F. Hebard, and M. S. Dresselhaus, Phys. Rev. B **55**, 16 439 (1997).
[15] V. V. Kveder, V. D. Negrii, É. A. Shteĭnman, A. N. Izotov, Yu. A. Osip'yan, and R. K. Nikolaev, Zh. Éksp. Teor. Fiz. **113**, 734 (1998) [JETP **86**, 405 (1998)].
[16] D. J. van den Heuvel, I. Y. Chan, E. J. J. Groenen, M. Matsushita, J. Schmidt, and G. Meijer, Chem. Phys. Lett. **233**, 284 (1995).
[17] F. Negri, G. Orlandi, and F. Zerbetto, J. Chem. Phys. **97**, 6496 (1992).
[18] B. Renker, F. Gompf, R. Heid, P. Adelmann, A. Heiming, W. Reichardt, G. Roth, H. Schober, and H. Rietschel, Z. Phys. B **90**, 325 (1993).
[19] L. Pintschovus, B. Benker, F. Gompf, R. Heid, S. L. Chaplot, M. Haluska, and H. Kuzmany, Phys. Rev. Lett. **69**, 2662 (1992).
[20] L. Pintschovus and S. L. Chaplot, Z. Phys. B **98**, 527 (1995).
[21] M. R. Philpott, J. Chem. Phys. **54**, 111 (1971).
[22] M. R. Philpott, in *Advances in Chemical Physics, Vol. 23*, I. Prigogine and S. A. Rice (eds.), Wiley, Chichester (1973), p. 227.
[23] K. Yabana and G. F. Bertsch, Chem. Phys. Lett. **197**, 32 (1992).
[24] M. Diehl, J. Degen, and H.-H. Schmidtke, J. Phys. Chem. **99**, 10 092 (1995).
[25] R. D. Bendale, J. F. Stanton, and M. C. Zerner, Chem. Phys. Lett. **194**, 467 (1992).
[26] S. Kazaoui, R. Ross, and N. Minami, Phys. Rev. B **52**(16), R11 665 (1995).
[27] M. Patrini, F. Marabelli, G. Guizzetti, M. Manfredini, C. Castoldi, and P. Milani, in *Proceedings of the ECS, Vol. 94-24*, K. M. Kadish and R. S. Ruoff (eds.), p. 632.

Translated and edited by P. Shelnitz

Underdoped manganites: canted antiferromagnetic ordering or two-phase ferro-antiferromagnetic state?

E. L. Nagaev*)

Institute for High Pressure Physics, 142092 Troitsk, Moscow Region, Russia
(Submitted 2 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2225–2237 (December 1998)

We calculate the energy of charge-carrier-induced canted ordering in conducting layered antiferromagnetic systems with double exchange. The quantum approach to the *d*-spins is used. In the jellium model the energy of the canted state is lower than the energies of both collinear ferro- and antiferromagnetic states over a certain range of charge carrier densities, beginning with arbitrarily small densities. Nevertheless, the canted state cannot be realized, because it is unstable against charge-carrier density fluctuations. The two-phase ferro-antiferromagnetic state can play the role of an alternative to canting. The case of an intermediate electronic-impurity phase separation is investigated. © *1998 American Institute of Physics.* [S1063-7761(98)02612-2]

1. INTRODUCTION

As is well known, charge carriers in magnetic semiconductors tend to establish ferromagnetic ordering at which their energy is minimal. If the ordering in the undoped semiconductor is antiferromagnetic, then with increasing charge carrier density first a magnetic state intermediate between the ferromagnetic and antiferromagnetic states should appear. At still higher densities, the ordering in the crystal becomes completely ferromagnetic. There are different points of view about the nature of the intermediate state at moderate doping (underdoped samples).

De Gennes[1] assumed canted antiferromagnetic ordering to be an intermediate state of degenerate antiferromagnetic semiconductors. He found, treating the *d*-spins as classical vectors, that the canting angle should be proportional to the charge carrier density. As an alternative to canting, in Ref. 2 the electronic phase separation model was proposed, according to which a degenerate antiferromagnetic semiconductor with frozen impurities in its ground state is separated into an insulating antiferromagnetic phase and high-conductivity ferromagnetic phases. Later the case of ferro-antiferromagnetic phase separation was considered in systems with mobile impurities and a high magnetic ordering temperature (e.g., oxygen in perovskites[3]). Then not only the charge carriers but also the ionized donors of acceptors are concentrated in the ferromagnetic portion of the crystal (impurity phase separation).

As for experimental verification of these theories, the electronic phase separation theory was confirmed by electric, magnetic and magnetooptic data on EuSe, EuTe, and so on (see Ref. 4). On the other hand, De Gennes[1] interpreted data of neutron studies of the doped lanthanum manganites $La_{1-x}Ca_xMnO_3$ (Ref. 5) as confirmation of canted antiferromagnetic ordering in them. This idea was accepted by many investigators, even those who are engaged in neutron studies, and now it is customary to refer to the results of Ref. 5 as providing verification of canted ordering.

In reality, this involves a misunderstanding: Wollan and Koehler[5] arrived just to the opposite conclusion. They pointed out that in principle the superposition of ferro- and antiferromagnetic peaks they observed at $x=0.18$ can be related to both the canted antiferromagnetic ordering and mixture of the ferro- and antiferromagnetic regions. To choose between these two possibilities, they investigated the behavior of the peaks in the magnetic field. They found that a field of about 4 kOe halves the height of the ferromagnetic peaks but does not influence the height of the antiferromagnetic peaks.

But in the case of canted ordering, the ferromagnetism vector is rigidly related to the antiferromagnetism vector, and the field should rotate these vectors simultaneously. Hence, both the ferro- and antiferromagnetic peaks should vary in intensity simultaneously. In contrast, in case of the phase-separated state, the ferromagnetism and antiferromagnetism vectors are independent. This led Wollan and Koehler[5] to conclude (and this conclusion they expressed in plain words) that the phase separation, rather than canting, takes place in their samples. But the nature of the phase separation remains as yet unknown.

Strictly speaking, a very small canting of the relativistic origin was discovered later[6] in the undoped $LaMnO_3$ but it has nothing in common with the charge-carrier-induced canting proposed in Ref. 1. Its existence was confirmed in Refs. 7 and 8 by neutron investigations of $La_{1-x}Sr_xMnO_3$ and $La_{1-x}Ca_xMnO_3$, respectively. For example, in the former at $x=0.04$, the ferromagnetic and antiferromagnetic peaks appear simultaneously at the same temperature (136 K), which unequivocally confirms the canted structure. But at $x=0.125$ the ferromagnetic peaks appear at 230 K and the antiferromagnetic peaks only at 150 K.[7]

The most natural explanation of this difference is that the ferromagnetism is not related to the antiferromagnetism as the corresponding regions are spatially separated. In other words, this result can be considered to be consistent with

conclusions of Ref. 5. But in Ref. 7 the hypothesis was advanced that as the temperature decreases first the ferromagnetic ordering is established, and then it is replaced by canted antiferromagnetic ordering. Investigations in a magnetic field similar to those carried out in Ref. 5 might address the question of whether the hypothesis of Ref. 7 is adequate.

In Ref. 8, short-range ferromagnetic correlations with a length of several lattice constants were discovered, which were attributed to moving magnetic polarons (ferrons). This result is very important as it directly confirms phase separation: the appearance of ferromagnetic regions inside antiferromagnetic crystals. In our opinion, these correlations should be attributed not to the moving magnetic polarons (ferrons) predicted in Ref. 2: their number is exponentially small, and they are unlikely to be observed in neutron studies. But ferromagnetic correlations can be attributed to ferrons bound to ionized acceptors.[2] Their number is several orders of magnitude larger than the number of free ferrons.

Not only neutron data, but also electric data on manganites conflict with the De Gennes scenario of the appearance of canted antiferromagnetic ordering due to mobile holes. In fact, contrary to the De Gennes assumption, incompletely magnetized materials are not highly conductive but insulating. In addition to Ref. 8, one should also mention other investigations with larger acceptor contents (see Ref. 9). But this is not inconsistent with the properties of the phase-separated semiconductors: at modest charge carrier density they behave like the insulators.[2–4]

In the present paper we describe a detailed investigation of the assumed canted layered antiferromagnetic ordering under the double exchange conditions typical of lanthanum manganites. Unlike that of Ref. 1, our treatment of the magnetic system will not be classical but quantum-mechanical. It will be shown that although in a certain charge-carrier density range the canted antiferromagnetic structure is more energetically favored than the collinear antiferromagnetic and ferromagnetic structures, it is nevertheless not stable against transformation to a nonuniform state. The instability of the uniform state is seen from the fact that the *s*-electron screening length diminishes with increasing charge carrier density, and formally becomes imaginary at moderate densities.

A new type of ferro-antiferromagnetic phase separation will be considered as a possible alternative to canting. In particular, manganites usually contain not only immobile acceptors (Ca, Sr, and so on) but also mobile acceptors in the form of excess oxygen. The situation intermediate between electronic and impurity phase separation will be considered: in the ferromagnetic portion of the crystal, holes of both immobile and mobile acceptors congregate with the mobile acceptors themselves.

2. DOUBLE-EXCHANGE-INDUCED CANTED ANTIFERROMAGNETIC ORDERING

Canted antiferromagnetic ordering in a layered antiferromagnetic structure will be considered, as this is just the structure realized in lanthanum manganites. The crystalline structure is assumed to be simple cubic.

The charge-carrier energy spectrum for double exchange will be found. It is commonly believed that this limiting case of extremely strong s-d coupling is realized in lanthanum manganites, though some experimental data point to inadequacy of the double exchange (see Ref. 9). Nevertheless, there are other data that support double-exchange scenario in these materials. For example, neutron studies suggest that the magnon spectrum of doped lanthanum manganites corresponds to nearest neighbors, which is inconsistent with standard RKKY indirect exchange.[10] On the other hand,[4,11] this is just the case for indirect exchange in double-exchange systems. We also take note of Ref. 12, in which it was found by optical studies that the Mn e_g^1 band is the highest of the filled bands so that holes only appear in this band.

The treatment is based on the standard s-d model with the Hamiltonian

$$H=-t\sum a^*_{\mathbf{g}\sigma}a_{\mathbf{g}+\Delta\sigma}-A\sum (\mathbf{sS_g})_{\sigma\sigma'}a^*_{\mathbf{g}\sigma}a_{\mathbf{g}\sigma'}$$
$$-\frac{1}{2}\sum I(\mathbf{f})\mathbf{S_g}\mathbf{S_{g+f}}, \qquad (1)$$

where $a^*_{\mathbf{g}\sigma}$, $a_{\mathbf{g}\sigma}$ are the s-electron operators corresponding to conduction electrons or holes located at atom $\mathbf{g}$ with the spin projection σ, $\mathbf{s}$ is its spin operator, $\mathbf{S_g}$ is that of the d-spin of atom $\mathbf{g}$, and $\mathbf{\Delta}$ is the vector connecting nearest neighbors. The d-d exchange interaction in Eq. (1) is taken in the form ensuring the existence of the ferromagnetically ordered (001) planes with alternating moments. In particular, in the Jahn–Teller systems to which the lanthanum manganites belong, the in-plane $d-d$ exchange can be ferromagnetic, whereas the out-of-plane $d-d$ exchange can be antiferromagnetic.[13] In this case the nearest-neighbor approximation is sufficient.

The double-exchange condition can be formulated mathematically as a requirement that the s-electron band width $W=2zt$ be small compared to AS, where S is the d spin magnitude and z is the coordination number. The $s-d$ exchange integral A is assumed to be positive. The inequality $W\gg|I|zS^2$ should also be met, as these quantities are of the first and second order in the small d-orbital overlap, respectively.

Unlike Ref. 1, where the d spins were considered as classical, here the d spins will be considered as quantum-mechanical, and the inequality $2S\gg1$ is not assumed. The quantum approach is necessary to investigate the stability of the canted antiferromagnetic ordering induced by double exchange. The wave function of an s electron in the completely antiferromagnetically ordered crystal is expanded in the eigenfunctions of the s-d exchange Hamiltonian [the second term in Eq. (1)]. Such a quantum-mechanical treatment was first carried out in Ref. 14 for a system of two atoms, and by the present author for a system consisting of an arbitrarily large number of atoms.[15] It is assumed that the moments of the two sublattices ($i=1$ or 2) make angles $\pm\theta$ with the total moment of the crystal.

For any atom, the z axis is aligned with the moment of the sublattice to which the atom belongs. The following eigenfunctions of the $s-d$ exchange Hamiltonian will be used as the basis for the wave function expansion:

$$\Phi_0(\mathbf{g})=a^*_{\mathbf{g}\uparrow}|0\rangle\prod_f \delta(S^z_f,S),$$

$$\Phi_1(\mathbf{g})=\frac{1}{\sqrt{2S+1}}[a^*_{\mathbf{g}\downarrow}+a^*_{\mathbf{g}\uparrow}S^-_{\mathbf{g}}]|0\rangle\prod_{\mathbf{f}} \delta(S^z_{\mathbf{f}},S), \qquad (2)$$

where $S^-=S^x-iS^y$, $\delta(n,m)=1$ for $n=m$ and 0 otherwise, and $|0\rangle$ is the s electron vacuum function.

