Coulomb dissociation of incident ^{16}O at 4.5 GeV/c

M. Nabil Yasin and M. M. Sherif

Physics Department Faculty of Science, Cairo University, Giza Egypt

S. M. Abd El-Halim

Phys. Department Faculty of Science, Bonha-Zagazig University, Egypt

M. A. Jilany

Sohag Faculty of Science, South Valley University, Sohag, Egypt

(Submitted 10 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 1153–1162 (April 1997)

Projectile multifragment breakup of ^{16}O, ^{2}C and ^{7}Li at energies 3.0–4.5 A-GeV is studied by means of the Weizsäcker–Williams method. The fragmentation channels of the ^{16}O projectile at 4.5 A-GeV are investigated and compared with that of ^{16}O at 200 A-GeV. The events characterized by $N_h=0$ and the events due to both Coulomb and diffraction dissociation have been selected and analyzed as a function of impact parameter. Also, the dependence of the electromagnetic dissociation cross-section on incident energy and the charge of projectile and target is found. © *1997 American Institute of Physics.* [S1063-7761(97)00104-2]

1. INTRODUCTION

In recent years, nucleus–nucleus collisions at high energies have been attracting more interest as a way to understand the important effects of nuclear interaction mechanisms at different impact parameters b. The reason for the study at small impact parameter is to observe the signatures of unusual forms of nuclear matter such as the quark–gluon plasma,[1] while studies at higher values of the impact parameters help in understanding the nuclear structure, and also, electromagnetic effects[1,2] considered as a source of background for nuclear interactions like the possible features for quark gluon plasma.

Heckman and Greiner[3] reported the first results of a Bevatron experiment on the fragmentation of nitrogen ions by carbon and hydrogen targets at 29 A-GeV. The point of interest in this work is the dislocation of fragments from the projectile nucleus. In their work, they tried to classify the reactions according to the following categories.

1) Stripping reactions, in which the spectator part of a projectile is diffracted inelastically while the other part suffers a strong interaction with the target nucleus. This reaction occurs at $R_p+R_t>b>R_p-R_t$ where R_p and R_t are the radii of projectile and target nucleus respectively.

2) More peripheral collisions, in which all projectile constituents are dislocated and associated with some hadrons, mostly pions. These interactions occur at