To find the energy to first order in t, the wave function is represented by the linear combination

$$\Psi=\sum_{i=1}^{2}\sum_{\mathbf{g}_i}[X(\mathbf{g}_i)\Phi_0(\mathbf{g}_i)+Y(\mathbf{g}_i)\Phi_1(\mathbf{g}_i)]. \qquad (3)$$

In fact, this is the Ritz variational procedure: In Eq. (3), terms are omitted that correspond to the d spins deviated from the moment of their sublattice in the absence of s electrons at them. These terms correspond to string-like motion of the charge carrier,[16] which is of vital importance at $A<0$ and $S=1/2$, but is not significant at $A>0$.[4]

To proceed further, transformation rules for the electron operators from one reference frame to the other are necessary:

$$a(\mathbf{g}+\mathbf{\Delta},\downarrow_{\mathbf{g}+\Delta})=\cos\theta a(\mathbf{g}+\mathbf{\Delta},\downarrow_{\mathbf{g}})+i\exp(i\mathbf{Qg})\times\sin\theta a(\mathbf{g}+\mathbf{\Delta},\uparrow_{\mathbf{g}}),$$

$$a(\mathbf{g}+\mathbf{\Delta},\uparrow_{\mathbf{g}+\Delta})=\cos\theta a(\mathbf{g}+\mathbf{\Delta},\uparrow_{\mathbf{g}})+i\exp(i\mathbf{Qg})\times\sin\theta a(\mathbf{g}+\mathbf{\Delta},\downarrow_{\mathbf{g}}), \qquad (4)$$

where $\mathbf{Q}$ is the antiferromagnetic wave vector, and the index of the spin projection points to the atom in whose reference frame this projection is measured. We put $\mathbf{g}=(\mathbf{r},z)$ and denote vectors connecting nearest neighbors in the plane and between the planes by $\mathbf{\Delta}$ and $\mathbf{n}$, respectively. Then with allowance for the relationships $t\gg|I|$ and $\mathbf{Q}=\pi(0,0,1)$, one obtains from Eqs. (1)–(4) for the s electron energy E_s

$$\left(E_s+\frac{AS}{2}\right)X_{\mathbf{g}}=-t\cos\theta\sum_{\mathbf{n}}X_{\mathbf{g}+\mathbf{n}}-t\sum_{\mathbf{\Delta}}X_{\mathbf{g}+\mathbf{\Delta}}-\frac{it\sin\theta e^{i\mathbf{Qg}}}{\sqrt{2S+1}}\sum_n Y_{\mathbf{g}+\mathbf{n}},$$

$$\left(E_s+\frac{AS}{2}\right)Y_{\mathbf{g}}=-\frac{t\cos\theta}{2S+1}\sum_{\mathbf{n}}Y_{\mathbf{g}+\mathbf{n}}-\frac{t}{2S+1}\times\sum_{\mathbf{\Delta}}Y_{\mathbf{g}+\mathbf{\Delta}}-\frac{it\sin\theta e^{i\mathbf{Qg}}}{\sqrt{2S+1}}\sum_{\mathbf{n}}X_{\mathbf{g}+\mathbf{n}}. \qquad (5)$$

Putting $X_{\mathbf{g}}=Xe^{i\mathbf{qg}}Y_{\mathbf{g}}=Ye^{i(\mathbf{q}+\mathbf{Q})\mathbf{g}}$, and omitting the constant $AS/2$ in Eq. (5), one can rewrite Eq. (5) in the form

$$(E_s+P_{\mathbf{k}})X=-R_{\mathbf{p}}\cos\theta X-\frac{R_{\mathbf{p}}\sin\theta}{\sqrt{2S+1}}Y,$$

$$\left(E_s+\frac{P_{\mathbf{k}}}{2S+1}\right)Y=-\frac{R_{\mathbf{p}}\cos\theta}{2S+1}Y+\frac{R_{\mathbf{p}}\sin\theta}{\sqrt{2S+1}}X \qquad (6)$$

with $\mathbf{q}=(k_x,k_y,p)$, $P_{\mathbf{k}}=2t[\cos(k_xa)+\cos(k_ya)]$, and $R_p=2t\cos(pa)$.

One obtains from Eqs. (6) the following expression for the electron energy:

$$E_s(\mathbf{k}p)=-\frac{(S+1)P_{\mathbf{k}}}{2S+1}\frac{MR_p}{2S+1}\pm\frac{1}{2S+1}[S^2P^2_{\mathbf{k}}+(M^2+2S+1)R^2_p+2M(S+1)P_{\mathbf{k}}R_p]^{1/2}, \qquad (7)$$

where $M=S\cos\theta$ is the magnetization per atom, and a is the lattice constant.

As can be seen from Eq. (7), the charge-carrier spectrum consists of two subbands. In complete ferromagnetic ordering, the lower and upper subbands correspond to the total spin projection of an s-electron-loaded atom, equal to $S+1/2$ or $S-1/2$, respectively. At arbitrary magnetization, one of the subbands remains well below the other. As the number of charge carriers is small compared to the number of magnetic atoms, only the low-energy electronic subband matters. If $2S\gg1$ and $M^2\gg2S$, one obtains from Eq. (7) the expression for the s electron energy, which coincides with the result obtained in Ref. 1:

$$E_s(\mathbf{k}p)=-P_{\mathbf{k}}-\frac{R_pM}{S}. \qquad (8)$$

For arbitrary M and S in the quadratic approximation, this energy can be written in the form ($\hbar=1$)

$$E(\mathbf{k}p)=B+\frac{k^2}{2m_{xy}}+\frac{p^2}{2m_z},$$

$$B=-t\frac{4(S+1)+2M+\sqrt{Z}}{2S+1},$$

$$\frac{1}{2m_{xy}}=\frac{ta^2}{2S+1}\left[S+1+\frac{4S^2+2M(S+1)}{\sqrt{Z}}\right],$$

$$\frac{1}{2m_z}=\frac{ta^2}{2S+1}\left[M+\frac{2(M^2+2S+1)+4M(S+1)}{\sqrt{Z}}\right],$$

$$Z=16S^2+4(M^2+2S^2+1)+16M(S+1). \qquad (9)$$

This result makes it possible to find the magnetization that minimizes the energy of the canted antiferromagnetic ordering. At this stage, keeping in mind that the number ν of s electrons per atom is small, one can put the total energy per atom equal to $E_t=\nu B-IS^2\cos2\theta$, where B is given by Eq. (9), and I is the interplane d-d exchange integral in the nearest-neighbor approximation. Minimizing E_t with respect to M, one obtains the equation for the equilibrium magnetization, which in the limit $2S\to\infty$ goes over into the De Gennes result[1]:

$$M=\frac{2\nu t}{(2S+1)|I|}\left[1+\frac{4(S+1)+2M}{\sqrt{Z}}\right]. \qquad (10)$$

In case considered, M is nonzero for all ν. In particular, for large S the magnetization is proportional to ν. From Eq. (10),

one obtains the following expression for the relative electron density ν_F at which complete ferromagnetic ordering is established:

$$\nu_F = \frac{|I|S(2S+1)}{4t}\left[1+\frac{1}{2(3S+1)}\right]^{-1}. \quad (11)$$

Obviously, the inequality $|I|S^2 \ll t$ should hold as the quantities I and t are of the second and first order in the small overlap of d orbitals of neighboring atoms, respectively. For this reason, the inequality $\nu_F \ll 1$ should also hold.

It is interesting to compare results for the layered structure just obtained with the results obtained earlier for a staggered structure.[17] In the latter case, the effective hopping integral for the collinear antiferromagnetic structure is equal to $t/\sqrt{2S+1}$. Hence, it can be rather large even at $2S \gg 1$, keeping in mind that $2S \leqslant 5$ for d shells. But for the layered structure, according to Eq. (9), it is equal to $t/4S$ for $2S \gg 1$, i.e., it is really small. This justifies the De Gennes approach.[1]

Another drastic difference between these two structures is the fact that for the staggered structure, canting is energetically favored starting not at some arbitrarily low charge-carrier density, but at some finite value. At lower densities, staggered antiferromagnetic collinear ordering is at least relatively stable, whereas layered collinear ordering is unstable at all densities.

3. INSTABILITY OF THE CANTED ANTIFERROMAGNETIC ORDERING

So far we have shown that a range of charge-carrier densities exists in which canted antiferromagnetic ordering is more energetically favored than the collinear antiferromagnetic or ferromagnetic ordering. But this does not necessarily ensure the stability of a canted structure in magnetic systems with an isotropic exchange interaction. In what follows, it will be proved that in reality, the canted state is absolutely unstable against arbitrarily small fluctuations of electric fields, if the charge carrier density is not too low. Hence, the results of the preceding section, which attest to the instability of the collinear antiferromagnetic and ferromagnetic states in a certain ν range, do not ensure stability of the canted state in this range.

The fact of its being energetically favored as compared with the collinear ferromagnetic and antiferromagnetic states simply implies the absolute instability of the collinear states, but does not prove the stability of the canted state.

The instability of the canted state may be due to the fact that the band bottom position B depends on the magnetization M and decreases with decreasing M (9). On the other hand, the local magnetization increases with local charge-carrier density (10). Hence, if a local density fluctuation lowers the local band bottom, a local potential well for carriers arises at that location. The carriers tend to increase the local density still more. But there are two factors hindering an increase in the fluctuations: a rise in the kinetic energy of the charge carriers, and the Coulomb interaction between them in the region of their enhanced density. Competition among all these factors determines whether the initial fluctuation will continue to increase or it will begin to decrease. In the former case, the uniform canted state will be destroyed.

Here, as is customary for degenerate semiconductors, the jellium model is used, which is applicable at $\mu \gg e^2 n^{1/3}/\varepsilon_0$, where μ is the Fermi energy, $n = \nu a^3$ is the charge-carrier density, and ε_0 is the dielectric constant. In the jellium model, instability of the canted state against fluctuations should be manifested mathematically by the screening length becoming imaginary. It means that any arbitrarily small electric field makes the uniform state unstable. This field can be caused by a fluctuation of the charge-carrier density, i.e., one may speak of instability against the density fluctuations. This points to the tendency for the system to go over into a nonuniform state, i.e., to a phase separation. Certainly, an imaginary screening length is a sufficient but not a necessary condition for instability of the uniform state.

The screening length for the canted state will be calculated in the Born–Oppenheimer approximation. This means that the magnetization $M(\mathbf{r})$ and the band bottom position $B(\mathbf{r})$ are smoothly varying functions of the coordinates. This is justified by the fact that the typical length over which they change is the screening length $1/\varkappa$, which greatly exceeds the lattice constant a in degenerate semiconductors. As both these quantities depend on $\mathbf{r}$ via $n(\mathbf{r})$, one may put

$$B(\mathbf{r}) = B + \frac{dB}{dM}\frac{dM}{dn}\delta n(\mathbf{r}),$$

$$n(\mathbf{r}) = n + \delta n(\mathbf{r}), \quad \delta n(\mathbf{r}) = \sum n(\mathbf{q}) e^{i\mathbf{qr}}, \quad (12)$$

where B and n are the average values of the corresponding quantities.

The requirement of constant electrochemical potential in the presence of an external electrostatic field with potential $\Phi(\mathbf{q})$,

$$\mu(\mathbf{r}) + B(\mathbf{r}) + e\phi(\mathbf{r}) = \text{const}, \quad (13)$$

after linearization with respect to $\delta n(\mathbf{r})$ takes the form

$$\frac{d\mu}{dn} + e\phi(\mathbf{q}) + B(\mathbf{q}) = 0, \quad \phi(\mathbf{q}) = \frac{\Phi(\mathbf{q})}{\varepsilon(\mathbf{q})}, \quad (14)$$

where the Fermi energy μ is measured from the bottom of the band.

Using Eq. (12), one obtains the relationship between the internal field $\phi(\mathbf{q})$ and the corresponding fluctuation of the electron density:

$$n(\mathbf{q}) = -\frac{dn}{d\mu}\frac{e\phi(\mathbf{q})}{1-\Gamma}, \quad (15)$$

where the magnetoelectric constant Γ is introduced:

$$\Gamma = -\frac{dB}{dM}\frac{dM}{dn}\frac{dn}{d\mu}. \quad (16)$$

In what follows, the standard pattern for calculating the screened potential is used. One introduces the field $\delta\phi(\mathbf{q})$ created by the s electrons polarized by the external field:

$$\delta\phi(\mathbf{q})=\phi(\mathbf{q})-\frac{\Phi(\mathbf{q})}{\varepsilon_0}=\left[1-\frac{\varepsilon(\mathbf{q})}{\varepsilon_0}\right]\phi(\mathbf{q}). \tag{17}$$

With allowance for Eq. (15), the Poisson equation takes the form

$$q^2\delta\phi(\mathbf{q})=-\kappa^2\phi(\mathbf{q}), \tag{18}$$

where

$$\kappa^2=\frac{4\pi e^2}{\varepsilon_0(1-\Gamma)}\frac{dn}{d\mu}.$$

As can be seen from Eqs. (17) and (18), the quantity $1/\kappa$ is just the screening length. Obviously, with increasing Γ, the screening length decreases and becomes imaginary when the magnetoelectric constant Γ exceeds unity. For the jellium model used when the compensating charge of ionized impurity is assumed to be distributed uniformly, the condition $\Gamma=1$ can be considered necessary for the stability of the canted structure. Hence, the problem consists in obtaining an explicit expression for this quantity in the canted structure.

First, the quantity $dn/d\mu$ will be found. We introduce the effective mass of the density of states, $m_{\rm dos}$, according to the standard procedure for semiconductors with an anisotropic effective mass, as in Eq. (9). To obtain an expression for $m_{\rm dos}$, the following equation for the density of states is used:

$$g(E)=\frac{Na^3}{(2\pi)^3}\int_{-\pi/a}^{\pi/a}dk_x dk_y dp\,\delta\left(\frac{k_x^2+k_y^2}{2m_{xy}}+\frac{p^2}{2m_z}-E\right). \tag{19}$$

This expression can be obtained from the expression for the density of states when $\pi^2/2m_z a^2\gg\mu$. If the inequality does not hold, one should consider the *s* electron motion to be two-dimensional. But this ''ultraclassical'' case emerges only when $2S\to\infty$, and for this reason it is purely of academic interest as $2S\leqslant 5$ for *d*-shells, and hence at fairy low densities the condition just mentioned is satisfied.

As follows from Eq. (19), the Fermi energy is given by

$$\mu=\frac{(6\pi^2 n)^{2/3}}{2m_{\rm dos}},\quad m_{\rm dos}=(m_{xy}^2 m_z)^{1/3}, \tag{20}$$

where according to Eq. (9), the quantity $m_{\rm dos}$ is a function of *M*, and thus of *n*.

To make subsequent calculations more transparent, we first consider the case in which Eq. (8) is valid ($M^2\gg 2S+1\gg 1$). Then

$$m_{\rm dos}=m\left(\frac{S}{M}\right)^{1/3},\quad \frac{1}{2m}=ta^2,$$

$$B=-4t-\frac{2tM}{S},\quad M=\frac{4\nu t}{(2S+1)|I|}=\frac{S\nu}{\nu_F}, \tag{21}$$

and one obtains from Eqs. (16) and (21)

$$\Gamma=\frac{2}{(6\pi^2)^{2/3}\nu_F^{2/3}}. \tag{22}$$

As can be seen from Eq. (22), the magnetoelectric constant Γ exceeds unity at $\nu_F<0.05$. Hence, the canted antiferromagnetic ordering at $M^2\gg 2S+1$ is absolutely unstable for such ν_F. But this is not the case for $0.05<\nu_F\ll 1$.

In the opposite limiting case, $2S\gg 1$ and $M\ll 2S$, one finds by a similar calculation

$$\Gamma=\frac{3\nu^{1/3}}{(6\pi^2)^{2/3}\nu_F}. \tag{23}$$

We see from Eq. (23) that when $\nu\to 0$, canted ordering remains at least relatively stable.

Similar qualitative conclusions were drawn in Ref. 9 about stability of the canted structure in a staggered antiferromagnet, although the case of very low ν is meaningless for them: at such densities collinear antiferromagnetic ordering is energetically more favorable than canted ordering. But in case of staggered ordering, a more realistic model with randomly distributed point impurities can be used instead of the jellium model to investigate the stability of the canted structure.[9]

In this model one takes into account that a charge carrier is attracted to the ionized impurity not only by the Coulomb force but also by the force related to the magnetization-dependent location of the bottom of the band (the expression for which is similar to Eq. (9) for *B*). This force arises because in the vicinity of an impurity the charge-carrier density is higher than its average value. Hence, according to an expression for *M* similar to Eq. (10), the same is true for the magnetization. As the total force attracting the charge carrier to the impurity exceeds the Coulomb force, the condition for Mott delocalization of the donor electrons is more stringent in the case of canted ordering, than in case of collinear ordering.

For this reason, canting can be unstable against localization of the charge carriers. But canted antiferromagnetic ordering was obtained under the assumption of delocalized charge carriers, which means that this assumption is invalid. Hence, the uniform canted ordering is unstable. Mathematically, the condition of Mott delocalization leads to the stability condition $\Gamma<1$, instead of $\Gamma=1$ as in the jellium model.

Unfortunately, a theory of Mott delocalization in anisotropic systems is lacking at present, so it is impossible to use this approach for a layered structure. One must merely be aware that $\Gamma=1$ is a sufficient (and not a necessary) condition for instability of the canted structure. Nevertheless, as seen from Eq. (22), in typical cases the margin of stability is very large.

4. MIXED ELECTRONIC-IMPURITY PHASE SEPARATION

We now study phase separation as an alternative to canting. As electronic and impurity phase separation were already investigated in Refs. 2 and 3, here we study a special situation that is likely to be typical of manganites. It will be assumed that two types of acceptors (donors) are present in the crystal: immobile (Ca, Sr, etc.) and mobile (oxygen). Mixed impurity-electronic ferro-antiferromagnetic phase separation should then occur. It is characterized by the fact that the ferromagnetic phase becomes the places where holes of the immobile acceptors and holes of the mobile acceptors

congregate with these ionized mobile acceptors. The holes and mobile acceptors are absent from the antiferromagnetic portion of the crystal.

This type of phase separation is of special interest for the following reason. The main difference between electronic and impurity phase separation lies in the fact that the local electric charge is nonzero in the former case and zero in the latter. Indeed, in electronic phase separation the positions of the impurity atoms are frozen. For this reason, the concentration of charge carriers in a certain phase leads to its becoming charged, and to the opposite charge of the other phase, due to the remaining ionized impurity atoms. Coulomb forces tend to mix both phases, but surface forces limit the mixing. If the crystal is isotropic, the phase of the smaller volume (''minor phase'') consists of small droplets several nanometers in size embedded in the host (''major phase'').[2,4]

As for impurity phase separation, here there are no Coulomb forces, since the local charges of the nonuniformly distributed electrons (holes) are balanced everywhere by the charge of ionized donors (acceptors). Nevertheless, the tendency to phase mixing exists here too: it reduces elastic forces resulting from the difference in the elastic moduli of the two phases. As a result, the minimum energy corresponds to a plane-parallel geometry (alternating layers of the two phases), but under typical conditions their size is several orders of magnitude larger than the radius of the droplets in the electronic phase separation.[9] As the Coulomb forces are much stronger than the elastic forces, in mixed phase separation the former should determine the geometry of the system as well.

To calculate the energy of the phase-separated state, a variational procedure will be used. It generalizes the procedure developed in Ref. 2. It is assumed that the minor phase consists of spheres of radius R arranged periodically inside the major phase. The second variational parameter is the ratio of volumes of the antiferromagnetic and ferromagnetic phases: $w=V_A/V_F$. If the minor phase is highly conductive, the electronic part of the wave function at $T=0$ is taken in the form of the antisymmetrized product of the single-electron wave functions $\psi(\mathbf{r})$ corresponding to the free motion of an electron inside a spherical region of radius R,

$$\Psi=\frac{1}{\sqrt{N_e!}}\mathrm{Det}[\,\psi_k(\mathbf{R}_i-\mathbf{r}_{ni})\,], \tag{24}$$

where $\mathbf{R}_i$ is the location of the center of the ith sphere, $\mathbf{r}_{ni}$ is the location of the nth electron inside the ith sphere, and N_e is the total number of electrons. In the ground state the indices k of the single-electron states correspond to the single-electron energies E_k below the Fermi energy μ.

Dirichlet boundary conditions are used for each sphere. This is justified even for fairly shallow potential wells if $\sqrt{2m(U-E_k)}R\gg 1$, where U is the potential-well depth and m is the electron effective mass. The wave function (24) is accurate at radii less than the screening length in ferromagnetic ordering. At larger R, it gives the upper bound on the energy of the phase-separated state.

If $\sqrt{2m\mu}R\gg 1$, the Born-Oppenheimer approximation can be used to calculate the electron kinetic energy E_K. In this approximation, a memory of the spatial quantization of the electron motion remains in the form of the surface electron energy E_S, which be added to the bulk energy E_B,

$$E_K=E_V+E_S, \tag{25}$$

$$E_V=\frac{3}{5}\mu(n)n(1+w)^{2/3}V,\quad \mu(n)=\frac{(6\pi^2 n)^{2/3}}{2m}, \tag{26}$$

$$E_S=\beta\left(\frac{\pi}{6}\right)^{1/3}\frac{5E_V}{16n^{1/3}(1+w)^{1/3}R}, \tag{27}$$

where $n=n_A+n_0$ is the mean hole (or electron) density. It is composed of the densities n_A and n_0 associated with immobile acceptor ions, and excess oxygen, respectively. Further, $n(1+w)$ is the charge-carrier density inside the ferromagnetic phase, V is the total volume of the sample, $\beta=3$ if the ferromagnetic phase is the minor one, with volume $V/(1+w)$.

The Coulomb energy is calculated using elementary electrostatics. For ferromagnetic spheres one obtains

$$E_C=\frac{2\pi n_A^2}{5\varepsilon_0}e^2R^2f(w)V,$$

$$f(w)=2w+3-3(1+w)^{2/3}. \tag{28}$$

In this case, in which the major phase is ferromagnetic and antiferromagnetic droplets are embedded in the host, the surface energy is given by Eq. (27) with $\beta=3w$ and the Coulomb energy is given by Eq. (28), in which the function $f(w)$ must be replaced by the function

$$g(w)=w[2+3w-3w^{1/3}(1+w)^{2/3}]. \tag{29}$$

In the same nearest-neighbor approximation as before, the $d-d$ exchange energy is

$$E_{dd}=\left[\frac{D-HS}{(1+w)v}-\frac{H^2S^2w}{4Dv(1+w)}\right],\quad D=2|I|S^2, \tag{30}$$

where $v=a^3$ is the unit cell volume and H is the external magnetic field. The first term in Eq. (30) represents an increase in $d-d$ exchange energy due to replacement of antiferromagnetic ordering by ferromagnetic; the second term is the energy of the antiferromagnetic phase in the field.

Finally, we present the s electron energy difference E_U between the ferromagnetic and antiferromagnetic state, which one easily obtains from Eq. (7):

$$E_U=-2(t-t_A)nV,\quad t_A=t\frac{2S}{2S+1}\left(\sqrt{1\frac{2S+1}{4S^2}}-1\right). \tag{31}$$

We now minimize the total energy

$$E_t=E_V+E_S+E_C+E_{dd}+E_U.$$

As seen from Eqs. (25)–(31), only the surface energy E_S and Coulomb energy E_C depend on R. This makes it possible to minimize the total energy with respect to R in explicit form. In ferromagnetic droplets, the optimized energy and radius are

$$E_R \equiv (E_S + E_C)_{\rm opt} = (2^{-2/3} + 2^{1/3}) \left(\frac{9m^2}{160\pi} \right)^{1/3} \times \frac{e^{2/3} n^{2/3} \mu^{4/3}(n) f^{1/3}(w) (1+w)^{2/9}}{\zeta_0^{1/3}}, \quad (32)$$

$$R_{\rm opt}^3 = \frac{135\pi^2 \zeta_0 (1+w)^{1/3}}{32 m^2 e^2 \mu(n) f(w)}, \quad (33)$$

where the effective dielectric constant corresponding to the mixed phase separation is

$$\zeta_0 = \varepsilon_0 \left(1 + \frac{n_0}{n_A} \right)^2. \quad (34)$$

Energy minimization with respect to w must be carried out numerically. But if the energy E_R is low, it only weakly influences the optimum value of w, which is determined mainly by E_V and E_{dd} in this case. One then obtains for $H \ll D$

$$\frac{V_F}{V} = \frac{1}{1+w} = \left[\frac{3\mu(1/v)}{5(D - HS)} \right]^{3/5} nv. \quad (35)$$

As seen from Eqs. (33)–(35), at fixed total charge-carrier density n, the volume of ferromagnetic droplet increases quadratically with the relative weight of the impurity phase separation n_0/n_A [Eq. (34)]. Hence, here the size of ferromagnetic droplets can be considerably larger than in pure electronic phase separation. Thus, these droplets can be manifested not by small-angle neutron scattering but by well-formed ferromagnetic peaks. As such peaks were observed in Ref. 5, one of the possible reasons for their appearance might be mixed phase separation.

According to Eq. (33), the droplet size decreases with increasing w. Hence, according to Eq. (35), the volume of the ferromagnetic part of the crystal also increases with the field. For this reason, beginning with a certain field strength, droplets should begin to make contact with one another, and the charge carriers acquire the ability to move freely from one droplet to another. This means that the magnetic-field-induced transition from the insulating state to the high-conductivity state occurs in the sample as a whole, and can be considered as a manifestation of the giant magnetoresistance.

We note here some other recent publications on phase separation in manganites. First, Ref. 18 is a continuation of Ref. 19. In these papers a simplified treatment is carried out as compared with Ref. 2 and the present paper: the Coulomb interaction and interphase surface energy are not taken into account (the authors of Refs. 18 and 19 are likely not to be acquainted with Ref. 2 and subsequent publications of the present author). Calculations in Ref. 18 are carried out only for the ground state, assuming double exchange. Different dimensionalities are treated, beginning with an one-dimensional system and proceeding to an infinite-dimensional system.

One can also consider the appearance of charge-ordered stripes in $La_{1-x}Ca_xMnO_3$ with $x > 0.5$ as a special sort of phase separation. This pattern of phase separation takes the form of extremely stable pairs of $Mn^{3+}O_6$ stripes separated periodically by stripes of undistorted $Mn^{4+}O_6$ octahedra.[20] Some comments on this subject are given in Ref. 21, but they do not pretend to be a theory of this interesting and complicated phenomenon.

The author is grateful to D. I. Khomskii and M. Yu. Kagan for valuable discussions.

This investigation was partially supported by the Russian Fund for Fundamental Research (Grant No. 98-02-16148) and by the Russian Ministry of Science (Grant No. 97-1076(072)).

*)E-mail: tsir@elch.chem.msu.ru

[1] P. De Gennes, Phys. Rev. **118**, 141 (1960).
[2] E. L. Nagaev, JETP Lett. **6**, 18 (1967); **16**, 394 (1972); Zh. Éksp. Teor. Fiz. **54**, 228 (1968) [Sov. Phys. JETP **27**, 122 (1968)]; V. A. Kashin and E. L. Nagaev, Zh. Éksp. Teor. Fiz. **66**, 2105 (1974) [Sov. Phys. JETP **39**, 1036 (1974)].
[3] E. L. Nagaev, Physica C **222**, 234 (1994).
[4] E. L. Nagaev, *Physics of Magnetic Semiconductors*, Mir, Moscow (1983); Usp. Fiz. Nauk **165**, 529 (1995) [Phys. Usp. **38**, 497 (1995)].
[5] E. Wollan and W. Koehler, Phys. Rev. **100**, 545 (1955).
[6] G. Matsumoto, J. Phys. Soc. Jpn. **29**, 606, 615 (1970).
[7] H. Kawano, R. Kajimoto, M. Kuboto, and H. Yoshizawa, Phys. Rev. B **5**, 2202 (1996).
[8] M. Hennion, F. Moussa, J. Rodrigues-Carvajal, L. Pinsard, and A. Revcolevschi, Phys. Rev. B **56**, R497 (1997).
[9] E. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996).
[10] T. Perring, G. Aeppli, S. Hayden *et al.*, Phys. Rev. Lett. **77**, 711 (1996).
[11] E. L. Nagaev, Fiz. Tverd. Tela **11**, 2779 (1969) [Sov. Phys. Solid State **11**, 2249 (1969)]; Zh. Éksp. Teor. Fiz. **58**, 1269 (1970) [Sov. Phys. JETP **31**, 682 (1970)].
[12] J. Jung, K. Kim, D. Eom *et al.*, Phys. Rev. B **55**, 15489 (1997).
[13] K. I. Kugel' and D. I. Khomskii, Usp. Fiz. Nauk **136**, 621 (1982) [Sov. Phys. Usp. **25**, 231 (1982)].
[14] P. Anderson and A. Hasegawa, Phys. Rev. **100**, 675 (1955).
[15] E. L. Nagaev, Zh. Éksp. Teor. Fiz. **56**, 1013 (1969) [Sov. Phys. JETP **56**, 545 (1969)]; Fizika Metallov i Metallovedenie **6**, 214, 586 (1968); Fiz. Tverd. Tela **11**, 2779 (1969) [Sov. Phys. Solid State **11**, 2249 (1969)]; Physica Status Solidi **65**, 11 (1974); E. L. Nagaev and E. B. Sokolova, Fiz. Tverd. Tela **16**, 1293 (1974) [Sov. Phys. Solid State **16**, 836 (1974)].
[16] L. N. Bulaevskii, E. L. Nagaev, and D. I. Khomskii, Zh. Éksp. Teor. Fiz. **54**, 1562 (1968) [Sov. Phys. JETP **27**, 836 (1968).
[17] E. L. Nagaev, Zh. Éksp. Teor. Fiz. **57**, 1274 (1969) [Sov. Phys. JETP **30**, 693 (1969)]; Fiz. Tverd. Tela **14**, 773 (1972) [Sov. Phys. Solid State **14**, 658 (1972)].
[18] S. Yunoki, J. Hu, A. Malvezzi, A. Moreo, N. Furukawa, and E. Dagotto, Phys. Rev. Lett. **80**, 845 (1998).
[19] A. Moreo, D. Scalapino, and E. Dagotto, Phys. Rev. B **43**, 11442 (1991).
[20] S. Mori, C. Chen, and S.-W. Cheong, Nature (London) **392**, 473 (1998).
[21] A. Millis, Nature (London) **392**, 438 (1998).

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

Kinetics of a phase transition into an inhomogeneous state in a ferroelectric plate

B. M. Darinskiĭ

Voronezh State Technical University, 394026 Voronezh, Russia

A. P. Lazarev

Voronezh State University, 394693 Voronezh, Russia

A. S. Sigov*)

Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University), 117454 Moscow, Russia

(Submitted 5 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 2238–2245 (December 1998)

The kinetics of the emergence of a polarization vector field after a ferroelectric sample is rapidly cooled is investigated in two cases, in which the rate of growth of the polarization vector is determined by viscous forces of phonon origin or by the diffusion of charged particles. Analytic expressions are obtained for the rate of growth of the polarization vector and the period of the inhomogeneous phase as a function of the degree of supercooling of the sample. The existence of a slow relaxation process in the domain structure is established and the time dependence of this process is found. © *1998 American Institute of Physics.* [S1063-7761(98)02712-7]

1. INTRODUCTION

In Refs. 1–5 it is shown that when a homogeneous ferroelectric plate is cooled, a ferroelectric phase transition occurs from the paraphase into an inhomogeneous state, in which the spontaneous polarization vector is a periodic function of the plane-wave type, and which during subsequent cooling of the sample transforms into a plane-parallel domain structure. A similar phenomenon also occurs in sufficiently thin ferroelectric plates.[6] This phenomenon has been investigated for the case of slow (quasistatic) cooling of the sampling, for which the phase transition temperature and geometry of the resulting phase do not depend on the cooling rate.

In the present work the kinetics of a phase transition after a ferroelectric crystal sample is rapidly cooled is studied. The kinetics of the phase transition is investigated both for a ferroelectric crystal with no free charges and taking account of the screening of the depolarizing electric field by mobile charged particles.

Note that the processes leading to the formation of an inhomogeneous state of matter in bulk samples under strongly nonequilibrium conditions by the mechanism of spinodal decomposition of a homogeneous solid solution have been studied in previous work.[7–10] The dimensions of inhomogeneities of the emerging phase are determined by the bulk characteristics of the material, and do not depend on the macroscopic dimensions of the sample. For the systems studied in the present work the parameters of the inhomogeneity of the low-temperature phase typically also depend on the sample size and shape.

2. KINETICS OF THE FORMATION OF AN INHOMOGENEOUS STATE IN A FERROELECTRIC PLATE WITH NO FREE CHARGES

Consider a uniaxial ferroelectric crystal sample in the form of a thin plate of thickness l. The plane of the plate is perpendicular to the spontaneous polarization vector. The laboratory coordinate system $x_1x_2x_3$ in this uniaxial ferroelectric crystal plate is oriented so that the x_3 axis is perpendicular to the plane of the plate, while the axes x_1 and x_2 lie in the plane of the plate.

Initially, the sample is in the paraphase. After the sample is rapidly cooled, a periodic spontaneous polarization vector field P_s with two nonzero components—P_3, the projection of the spontaneous polarization vector on the x_3 axis, and P_1, the projection on the x_1 axis, which is oriented in the direction of the projection of the wave vector q determining the periodic structure of the polarization vector field—arises and grows in amplitude.

The nonequilibrium thermodynamic potential Φ of the ferroelectric plate can be represented in the form[4,5]

$$\Phi=\int\left[\frac{\alpha_1}{2}P_1^2-\frac{\alpha_3}{2}P_3^2+\frac{\kappa}{2}(\nabla_1 P_3)^2+\frac{(\nabla\varphi)^2}{8\pi}\right]dV. \quad (1)$$

In Eq. (1) α_1,α_3, and κ are coefficients in the series expansion of the thermodynamic potential in powers of the components of the polarization vector, all coefficients being positive. The temperature-dependent coefficient is α_3, where $\alpha_3=\alpha_0(T_c-T)$, T is the sample temperature, T_c is the Curie temperature, $\alpha_0=(2C)^{-1}$, and C is the Curie–Weiss constant.[11] The third term in Eq. (1) is the crystal energy

associated with the inhomogeneity of the polarization vector field and reflecting the correlation of the values of the vector P_3 at neighboring points with different coordinates x_1. In order of magnitude $\kappa \sim a^2$, where a is the lattice constant.[12] In the integral (1) the correlations between the polarization vectors at points with different coordinates x_3 are dropped; this is justified if the width of the domains is much less than the thickness of the plate. This is a typical situation for the present case of a uniaxial ferroelectric plate.[1–5] The last term in Eq. (1) is the energy of the depolarizing electric field and φ is the potential of this field. The low-temperature phase grows by the mechanism of instability of the high-temperature state of the material after the coefficient α_3 changes rapidly as a result of rapid cooling of the sample. To study the initial emergence of the new phase, it is therefore sufficient to include in the thermodynamic potential (1) only terms quadratic in the components of the polarization vector.[2–5]

In the bulk of the crystal the components P_i of the polarization vector and the potential φ of the depolarizing electric field are related by the electrostatic equation

$$\Delta\varphi = 4\pi\nabla_i P_i. \tag{2}$$

Combining Eqs. (1) and (2) by introducing a Lagrange multiplier $\mu(\mathbf{r})$ that depends on the radius vector $\mathbf{r}$ of the point of interest gives a new functional $\tilde{\Phi}$ in which P_1, P_3, and φ are independent variables:

$$\tilde{\Phi} = \Phi + \frac{\mu(\mathbf{r})}{4\pi}(\Delta\varphi - 4\pi\nabla_i P_i). \tag{3}$$

Varying $\tilde{\Phi}$ (3) with respect to the variables P_1 and P_3 gives the following expressions for the components E_1 and E_3 of the thermodynamic force on the components of the polarization vector:

$$E_1 = -\frac{\delta\tilde{\Phi}}{\delta P_1} = -\alpha_1 P_1 - \nabla_1\mu, \tag{4}$$

$$E_3 = -\frac{\delta\tilde{\Phi}}{\delta P_3} = \alpha_3 P_3 + \kappa\nabla_1^2 P_3 - \nabla_3\mu. \tag{5}$$

Varying Eq. (3) with respect to φ gives

$$\frac{\delta\tilde{\Phi}}{\delta\varphi} = -\frac{\Delta\varphi}{4\pi} + \frac{\Delta\mu}{4\pi} = 0, \tag{6}$$

whence $\varphi = \mu$.

If the rate of change of the spontaneous polarization vector is determined by viscous forces of phonon origin, the following system of kinetic equations is obtained:[13]

$$\gamma_1\dot{P}_1 = -\alpha_1 P_1 - \nabla_1\varphi, \tag{7}$$

$$\gamma_3\dot{P}_3 = \alpha_3 P_3 + \kappa\nabla_1^2 P_3 - \nabla_3\varphi, \tag{8}$$

where γ_1 and γ_3 are the coefficients of viscosity at the temperature of supercooling.

Equations (2), (7), and (8) constitute a complete set that makes it possible to determine the coordinate and time dependences of the electric potential φ and the polarization vector $\mathbf{P}$. These dependences can be represented as

$$P_j, \varphi \propto \exp(i(q_1 x_1 + q_3 x_3) + \nu t). \tag{9}$$

In Eq. (9) the wave vector q_1 determines the periodicity of the polarization vector field in the plane of the plate, as indicated above, and can be determined by requiring that the rate of growth of the low-temperature phase be maximum. The other wave number is determined by the condition $q_3 l = \pi$.[1,5] The rate of growth of the polarization field is found by equating the determinant of the system (7) and (8) to zero, taking account of Eq. (2):

$$\begin{vmatrix} \alpha_1 + \nu\gamma_1 + 4\pi q_1^2/q^2 & 4\pi q_1 q_3/q^2 \\ 4\pi q_1 q_3/q^2 & -\alpha_3 + \kappa q_1^2 + \nu\gamma_3 + 4\pi q_3^2/q^2 \end{vmatrix} = 0, \tag{10}$$

which is constructed from the coefficients of P_1 and P_3.

The condition (10) gives the function $\nu = \nu(q_1, q_3)$ in the form of the equation

$$\gamma_1\gamma_3\nu^2 + [\gamma_1(-\alpha_3 + \kappa q_1^2 + 4\pi q_3^2/q^2) + \gamma_3(\alpha_1 + 4\pi q_1^2/q^2)]\nu + 4\pi\alpha_1 q_3^2/q^2 + (\alpha_1 + 4\pi q_1^2/q^2)(-\alpha_3 + \kappa q_1^2) = 0. \tag{11}$$

For fixed supercooling α_3, the rate of growth of the amplitude of the periodic phase depends on the length of the period. The phase at which the period of the variation of the polarization vector corresponds to the maximum rate of growth will reach the region where the thermodynamic potential is a nonlinear function of the components of the polarization vector sooner than other phases. The plane-parallel domain structure formed in the ferroelectric plate therefore acquires just this period. However, the equation $\partial\nu/\partial q_1 = 0$ turns out to be too unwieldy for analysis. It is simpler to analyze the function $\alpha_3 = \alpha_3(\nu, q_1)$ given by

$$\alpha_3 = \kappa q_1^2 + \frac{4\pi(\alpha_1 + \gamma_1\nu)q_3^2}{(\alpha_1 + \nu\gamma_1 + 4\pi q_1^2/q^2)q^2} + \gamma_3\nu, \tag{12}$$

which is likewise obtained from Eq. (11).

The supercooling α_3 and the wavelength of the periodic phase with fixed ν can be found from the condition for a minimum $\partial\alpha_3/\partial q_1 = 0$, where $\alpha_3 = \alpha_3(q_1, \nu)$. A single extremal value $q_1 = 0$, corresponding to a transition into a homogeneous state, always exists for Eq. (12). The condition under which $q_1 = 0$ corresponds to a minimum of the function $\alpha_3(q_1)$ can be determined by expanding α_3 near the point $q_1 = 0$:

$$\alpha_3 = (\partial\alpha_3/\partial q_1^2)q_1^2 + 4\pi + \gamma_3\nu, \tag{13}$$

where

$$\frac{\partial\alpha_3}{\partial q_1^2} = \kappa - \frac{4\pi(\alpha_1 + \gamma_1\nu + 4\pi)}{(\alpha_1 + \gamma_1\nu)q_3^2}. \tag{14}$$

As indicated above, $\kappa \sim a^2$ and $q_3 = \pi/l$, as a result of which $\partial\alpha_3/\partial q_1^2 < 0$ in the case $l > a$ for all values of the rate of growth of the low-temperature phase. Thus, for macroscopic plates the point $q_1 = 0$ corresponds to the maximum

of the function $\alpha_3(q_1)$. Regardless of the degree of supercooling of the sample, the phase transition in a ferroelectric plate of macroscopic thickness therefore proceeds into an inhomogeneous state.

The wave number q_1 of the inhomogeneous ferroelectric phase, which precipitates with rate u after the sample is rapidly cooled, can be found by minimizing the function $\alpha_3(q_1)$. The condition $\partial\alpha_3/\partial q_1=0$ from Eq. (12) with $l\gg a$ and $q\simeq q_1$ gives q_1 in the form

$$q_1=(4\pi^3/\varepsilon_1^*\kappa l^2)^{1/4}, \tag{15}$$

where

$$\varepsilon_1^*=1+4\pi/(\alpha_1+\gamma_1\nu). \tag{16}$$

Substituting the expression (15) into Eq. (12) gives an expression for α_3:

$$\alpha_3=2(4\pi^3\kappa/\varepsilon_1^* l^2)^{1/2}+\gamma_3\nu. \tag{17}$$

For $\nu=0$, which corresponds to quasistatic cooling, the expressions (15) and (17) correspond to the well-known expressions for the period of the resulting ferroelectric phase and the displacement of the transition temperature from T_c.[1,5]

We note that different periods of variation of the polarization field of the ferroelectric phase arising can be obtained by varying the degree of supercooling of the sample. It follows from Eq. (16) that as ν increases, which corresponds to increasing α_3 (17) and increasing supercooling of the sample, the quantity ε_1^* decreases from $\varepsilon_1=1+4\pi/\alpha_1$ to 1. According to Eq. (15), as ε_1^* decreases, the wavelength λ of the polarization vector decreases, since $\lambda=2\pi/q_1$. The period of the domain structure formed in the ferroelectric plate at a phase transition thereby depends on the degree of supercooling of the sample, making it possible to control the domain structure of the ferroelectric plate by choosing the appropriate thermal regime of the sample.

Qualitative agreement between the decrease in domain sizes found on the basis of our theoretical model and an increasing degree of supercooling was obtained experimentally.[14,15] The patterns found can be further checked experimentally on ferroelectrics for which the paraphase state can be preserved after sufficiently rapid cooling of the sample. Examples of such ferroelectrics are those of the ordering type, in which a polar phase forms by means of thermal activation of atomic groups, so that the characteristic formation time of the low-temperature phase can be less than the cooling time of the plate.

In conclusion, we note that in the initial stage of growth of the low-temperature periodic phase, after rapid cooling of the sample, the absolute value of the rate of growth of the polarization vector is proportional to its initial values. These values are determined by fluctuations associated for the most part with the random character of the distribution of various defects of crystal structure in the bulk and at the surface of the sample. Since the fluctuations are characterized by a set of different spectral lines with differing values of the wave vector q_1, waves not only with wave numbers q_1 from Eq. (15) but also with close values of the wave number participate in the formation of the new phase. The width Δq_1 of the wave packet is determined from the condition $\Delta\nu\sim\nu$, whence

$$\Delta q=\frac{2\nu}{\partial^2\nu/\partial q_1^2}.$$

From Eq. (11) the curvature of the function $\nu(q_1)$ is $\partial^2\nu/\partial q_1^2\sim\kappa/\gamma$. Then $\Delta q\sim(\nu\gamma/\kappa)^{1/2}$, so that as the rate of precipitation of the phase increases, the width of the packet of waves comprising the domain structure increases. Since the domain structure is not periodic, a relaxation process associated with the displacement ξ_i of the domain walls commences after the structure is formed. This process is described by the equation

$$\dot{\xi}_i=\frac{\eta}{2}(\xi_{i+1}+\xi_{i-1}-2\xi_i), \tag{18}$$

where the coefficient η takes account of the interaction force between neighboring walls and viscous forces arising as the walls move. This relaxation process is characterized by a spectrum of relaxation times. This is due to the existence of a size spectrum of the regions relaxing to a periodic domain structure. The asymptotic time dependence of the relaxation process that is characteristic of diffusion processes is $\xi\propto t^{-1/2}$, so that the relaxation process is slow and is not completed under real conditions.

3. INFLUENCE OF THE DIFFUSION OF CHARGED PARTICLES ON THE KINETICS OF THE FORMATION OF A FERROELECTRIC PHASE IN A FERROELECTRIC PLATE

When charged particles, such as electrons in the conduction band and holes in the valence band, are present in the ferroelectric sample, the equations of electrostatics and kinetics of the charge have the form

$$\Delta\varphi=4\pi\nabla_i P_i-4\pi\rho, \tag{19}$$

$$\dot{\rho}=D\Delta\rho+(ne^2/kT)D\Delta\varphi. \tag{20}$$

Here e is the elementary charge of a particle, m is the average density of charged particles participating in the screening of the depolarizing electric field, ρ is the local volume density of electric charge, $D\Delta=D_1\Delta_1^2+D_3\Delta_3^2$, D_1 and D_3 are the principal values of the diffusion coefficient of the charged particles, and k is Boltzmann's constant. To solve this problem, the previously derived Eqs. (4) and (5) must be added to Eqs. (19) and (20), but written in the form

$$\alpha_1 P_1+\nabla_1\varphi=0, \tag{21}$$

$$\alpha_3 P_3-\kappa q_1^2 P_3-\nabla_3\varphi=0. \tag{22}$$

The solution of the system of equations (19)–(22) has the form (9). The rate of growth of the low-temperature phase after rapid cooling of the sample is determined by the equation

$$\nu=(D_1q_1^2+D_3q_3^2)$$
$$\times\frac{(\varepsilon_1/R^2+\varepsilon_1q_1^2+q_3^2)(\alpha_3-\kappa q_1^2)-4\pi q_3^2}{4\pi q_3^2-(\varepsilon_1q_1^2+q_3^2)(\alpha_3-\kappa q_1^2)}, \qquad (23)$$

where

$$R=(4\pi e^2n/\varepsilon_1kT)^{-1/2}$$

is the screening radius of the electric charge.

The function $\alpha_3(q_1)$ can be obtained from Eq. (23) in the form

$$\alpha_3=\kappa q_1^2+\frac{4\pi q_3^2}{\varepsilon_1q_1^2+q_3^2+(1+\nu/\hat{D})^{-1}\varepsilon_1R^{-2}}, \qquad (24)$$

where $\hat{D}=D_1q_1^2+D_3q_3^2$.

The value $\nu=0$ corresponds to quasistatic cooling of the sample, whereupon Eq. (23) yields an expression for α_3 in the form

$$\alpha_3=\kappa q_1^2+\frac{4\pi q_3^2}{\varepsilon_1(q_1^2+R^{-2})}, \qquad (25)$$

which is identical to the analogous expression for α_3 in Refs. 3 and 16. The expression (25) is written for the case $l\gg R$.

The minimization condition $\partial\alpha_3/\partial q_1=0$ yields

$$q_1^{*2}=2\pi(\pi/\varepsilon_1\kappa)^{1/2}l^{-1}-R^{-2}=q_1^2-R^{-2}, \qquad (26)$$

where q_1 corresponds to Eq. (15). It follows from Eq. (26) that an inhomogeneous phase $q_1^*>0$ arises for plate thicknesses l satisfying

$$l^2<4\pi^3R^4/\varepsilon_1\kappa=l_*^2. \qquad (27)$$

It follows from (26) and (27) that thin enough plates, in the sense (27), always transform into an inhomogeneous state by the phase-transition mechanism considered above, while thick plates can undergo a transition into a homogeneous state.

The quasistatic supercooling of a sample for which a polarized state arises is determined by the following expressions obtained by substituting Eq. (26) into Eq. (25):

$$\alpha_3=\alpha_*-\frac{\kappa}{R^2}, \quad \alpha_*=\frac{4\pi}{l}\left(\frac{\pi\kappa}{\varepsilon_1}\right)^{1/2}, \quad l<l_*, \qquad (28)$$

$$\alpha_3=4\pi^3R^2/\varepsilon_1l^2, \quad l>l_*, \quad q_1^*=0. \qquad (29)$$

For significant supercooling, an inhomogeneous ferroelectric phase arises rapidly and the process is not accompanied by a redistribution of the electric charges. The expressions for q_1 and q_3 are determined in this case by Eq. (24) with $\nu=\infty$ or by Eqs. (26) and (28) with the equivalent condition $R=\infty$. For samples which are thin in the sense (27) and for supercooling in the range $(\alpha_*-\kappa/R^2,\alpha_*)$, the wave vector q_1 varies over the range $((q_1^2-R^{-2})^{1/2},q_1)$.

In thick films, an increase in supercooling from the value (28) at first preserves the transition into the homogeneous state. However, the temperature range for transitions into the homogeneous phase is narrow. The value of α_3 for which a transition into the inhomogeneous phase starts is determined by a relation obtained by setting the derivative $\partial\nu/\partial q_1$ from Eq. (24) to zero. This yields

$$\alpha_3=4\pi^3R^2/\varepsilon_1l^2+\pi^2\kappa D_3/D_1l^2. \qquad (30)$$

The expression (30) shows that the supercooling for which an inhomogeneous stage arises in thick films differs from Eq. (29) by a small quantity $\sim a/l$.

As the supercooling of the sample increases further, the resulting low-temperature phase is inhomogeneous.

For a high rate of formation of the phases $\nu\gg D_1q_1^2+D_3q_3^2$, it follows from Eq. (24) that

$$\alpha_3=\kappa q_1^2+\frac{4\pi q_3^2}{\varepsilon_1q_1^2(1+D_1/\nu R^2)}. \qquad (31)$$

The minimization condition $\partial\alpha_3/\partial q_1=0$ from Eq. (31) yields expressions for q_1 and α_3:

$$q_1^2=\frac{2\pi}{l}\left(\frac{\pi}{\varepsilon_1\kappa(1+D_1/\nu R^2)}\right)^{1/2}, \qquad (32)$$

$$\alpha_3=\frac{4\pi}{l}\left(\frac{\pi\kappa}{\varepsilon_1(1+D_1/\nu R^2)}\right)^{1/2}. \qquad (33)$$

A comparison of (32) and (33) yiedls a simple relation between α_3 and q_1 in the range of supercooling studied,

$$q_1=(\alpha_3/\kappa)^{1/2}. \qquad (34)$$

4. CONCLUSIONS

The following conclusions can be drawn from the above investigation of the kinetics of the phase transition in a ferroelectric plate.

In the case of rapid cooling of a macroscopic ferroelectric plate in which there are no screening charges, the material transforms into an inhomogeneous phase for all degrees of supercooling. The period of the inhomogeneity depends on the ratio of the viscosity coefficients as the longitudinal and transverse components of the polarization vector vary, and it decreases with increasing supercooling of the sample. The period of the resulting phase can be varied severalfold by varying the supercooling of the sample.

Ferroelectric plates containing mobile charge carriers can be divided into thick and thin plates according to the criterion (27). Upon cooling, thin plates always pass into an inhomogeneous ferroelectric state. Thick plates subjected to suffiiciently small supercooling pass, via the phase-transition mechanism studied, into a uniformly polarized state. As the degree of supercooling increases, the emergent phase becomes inhomogeneous, and its period decreases, as follows, for example, from Eq. (34).

As a result of the existence of fluctuations with differing wave vectors, an aperiodic domain structure will emerge in a sample after supercooling. This initiates a slow ongoing process whereby the domain structure relaxes to a periodic structure.

This work was supported by the Russian Fund for Fundamental Research (Project No. 96-02-18774), and in part by the International Soros Program for Education in the Exact Sciences.

*)E-mail: sigov@exodus.mirea.ac.ru

[1] E. V. Chenskiĭ and V. V. Tarasenko, Zh. Éksp. Teor. Fiz. **83**, 1083 (1982) [Sov. Phys. JETP **56**, 618 (1982)].
[2] B. M. Darinskii, A. P. Lazarev, and A. S. Sidorkin, Ferroelectrics **98**, 193 (1989).
[3] B. M. Darinskiĭ, A. P. Lazarev, and A. S. Sidorkin, Kristallografiya **36**, 757 (1991) [Sov. Phys. Crystallogr. **36**, 423 (1991)].
[4] B. M. Darinskiĭ, A. P. Lazarev, and A. S. Sidorkin, Fiz. Tverd. Tela (St. Petersburg) **35**, 1942 (1993) [Phys. Solid State **35**, 969 (1993)].
[5] B. M. Darinskiĭ, A. P. Lazarev, and A. S. Sigov, Izv. Ross. Akad. Nauk, Ser. Fiz. **60**, 186 (1996).
[6] B. M. Darinskiĭ and A. P. Lazarev, Izv. Ross. Akad. Nauk, Ser. Fiz. **59**, 37 (1995).
[7] J. W. Cahn and J. E. Hilliard, J. Chem. Phys. **28**, 258 (1958).
[8] L. I. Stefanovich and É. P. Fel'dman, Zh. Éksp. Teor. Fiz. **104**, 2774 (1993) [JETP **77**, 273 (1993)].
[9] A. G. Khachaturyan, *Theory of Phase Transformations and the Stucture of Solid Solutions* [in Russian], Nauka, Moscow (1974).
[10] V. G. Vaks, S. V. Beicler, and V. Yu. Dobretsov, JETP Lett. **61**, 68 (1995).
[11] G. A. Smolenskiĭ, V. A. Bokov, V. A. Isupov *et al.*, *The Physics of Ferroelectric Phenomena* [in Russian], Nauka, Leningrad (1985), p. 35.
[12] V. A. Zhirnov, Zh. Éksp. Teor. Fiz. **35**, 1175 (1958) [Sov. Phys. JETP **8**, 822 (1959)].
[13] A. P. Lazarev, A. S. Sidorkin, and V. N. Fedosov, Zh. Tekh. Fiz. **52**, 1623 (1982) [Sov. Phys. Tech. Phys. **27**, 993 (1982)].
[14] N. Tomita, H. Orihara, and Y. Ishibashi, J. Phys. Soc. Jpn. **58**, 1190 (1989).
[15] B. M. Darinskiĭ, A. A. Krivonos, A. P. Lazarev *et al.*, Izv. Ross. Akad. Nauk, Ser. Fiz. **57**, 75 (1993).
[16] A. S. Sidorkin, B. M. Darinskiĭ, and A. S. Sigov, Fiz. Tverd. Tela (St. Petersburg) **39**, 922 (1997) [Phys. Solid State **39**, 823 (1997)].

Translated by M. E. Alferieff

Lateral superlattices in a strong electromagnetic field: self-induced transparency, multistability, and frequency multiplication

E. P. Dodin, A. A. Zharov, and A. A. Ignatov

Institute of Microstructure Physics, Russian Academy of Sciences, 603600 Nizhniĭ Novgorod, Russia
(Submitted 21 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 2246–2262 (December 1998)

Self-action and generation of the harmonics of intense electromagnetic radiation interacting with quantum lateral semiconductor superlattices are investigated theoretically using a self-consistent approach. The calculations are based on a semiclassical description of electron transport in a conduction miniband in an electrodynamic model of an equivalent current screen. Multistep, multistable, self-induced transparency is obtained. The effect is associated with the dynamic localization of electrons in the self-consistent electromagnetic field and it appears only if the density of free current carriers in the superlattice is higher than a certain critical value. It is shown that the bleaching of the structure is accompanied by efficient generation of odd-numbered radiation harmonics, whose intensity peaks lie near the transparency threshold. © *1998 American Institute of Physics.* [S1063-7761(98)02812-1]

1. INTRODUCTION

The nonlinear interaction of quantum periodic semiconductor structures — quantum superlattices — with electromagnetic radiation is an important direction of research in modern quantum electronics. Above all, this is due to the latest technological advances in the preparation of extremely high-quality periodic structures demonstrating the very strongest nonlinear properties in millimeter, submillimeter, infrared, and optical electromagnetic fields.[1–4] The additional potential produced in quantum superlattices by the periodic variation of the chemical composition of the sample gives rise to restructuring of the energy spectrum of the current carriers. The energy spectrum breaks up into a set of narrow allowed and forbidden minibands.[5,6] The nonlinearity of the superlattice is due to the fact that the electron energy in a miniband is bounded, which gives rise to, on the one hand, an oscillatory motion of electrons in a constant magnetic field (Bloch oscillations) and, on the other, a strongly nonquadratic dispersion law, a consequence of which is an *N*-shaped current–voltage characteristic.[7]

These two circumstances in fact determine the mechanism of the nonlinearity of quantum superlattices. In a rapidly varying electric field the dynamics of the electrons in a miniband becomes very complicated and can be described analytically only in special cases, for example, in the limit of very infrequent collisions. Thus, if the frequency of the field is much higher than the frequency of elastic and inelastic electron scattering, there occurs a dynamic localization of electrons[8,9] under the conditions of which the superlattice behaves as a linear dielectric. A macroscopic manifestation of dynamic localization is so-called self-induced transparency. When ac and dc electric fields are present simultaneously, regions of absolute negative conductivity can appear in the static current–voltage characteristic,[9] showing that superlattices have active properties in these regions. Therefore the nonlinear and active properties of quantum superlattices can in principle be used to transform the frequency and the spatial spectra of the radiation incident on them, to amplify the radiation, and to generate secondary waves in the regions of absolute negative conductivity.

In the present paper we study the self-action and harmonic generation associated with the dynamic localization of electrons accompanying the reflection and transmission of an electromagnetic plane wave through a lateral superlattice. Lateral (planar) superlattices, where an additional potential is produced for two-dimensional electrons localized near a surface of the semiconductor, are now attracting considerable attention. In the present situation the electron motion is localized in a direction perpendicular to the surface, and at the same time the additional potential leads to the formation of a miniband spectrum for electron motion in the localization plane. Superlattices consisting of one-dimensional chains of identical and identically coupled GaAs/AlGaAs quantum dots (Fig. 1a)—a form of quantum wires and in the present geometry sometimes called quantum boxes[10]—occupy a special place among various lateral structures. In such a one-dimensional system electrons move along wires whose periodic potential forms a miniband energy spectrum. It is shown in Ref. 10 that such structures have unique properties associated with the possibility of the suppression of both intra- and interminiband scattering by optical phonons at room temperature, as a result of which the corresponding relaxation rates of the electron velocity can reach anomalously low values $\simeq 3\times 10^{11}\ \mathrm{s}^{-1}$. Elastic electron scattering is also found to be considerably suppressed. As will be shown below, lateral quantum superlattices of this type in the field of an incident wave demonstrate complicated dynamical behavior leading to multistep and multistable self-induced transparency accompanied by efficient generation of odd-numbered harmonics. In contrast to the articles cited above[8,9] and Ref. 11, where the fixed monochromatic field approximation was used on the basis of quasielectrostatics, in the

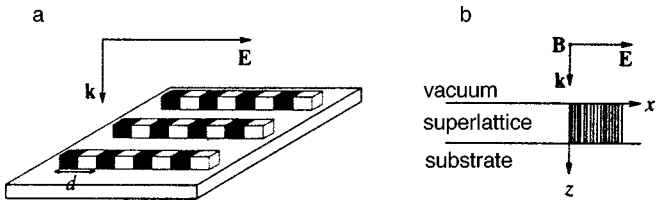

FIG. 1. a) Example of a lateral GaAs/AlGaAs semiconductor structure (one-dimensional systems of coupled quantum dots); b) geometry of the model formulation of the problem.

present paper a self-consistent wave approach taking account of the influence of the nonlinear screening current flowing along the axis of a lateral superlattice is developed.

2. HYDRODYNAMICS OF ELECTRONS IN A CONDUCTION MINIBAND. MATERIAL EQUATIONS

The equations of electron motion in an energy miniband of a superlattice can be derived using a semiclassical description of electron transport on the basis of the Boltzmann equation[9]

$$\frac{\partial f}{\partial t}+eE\frac{\partial f}{\partial p}=\mathrm{St}(f), \tag{1}$$

where $f(p,t)$ is the distribution function, p is the electron quasimomentum, E is the electric field applied along the axis (perpendicular to the layers) of the superlattice, St (f) is the collision integral, and the electron dispersion relation in the tight-binding approximation is

$$\epsilon(p)=\frac{\Delta}{2}\left(1-\cos\frac{pd}{\hbar}\right), \tag{2}$$

where Δ is the energy width of the miniband and d is the superlattice period. In writing down the expression (2) we assumed that all electrons are concentrated in the bottom miniband (the criterion for this to occur is given in Ref. 5). Semiclassical equations can be used if

$$\hbar\omega<\Delta_g, \quad eEd<\Delta_g, \quad \text{and} \quad eEd<\Delta,$$

where Δ_g is the width of the forbidden miniband. The first two conditions make it possible to neglect interminiband transitions and interminiband tunneling under the action of an intense field. The third condition makes it possible to treat the electron motion within a miniband semiclassically, though as shown recently[12] this condition is not necessary in some cases (*sequential tunneling*).

To take account of elastic and inelastic scattering, we introduce a collision integral in the model form

$$\mathrm{St}(f)=-\nu_\epsilon\{f(p,t)-f_0(p)\}-\frac{\nu_{el}}{2}\{f(p,t)-f(-p,t)\}, \tag{3}$$

where

$$f_0(p)=\frac{d}{2}\left\{\pi\hbar I_0\left(\frac{\Delta}{2kT}\right)\right\}^{-1}\exp\left\{\frac{\Delta\cos(pd/\hbar)}{2kT}\right\}$$

is the equilibrium distribution function, T is the lattice temperature, ν_ϵ is the effective energy relaxation rate, ν_{el} is the effective elastic collision frequency, $I_0(x)$ is a modified Bessel function, and k is Boltzmann's constant. The semiclassical instantaneous electron velocity $v(p)$ along the axis of the superlattice is a periodic function of the quasimomentum:

$$v(p)=\frac{\partial\epsilon}{\partial p}=\frac{d\Delta}{2\hbar}\sin\left(\frac{pd}{\hbar}\right). \tag{4}$$

To construct a theory of the interaction of radiation with a superlattice, it is necessary to obtain from the Boltzmann equation (1) the material equations for the electromagnetic field or, in other words, a relation between the current flowing through the superlattice and the electric field. The simplest way to do this is to transform to the hydrodynamic equations. The average (hydrodynamic) velocity V and energy W of the electrons are found by calculating the corresponding moments of the distribution function

$$V(t)=\int v(p)f(p,t)\,dp, \tag{5}$$

$$W(t)=\int \epsilon(p)f(p,t)\,dp, \tag{6}$$

where the integrals extend over the Brillouin zone $-\pi\hbar/d\leqslant p\leqslant\pi\hbar/d$, and the perturbed and unperturbed distribution functions satisfy the normalization conditions $\int f dp=1$ and $\int f_0 dp=1$. The equations for $V(t)$ and $W(t)$ are obtained by directly integrating Eq. (1), using Eq. (4) and the relation

$$\frac{\partial v(p)}{\partial p}=\frac{\Delta d^2}{2\hbar^2}\left[1-\frac{2\epsilon(p)}{\Delta}\right], \tag{7}$$

which follows directly from Eq. (4). After simple calculations we obtain

$$\frac{dV}{dt}=\frac{eE}{m(W)}-\nu_V V,$$

$$\frac{dW}{dt}=eEV-\nu_\epsilon(W-W_T), \tag{8}$$

where $W_T=\Delta(1-\mu_0)/2$ is the average thermal energy of an electron in the absence of an electric field, $\mu_0=I_1(\Delta/2kT)/I_0(\Delta/2kT)$, $\nu_V=\nu_\epsilon+\nu_{el}$ is the collisional relaxation rate of the average velocity, $m(W)=m_0/(1-2W/\Delta)$ is the energy dependence of the electron effective mass, and $m_0=2\hbar^2/d^2\Delta$ is the effective mass at the miniband bottom. The energy dependence of the electron effective mass is determined by the dispersion relation (2); physically, it is related to Bragg reflection. In what follows, we assume that the density of miniband conduction electrons is not too low, and that the polarizability of the superlattice is determined completely by the corresponding electric currents whose density is proportional to the hydrodynamic velocity V:

$$j=en_eV, \tag{9}$$

where e is the electron charge and n_e is the electron density converted to the entire surface of the structure. Thus, the

equations (8) can be treated as the material equations for an electromagnetic field polarized in the direction of the superlattice axis.

3. FORMULATION OF THE PROBLEM. INITIAL EQUATIONS

Consider a lateral superlattice of thickness h on a dielectric substrate with permittivity ε_s (in what follows we neglect the dispersion in the substrate, setting $\varepsilon_s=$ const). Let an electromagnetic plane wave with electric field vector oriented in the direction of the superlattice axis be incident normally in the direction from the vacuum onto the superlattice (the geometry of the problem is shown schematically in Fig. 1b). We assume that the superlattice thickness h is much less than the characteristic scale of the electromagnetic field in the medium. Very simple estimates show that this condition is satisfied conservatively for essentially any real structure in the millimeter, submillimeter, and far-infrared wavelength ranges. For definiteness we orient the z axis of a Cartesian coordinate system along the incident wave vector $\mathbf{k}$, and the x axis in the direction of the superlattice axis. We position the origin $z=0$ on the ''illuminated'' surface of the superlattice. We describe the electromagnetic field using Maxwell's equations

$$\nabla\times\mathbf{E}=-\frac{1}{c}\frac{\partial\mathbf{B}}{\partial t},\quad \nabla\times\mathbf{B}=\frac{1}{c}\frac{\partial\mathbf{E}}{\partial t}+\frac{4\pi}{c}\mathbf{j}, \tag{10}$$

where $\mathbf{j}$ is the electric current density in the corresponding medium. Here $\mathbf{j}=0$ for $z<0$ (vacuum). According to Eq. (9), $\mathbf{j}=en_e\mathbf{V}$ for $0\leqslant z\leqslant h$, and

$$\mathbf{j}=\frac{1}{c}\frac{\partial\mathbf{P}}{\partial t}=\frac{\varepsilon_s-1}{4\pi}\frac{\partial\mathbf{E}}{\partial t}\quad \text{for}\quad z>h$$

is the polarization current in the substrate. The media are assumed to be nonmagnetic ($\mathbf{B}=\mathbf{H}$).

Before deriving the equations describing electron dynamics in the superlattice in the self-magnetic field and the equations for the reflection and transmission coefficients for the electromagnetic field, for convenience we transform in Eqs. (8) and (10) to the dimensionless variables

$$\tau=\omega t,\quad \eta=\frac{\omega z}{c},\quad U=\frac{E}{E_C},\quad w=\frac{V}{V_0},$$

$$E_C=\frac{\hbar\omega}{ed},\quad V_0=\frac{d\Delta}{2\hbar},\quad \zeta=\frac{2(W-W_T)}{\Delta},$$

where ω is the frequency of the incident field. Since the problem is one-dimensional, we write in the new variables the wave equation describing the high-frequency electric field over all space

$$\varepsilon(\eta)\frac{\partial^2U}{\partial\tau^2}-\frac{\partial^2U}{\partial\eta^2}=-\theta\frac{\partial w}{\partial\tau}, \tag{11}$$

where $\theta=\omega_{pe}^2/\omega^2$, $\omega_{pe}^2=4\pi e^2n_e/m_0$ is the square of the electron plasma frequency in the miniband; $\varepsilon(\eta<\omega h/c)=1$, $\varepsilon(\eta>\omega h/c)=\varepsilon_s$, and $\theta(\eta<0,\eta>\omega h/c)=0$. To obtain a relation between the fields of the incident, reflected, and transmitted waves, taking into consideration the small thickness of the superlattice and neglecting diffraction effects due to the discreteness of the structure in the y direction, we treat the structure as an equivalent current screen of infinitesimal thickness, transforming from volume electric currents flowing through the superlattice to equivalent surface currents. We set the right-hand side of Eq. (11) to

$$\theta\frac{\partial w}{\partial\tau}=\theta_0\frac{\partial w}{\partial\tau}\delta(\eta), \tag{12}$$

where $\theta_0=\theta\omega h/c$ and $\delta(\eta)$ is the Dirac delta function.

The dimensionless parameter θ_0 is the key parameter of the theory. It determines the role played by nonlinear screening currents in the superlattice and, as will be shown below, greatly influences the self-action and harmonics generation processes. A numerical estimate of the parameter θ_0, which is proportional to the surface electron density ($N_s=n_eh$), for $n_e=10^{17}$ cm^{-3}, $h=10^{-5}$ cm, $\Delta=100$ meV, $d=10^{-6}$ cm, $\omega=10^{13}$ s^{-1} gives $\theta_0\sim1$. The boundary conditions for the electric field at the current screen can be obtained from Eqs. (11) and (12):

$$[U]_{\eta=0}=0,\quad \left[\frac{\partial U}{\partial\eta}\right]_{\eta=0}=\theta_0\frac{\partial w}{\partial\tau}, \tag{13}$$

corresponding to continuity of the electric field and jumps in the magnetic field at the surface current (the brackets denote jumps in the enclosed quantity). To substitute the incident (U_i), reflected (U_r), and transmitted (U_t) electromagnetic fields into the boundary conditions (13), we describe the fields separately using first-order wave equations

$$\begin{aligned}&\frac{\partial U_i}{\partial\tau}+\frac{\partial U_i}{\partial\eta}=0,\\ &\frac{\partial U_r}{\partial\tau}-\frac{\partial U_r}{\partial\eta}=0,\\ &\frac{\partial U_t}{\partial\tau}+\frac{1}{\sqrt{\varepsilon_s}}\frac{\partial U_t}{\partial\eta}=0,\end{aligned} \tag{14}$$

where the fields U_i, U_r, and U_t are related at $\eta=0$ by the condition (13).

For homogeneous linear media the transition from the initial equation (11) to Eqs. (14) is completely correct. Substituting the corresponding quantities $U_{i,r,t}$ into Eq. (13) and expressing the space derivatives in terms of the time derivatives from Eqs. (14), we obtain the explicit form of the boundary conditions at $\eta=0$:

$$U_i+U_r=U_t, \tag{15}$$

$$-\sqrt{\varepsilon_s}\frac{\partial U_t}{\partial\tau}-\frac{\partial U_r}{\partial\tau}+\frac{\partial U_i}{\partial\tau}=\theta_0\frac{\partial w}{\partial\tau}. \tag{16}$$

The equation (16) can be integrated over time. We obtain

$$-\sqrt{\varepsilon_s}\,U_t-U_r+U_i=\theta_0w+C. \tag{17}$$

The arbitrary integration constant is determined by the initial conditions. In what follows, we set $C=0$, which corresponds to vanishing initial conditions. We note that the electric field in Eqs. (8) has the same magnitude as the field of the transmitted wave, which in the renormalized variables corre-

sponds to U_t (in what follows the index t will be dropped: $U_t \equiv U$). Expressing U_r from Eq. (15) and substituting into Eq. (17), we have finally

$$\frac{2U_i(\tau)}{1+\sqrt{\varepsilon_s}} = U + \frac{\theta_0}{1+\sqrt{\varepsilon_s}}\, w, \tag{18}$$

which, together with Eqs. (8)—which in dimensionless variables are

$$\frac{\partial w}{\partial \tau} = U(\mu_0 - \zeta) - \nu_1 w,$$

$$\frac{\partial \zeta}{\partial \tau} = Uw - \nu_2 \zeta, \tag{19}$$

($\nu_{1,2} = \nu_{V,\epsilon}/\omega$)—completely describe the self-consistent interaction of a normally incident electromagnetic plane wave with a thin lateral superlattice.

The incident field $U_i(\tau)$ is assumed to be a given function of time. After making the substitutions

$$\frac{2U_i(\tau)}{1+\sqrt{\varepsilon_s}} \equiv \overline{U}_i(\tau), \quad \frac{\theta_0\mu_0}{1+\sqrt{\varepsilon_s}} \equiv \overline{\theta}_0,$$

$$\zeta = \mu_0\overline{\zeta}, \quad w = \mu_0\overline{w}$$

(in what follows the overbars are omitted), Eqs. (18) and (19) can be rewritten in the simpler form

$$\dot{w} = U(1-\zeta) - \nu_1 w,$$

$$\dot{\zeta} = Uw - \nu_2 \zeta, \tag{20}$$

$$U_i(\tau) = U + \theta_0 w.$$

Our main objective in this paper is to investigate the solutions of Eqs. (20). As one can see from this system, the fixed-field approximation employed in Refs. 9 and 11 corresponds to $\theta_0 = 0$, for which there is no effect due to the nonlinear screening current $\sim \theta_0 w$.

4. INTERACTION OF ELECTROMAGNETIC PLANE WAVES WITH A LATERAL QUANTUM SUPERLATTICE

Linear interaction

Let a monochromatic plane wave

$$U_i(\tau) = U_0 e^{i\tau} + \text{c.c.},$$

be incident from the vacuum onto the structure of interest. The amplitude U_0 of this wave is small enough that Eqs. (20) can be linearized (the corresponding criterion for the amplitude will be obtained below). After some obvious transformations, we obtain the following expression for the energy reflection and transmission coefficients (in terms of the energy flux density):

$$R = \frac{\left[1-\sqrt{\varepsilon_s} - \nu_1 \dfrac{\theta_0(1+\sqrt{\varepsilon_s})}{1+\nu_1^2}\right]^2 + \dfrac{\theta_0^2(1+\sqrt{\varepsilon_s})^2}{1+\nu_1^2}}{\left[1+\sqrt{\varepsilon_s} + \nu_1 \dfrac{\theta_0(1+\sqrt{\varepsilon_s})}{1+\nu_1^2}\right]^2 + \dfrac{\theta_0^2(1+\sqrt{\varepsilon_s})^2}{1+\nu_1^2}}, \tag{21}$$

$$T = \frac{4\sqrt{\varepsilon_s}}{(1+\sqrt{\varepsilon_s})^2} \frac{1+\nu_1^2}{(\nu_1+\theta_0)^2+1}. \tag{22}$$

Equation (21) and (22) show that in the linear approximation the superlattice behaves like an ordinary plasma film (whose plasma frequency is determined by the density and effective mass of the conduction electrons) on the surface of a dielectric with permittivity ε_s. The film makes the main contribution to reflection if the parameter θ_0 is sufficiently large,

$$\theta_0^2 > (1+\nu_1^2)\left(\frac{1-\sqrt{\varepsilon_s}}{1+\sqrt{\varepsilon_s}}\right)^2 = \theta_0^{*2},$$

and for $\theta_0 \gg \theta_0^*$ the film provides essentially perfect radiation shielding.

Transient processes in strong fields

Let us consider the excitation of the superlattice from vanishing initial conditions by an incident harmonic field. Expressing U from the last relation of the system (20) and substituting into the two other equations, we obtain

$$\dot{w} = U_i(1-\zeta) + \theta_0\zeta w - (\nu_1+\theta_0)w,$$

$$\dot{\zeta} = U_i w - \theta_0 w^2 - \nu_2\zeta. \tag{23}$$

For definiteness we consider the situation of greatest interest, $\theta_0 \gg 1, \nu_1$. Initially, when $\theta_0\langle\zeta\rangle \ll \theta_0$ ($\langle\zeta\rangle$ is the electron energy averaged over one period), the electrons are not heated and the time derivatives in the first of Eqs. (23) can be neglected (the velocity is established in a time $\tau_w \sim 1/\theta_0 \ll 1$ determined by radiation losses). As a result, we have

$$w \approx \frac{U_i(1-\zeta)}{\nu_1 + \theta_0(1-\zeta)}. \tag{24}$$

Let us substitute the expression (24) into the second of Eqs. (23), which gives

$$\dot{\zeta} = \frac{\nu_1 U_0^2(1-\zeta)}{\{\nu_1+\theta_0(1-\zeta)\}^2} - \nu_2\zeta. \tag{25}$$

Since $\zeta \ll 1$ by assumption, the expression (25) simplifies to

$$\dot{\zeta} \approx \frac{\nu_1 U_i^2(\tau)}{\theta_0^2} - \nu_2\zeta, \tag{26}$$

where we choose U_i in the form $U_i(\tau) = U_m \sin\tau$. An expression for the period-averaged electron energy follows from Eq. (26):

$$\langle\zeta\rangle = \frac{\nu_1 U_m^2}{2\nu_2\theta_0^2}[1-\exp(-\nu_2\tau)]. \tag{27}$$

Therefore the characteristic electron heating time is $\tau_\zeta \sim 1/\nu_2 \gg \tau_w$.

The threshold amplitude of the incident field at which nonlinear effects first appear can be estimated. The maximum average electron energy following from Eq. (27) should be close to 1:

$$\langle \zeta \rangle_{\max} = \frac{\nu_1 U_m^2}{2\nu_2 \theta_0^2} \sim 1. \tag{28}$$

The expression (28) gives us a corresponding estimate of the amplitude of the incident field:

$$U_m = U_m^{(C)} = \sqrt{\frac{2\nu_2}{\nu_1}}\ \theta_0. \tag{29}$$

As the electrons in the miniband are heated, their energy undergoes small oscillations about the mean at twice the frequency of the pump field. The amplitude of these oscillations can be easily obtained from Eq. (26):

$$\zeta_{\sim} = \frac{\nu_1 U_m^2}{4\theta_0^2}. \tag{30}$$

The electron heating-induced change in the transmission coefficient for radiation passing through the structure can be taken into account on the basis of the approximations employed. For this, the substitution

$$\theta_0 \to \theta_0 \left(1 - \frac{U_m^2}{(U_m^{(C)})^2} \right) \tag{31}$$

must be made in Eq. (22). Then it follows that as $U_m \to U_m^{(C)}$ the screening action of the superlattice becomes negligible, and self-induced transparency due to dynamic electron localization occurs.[8] Strictly speaking, these estimates do not hold for incident radiation amplitudes close to $U_m^{(C)}$, but they correctly describe the tendency for the transmission coefficient to increase with U_m and the characteristic times over which the corresponding interaction regimes are established. We shall study these interaction regimes in greater detail.

Steady-state periodic nonlinear interaction regimes

We now return to Eqs. (20), where the amplitude of the incident field can be large enough that the dynamical system under study can demonstrate strongly nonlinear behavior. We also assume that the system ''forgets'' the initial conditions (a situation brought about by collisions, which destroy the coherence of the electron motion) and undergoes periodic motion in the harmonic field of the incident wave. We study the approximate situation in which the relaxation times of the velocity and energy are equal ($\nu_1 = \nu_2 \equiv \nu$). (This will enable us to obtain certain analytic relations without qualitatively destroying the dynamics of the system. In addition, in reality, the frequencies ν_1 and ν_2 are, as a rule, close to one another, $|\nu_1 - \nu_2| \ll \nu_{1,2}$.) Then an exact expression for the velocity w can be obtained from the first two equations (20) with an arbitrary function $U(\tau)$:

$$w(\tau) = \nu \int_{-\infty}^{\tau} \exp[\nu(\tau' - \tau)] \sin\left(\int_{\tau'}^{\tau} U(\tau'')\, d\tau'' \right) d\tau'. \tag{32}$$

Substituting the expression (32) into the last of Eqs. (20) and making an obvious substitution of variables in the integrand, we obtain

$$U_i(\tau) = U(\tau) + \nu\theta_0 \int_0^{\infty} e^{-\nu\xi} \sin\left(\int_{\tau-\xi}^{\tau} U(\xi')\, d\xi' \right) d\xi. \tag{33}$$

As a result, we have reduced the original problem to the integral equation (33) for the field $U(\tau)$ in the superlattice. In the general case, Eq. (33) cannot be solved analytically. However, if the function $U(\tau)$ is assumed to be periodic in the field of the incident wave $U_i(\tau)$ and the amplitudes of the high harmonic components of $U(\tau)$ are sufficiently small compared with the amplitude of the fundamental harmonic, Eq. (33) can be solved by substituting into the integrand the first harmonic of $U(\tau)$, which for definiteness we choose in the form

$$U_1(\tau) = U_1 \cos\tau. \tag{34}$$

We now represent the complete expression for $U(\tau)$ in the form of a harmonic Fourier series

$$U(\tau) = U_1(\tau) + \sum_{k=1}^{\infty} a_{2k+1} \sin(2k+1)\tau + \sum_{k=1}^{\infty} b_{2k+1} \cos(2k+1)\tau. \tag{35}$$

Likewise, we write the integral on the right-hand side of Eq. (33) as the Fourier series

$$\nu\theta_0 \int_0^{\infty} e^{-\nu\xi} \sin\left\{ \int_{\tau-\xi}^{\tau} U_1(\xi')\, d\xi' \right\} d\xi = \sum_{p=1}^{\infty} A_p \sin p\tau + \sum_{p=1}^{\infty} B_p \cos p\tau. \tag{36}$$

Straightforward but unwieldy transformations yield expressions for the expansion coefficients A_p and B_p:

$$A_p = 2\pi\nu\theta_0 \int_0^{\infty} e^{-\nu\xi} \sin\left(\frac{p\xi}{2}\right) J_p\left(2U_1 \sin\frac{\xi}{2}\right) d\xi,$$

$$B_p = 2\pi\nu\theta_0 \int_0^{\infty} e^{-\nu\xi} \cos\left(\frac{p\xi}{2}\right) J_p\left(2U_1 \sin\frac{\xi}{2}\right) d\xi, \tag{37}$$

where $p = 2k+1$, $k = 0,1,2,3,\ldots$, and $J_p(x)$ are Bessel functions. Since the integrands in the expressions for the coefficients A_p and B_p in Eq. (37) contain the product of an exponential and a periodic function, the corresponding infinite integrals can be written as integrals over a period:

$$A_p = \frac{2\pi\nu\theta_0}{1 - \exp(-4\pi\nu)} \int_0^{4\pi} e^{-\nu\xi} \sin\left(\frac{p\xi}{2}\right) J_p\left(2U_1 \sin\frac{\xi}{2}\right) d\xi,$$

$$B_p = \frac{2\pi\nu\theta_0}{1 - \exp(-4\pi\nu)} \int_0^{4\pi} e^{-\nu\xi} \cos\left(\frac{p\xi}{2}\right) J_p\left(2U_1 \sin\frac{\xi}{2}\right) d\xi. \tag{38}$$

In what follows, we consider two limiting cases: weak ($4\pi\nu \ll 1$) and strong ($4\pi\nu \gg 1$) collisions.

Infrequent collisions. Neglecting the small change in the exponential factor in the interval $(0,\ 4\pi)$ in the integrals (38), we obtain for A_{2k+1} and B_{2k+1}

$$A_{2k+1} = 2\theta_0 J_0(U_1) J_{2k+1}(U_1),$$

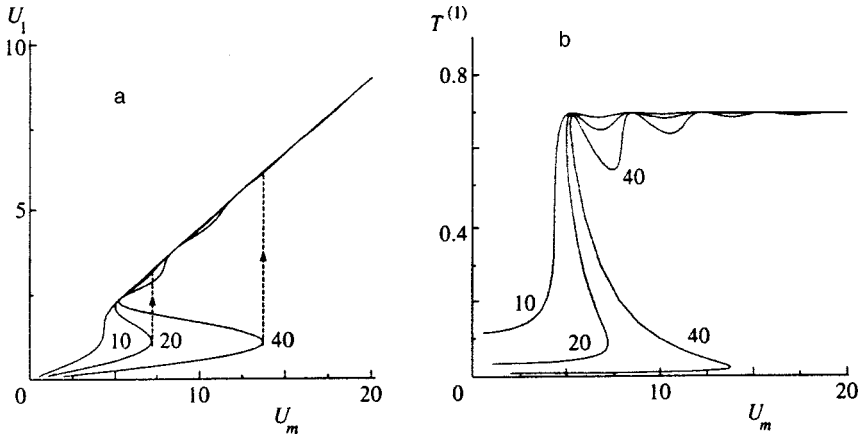

FIG. 2. a) Field amplitude U_1 in the superlattice at the first harmonic versus the amplitude U_m of the monochromatic wave incident on the structure. The curve was calculated using Eq. (41) for parameter values $\theta_0=10, 20, 40$. b) Transmission coefficient $T^{(1)}$ for radiation at the fundamental frequency versus the amplitude U_m of the incident wave. The curve was obtained from Eq. (41) and (46) for the same values of θ_0.

$$B_{2k+1}=0. \tag{39}$$

Substituting expressions (35), (36), and (39) into Eq. (33) gives

$$U_i(\tau)=U_1\cos\tau+\sum_{k=1}^{\infty}a_{2k+1}\sin(2k+1)\tau+\sum_{k=1}^{\infty}b_{2k+1}\cos(2k+1)\tau+2\theta_0J_0(U_1)\sum_{k=0}^{\infty}J_{2k+1}(U_1)\sin(2k+1)\tau. \tag{40}$$

We call attention to the factor $\sim J_0(U_1)$ in the nonlinear current. The amplitudes $U_1^{(s)}$ corresponding to the zeros of the Bessel function $J_0(U_1^{(s)})=0$ correspond to the so-called dynamical localization of electrons, which can be interpreted as the collapse of the energy band in a high-frequency field. It can also be interpreted as a specific effect in which the phases of individual electrons are mixed, as a result of which the macroscopic current vanishes. Of course, strict localization disappears if higher harmonics are taken into account in the integrand in Eq. (33) in addition to the fundamental. Higher harmonics play an ever greater role as the parameter θ_0 increases. Unfortunately, no precise criterion for the applicability of the present representation can be given, but the typical value of θ_0 for which the corresponding approximation first breaks down can be obtained from the rigorous numerical calculations presented below.

The expression (40) makes it possible to write out the relation between the amplitude of the incident wave and the amplitudes of the harmonics in the superlattice:

$$U_m=\sqrt{U_1^2+4\theta_0^2J_0^2(U_1)J_1^2(U_1)}, \tag{41}$$

$$\varphi_i=\cos^{-1}\frac{U_1}{\sqrt{U_1^2+4\theta_0^2J_0^2(U_1)J_1^2(U_1)}}, \tag{42}$$

$$a_{2k+1}=-2\theta_0J_0(U_1)J_{2k+1}(U_1), \tag{43}$$

$$b_{2k+1}=0, \tag{44}$$

where φ_i is the phase shift between the incident field and the field at the fundamental frequency inside the superlattice. Equation (41) yields the amplitude of the field at the fundamental frequency in a superlattice as an inverse function of the amplitude of the incident wave, while Eq. (43) (together with Eq. (41)) yields the corresponding parametric dependence for the amplitudes of the harmonics.

The curves constructed using Eq. (41) for various values of the parameter θ_0 are displayed in Fig. 2a. As one can see, there exists a critical value $\theta_0^{(C)}\sim 10$ beyond which the function $U_1(U_m)$ is no longer single-valued. This corresponds to the emergence of bi- or multistability in the system. The

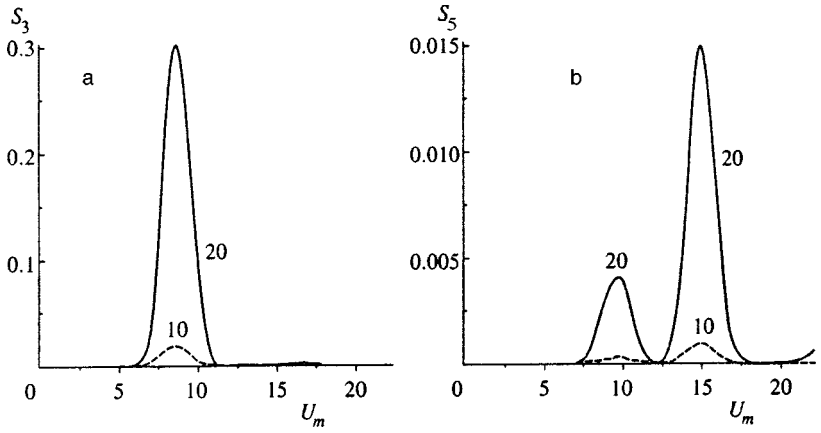

FIG. 3. Transformation coefficients S_{2k+1} (47) versus the amplitude U_m of the incident wave: a) into the third harmonic, b) into the fifth harmonic, for $\theta_0=10$ and $\theta_0=20$.

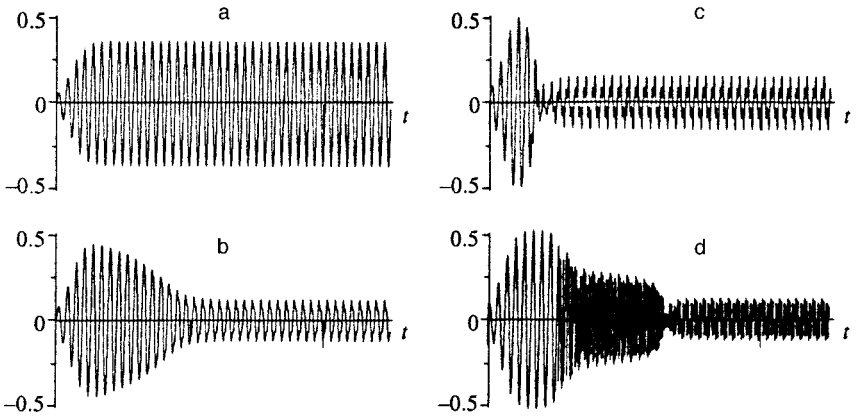

FIG. 4. Oscillograms of the electron velocity for various values of the amplitude of the incident wave for $\theta_0=80, \nu_1=0.2, \nu_2=0.1$: a) $U_m<U_m^*\simeq 15$ (quasilinear regime); b) $U_m\geqslant U_m^*\simeq 15$ (transition to the bleaching regime near threshold); c) $U_m\geqslant 18$ (transition to the bleaching regime); d) $U_m\geqslant 20$ (two-step transition to the self-induced transparency regime through an intermediate metastable state).

points at which a transition occurs from one quasistationary periodic regime to another (the regimes are marked in Fig. 2a by arrows) and the number of such points can be found by solving the transcendental equation

$$1+\frac{2\theta_0^2}{U_1}\frac{\partial}{\partial U_1}\{J_0^2(U_1)J_1^2(U_1)\}=0, \tag{45}$$

where the relation between U_1 and U_m is given by Eq. (41).

Equations (41)–(44) make it possible to write out expressions for the transmission coefficient (T) for radiation at the fundamental frequency and the transformation coefficients (S_{2k+1}) into higher harmonics:

$$T=\frac{4\sqrt{\varepsilon_s}}{(1+\sqrt{\varepsilon_s})^2}\frac{U_1^2}{U_1^2+4\theta_0^2J_0^2(U_1)J_1^2(U_1)}, \tag{46}$$

$$S_{2k+1}=\frac{16\theta_0^2J_0^2(U_1)J_{2k+1}^2(U_1)}{U_1^2+4\theta_0^2J_0^2(U_1)J_1^2(U_1)}, \tag{47}$$

where the amplitude of the field at the fundamental frequency in the superlattice is related to the amplitude of the in the incident field by (41). The function (46) is displayed in Fig. 2b and the function (47) is displayed in Fig. 3 for the third and fifth harmonics at various values of θ_0. The existence of hysteresis in the function $U_1(U_m)$ for $\theta_0>\theta_0^{(C)}$ results in both $T(U_m)$ and $S_{2k+1}(U_m)$ begin multivalued.

Thus, in the case of sufficiently infrequent collisions, which do not destroy the coherent motion of current carriers over the course of many periods of the high-frequency field, the system demonstrates complicated behavior that depends strongly on the electron density. The interaction of intense electromagnetic radiation with a lateral superlattice is accompanied by self-action effects, which lead to multistable self-induced transparency and to generation of odd-numbered harmonics, whose intensity is a nonmonotonic function of the amplitude of the incident wave.

Frequent collisions. In the other limiting case, $4\pi\nu\gg 1$, the integrals (38), which give the Fourier coefficients in the expansion of the current in the superlattice, can likewise be approximated by expanding the trigonometric functions and Bessel functions in the integrands in power series and retaining the leading terms. The result is

$$A_{2k+1}=\pi\theta_0\frac{(k+1)(2k+1)U_1^{2k+1}}{2^{2k}\nu^{2(k+1)}}, \tag{48}$$

$$B_{2k+1}=\pi\theta_0\frac{U_1^{2k+1}}{2^{2k}\nu^{2k+1}}. \tag{49}$$

The harmonics of the nonlinear current described by Eqs. (48) and (49) are monotonic functions of U_1, suggesting a lack of hysteresis. In principle, a more rigorous equation describing the dynamics of the superlattice can be obtained in the limit of very frequent collisions ($\nu\gg 1$). For this, a quasistationary current–voltage characteristic, which is obtained by neglecting the time derivatives in Eqs. (20), must be subsituted for the integral operator on the right-hand side of the relation (33). This yields a time-local algebraic relation between the incident field and the field inside the superlattice:

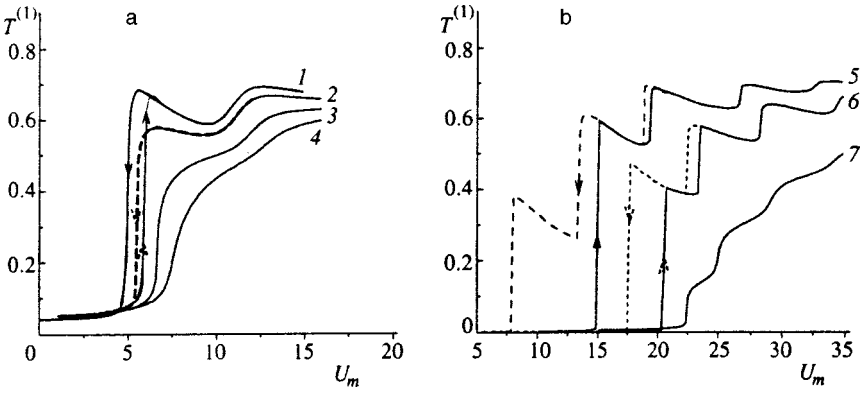

FIG. 5. Numerically computed curves of the transmission coefficients for a wave at the fundamental frequency as a function of incident field amplitude for a) $\theta_0=20$ (*1*—$\nu_1=\nu_2=0.01$; *2*—$\nu_1=\nu_2=0.1$; *3*—$\nu_1=\nu_2=0.3$; *4*—$\nu_1=\nu_2=0.5$) and b) $\theta_0=80$ (*5*—$\nu_1=0.2$, $\nu_2=0.1$; *6*—$\nu_1=\nu_2=0.3$; *7*—$\nu_1=\nu_2=1$).

$$U_i(\tau)=U(\tau)+\frac{\nu\theta_0 U(\tau)}{\nu^2+U^2(\tau)}. \quad (50)$$

This equation also describes self-induced transparency due to the presence of a descending section in the static current–voltage characteristic.

5. NUMERICAL RESULTS

Since the analytic results presented above were obtained for individual limiting cases, they cannot cover the entire picture of the interaction of the incident radiation with a lateral quantum superlattice. It is probably impossible to concoct a more complete picture of the nature of this interaction without recourse to numerical simulation, which can also be used to check the accuracy of the analytic model. The numerical calculations presented below are based on the solution of (20) with vanishing initial conditions.

Figure 4 displays oscillograms of the electron velocity, illustrating the transient processes in the system under study for various values of the parameter θ_0 and amplitudes U_m of the incident high-frequency field. The calculation was performed for a regime in which the external field is turned on gradually and the velocity w at the linear stage ''tracks'' the variation of the field in time with essentially no delay. Figure 4a demonstrates the subthreshold and, as we can see, almost linear regime of electromagnetic field interaction with the superlattice. Above the threshold U_m^* for self-induced transparency (Fig. 4b–4d) the dynamics of the system and the nature of the transient processes depend strongly on the amplitude of the incident field. For a field that exceeds the threshold value by a comparatively small amount, the amplitude of the electron velocity (and therefore the screening current) decreases comparatively rapidly (over several periods) to a certain stationary value and the motion of the current carriers becomes anharmonic (Fig. 4b). As the amplitude of the incident field increases, the duration of the corresponding transient process decreases and the anharmonicity of the motion in the established regime becomes stronger (Fig. 4c). When multistability is present, repeated switching in a strong incident field is possible in the system (this corresponds to a large θ_0). Figure 4d shows a two-step transient process in which the stationary state is reached through an intermediate quasistationary state.

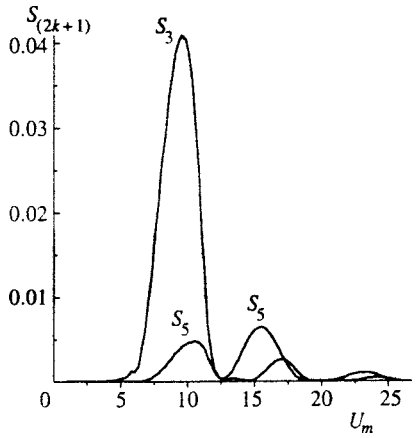

FIG. 6. Numerically computed transformation coefficients into the third and fifth harmonics versus the incident field amplitude for $\theta_0=20$, $\nu_1=\nu_2=0.1$.

Figure 5 shows the transmission coefficients for a wave at the fundamental frequency as a function of the amplitude of the incident electric field. The corresponding curves are shown for superlattices with various free-carrier densities and relaxation frequencies. We call attention to Fig. 5a, where the parameters (for curve *1*) correspond to the analytic calculations (Fig. 2b). As one can see, Eq. (46) predicts well the behavior of the system right up to $\theta_0\sim 20$. These calculations show that a characteristic feature of the self-reduced transparency effect is the multistep (in the amplitude of the incident wave) character of the bleaching. This is due to the multistability that emerges in the system at sufficiently high free-electron densities and low relaxation rates. In very strong fields (much greater than the bleaching threshold), the screening action of the superlattice becomes negligible and

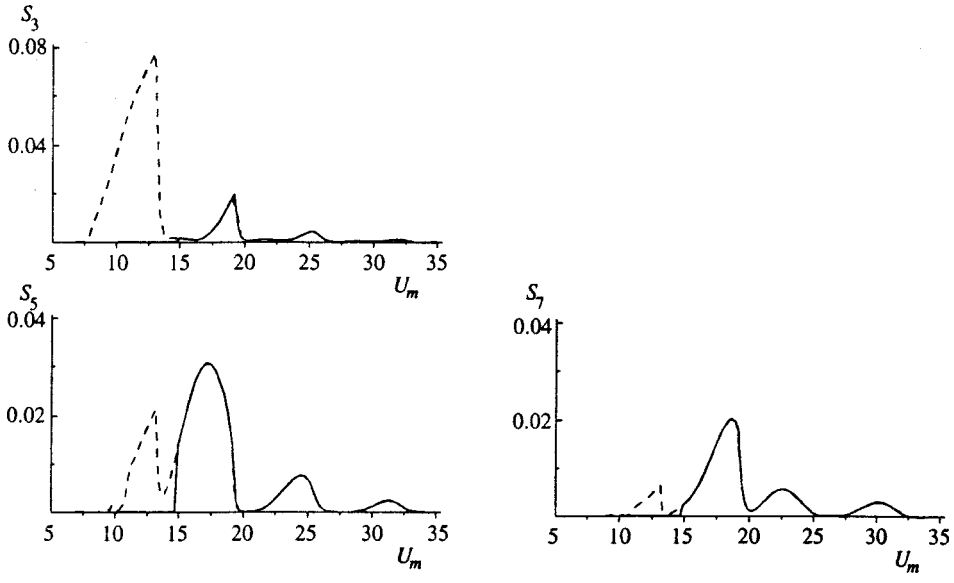

FIG. 7. Numerically computed transformation coefficients into the third, fifth, and seventh harmonics versus the incident field amplitude for $\theta_0=80$, $\nu_1=\nu_2=0.1$.

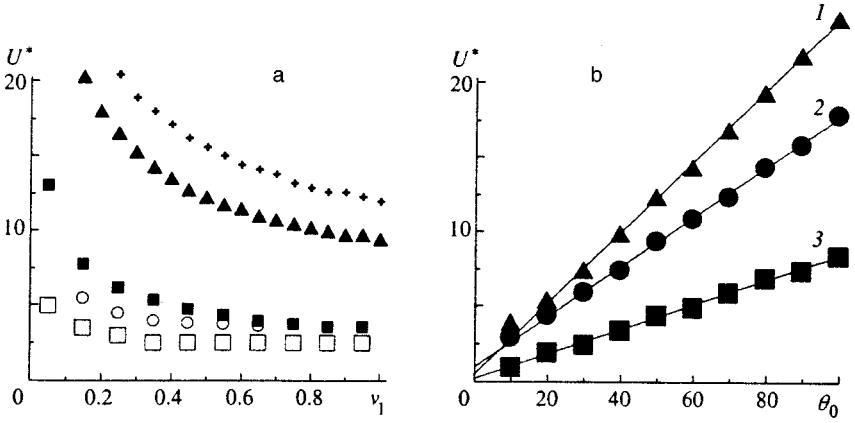

FIG. 8. Bleaching threshold U^* versus a) relaxation rates [$\theta_0=80$: $\nu_2=0.25$ (+); $\nu_2=0.15$ (▲); $\nu_2=0.01$ (■), and $\theta_0=20$: $\nu_2=0.1$ (○); $\nu_2=0.01$ (□)]; b) free-carrier density [$\nu_1=\nu_2=0.1$ *(1)*; $\nu_1=0.2$, $\nu_2=0.1$ *(2)*; $\nu_1=0.1$, $\nu_2=0.01$ *(3)*].

the transmission coefficient asymptotically approaches a value corresponding to the transmission coefficient for radiation propagating into the substrate with no film on its surface. We note that the reflection coefficient as a function of the amplitude of the incident radiation, which can be found using Eqs. (15) and (41), has all of the same features (multistep nature, hysteresis behavior) as the corresponding transmission coefficients.

Efficient generation of high (odd-numbered) harmonics is observed between the bleaching thresholds of the superlattice. We calculated the transformation coefficients into the $(2k+1)$-st harmonic as a function of the amplitude on the incident wave. Figures 6 and 7 show the results for S_{2k+1} for $2k+1=3$, 5, and 7 for θ_0 and $\nu_{1,2}$ corresponding to Fig. 5. We note a characteristic feature of the generation of high harmonics in the hysteresis region $U_m<U_m^*$ (U_m^* is the bleaching threshold), where as the amplitude of the external field decreases, the intensity of the harmonics is redistributed so that the third harmonic becomes strongest and the maximum intensity for $U_m<U_m^*$ is greater than in the region $U_m>U_m^*$. For example, for $\theta_0=80, \nu_1=0.2$, and $\nu_2=0.1$ the intensities of the fifth and seventh harmonics decrease, while the intensity of the third harmonic for $U_m\approx 10<U_m^*$ more than doubles compared with the above-threshold value (the dashed curve in Fig. 7).

In conclusion, we present in Fig. 8 the bleaching threshold as a function of the collision frequency (a) and density (b). Note that while the bleaching threshold tends to decrease as the frequency of elastic collisions increases and/or the density of free current carriers decreases, the relative change in the transmission coefficient of the structure decreases (see curves *3* and *4* in Figs. 5a and curve *7* in Fig. 5b).

6. CONCLUSIONS

This paper contains results of an investigation of self-action and harmonic generation accompanying the interaction of electromagnetic waves with lateral quantum semiconductor superlattices. The calculations were performed in a self-consistent formulation based on material equations obtained in the semiclassical approximation. Multistable self-induced transparency accompanied by efficient generation of odd-numbered harmonics was found. This effect is due to the influence of the nonlinear screening current flowing through the superlattice. We found the threshold relations determining the range of incident field amplitude and free-carrier density where the corresponding effects materialize.

This work was supported by the Russian Fund for Fundamental Research (Grant No. 97-02-17572).

[1] A. A. Ignatov, J. Genzer, E. Schomburg *et al.*, Ann. Phys. **5**, 173 (1995).
[2] M. Hadjazi, J. F. Palmier, A. Sibille *et al.*, Electron. Lett. **29**, 648 (1993).
[3] M. L. Wanke, S. J. Allen, K. Maranowski *et al.*, in *Physics of Semiconductors*, M. Scheffter and R. Zimmerman (Eds.), World Scientific, Singapore (1996), p. 1791.
[4] H. Schneider, K. Fujiwara, H. T. Grahn *et al.*, Appl. Phys. Lett. **56**, 605 (1990).
[5] A. Ya. Shik, Fiz. Tekh. Poluprovodn. **8**, 1841 (1974) [Sov. Phys. Semicond. **8**, 1195 (1974)].
[6] L. Esaki and L. L. Chang, Thin Solid Films **36**, 285 (1976).
[7] L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 61 (1970).
[8] A. A. Ignatov and Yu. A. Romanov, Phys. Status Solidi B **73**, 327 (1976).
[9] A. A. Ignatov, J. Genzer, K. F. Renk *et al.*, Z. Phys. B **98**, 187 (1995).
[10] H. Noguchi, J. P. Leburton, and H. Sakaki, Inst. Phys. Conf. Ser. No. 129 (1993), Chap. 5, p. 299.
[11] A. A. Ignatov, E. P. Dodin, and V. I. Shashkin, Mod. Phys. Lett. **5**, 1087 (1991).
[12] A. Wacker and Antti-Pekka Jauho, Phys. Rev. Lett. **80**, 369 (1998).

Translated by M. E. Alferieff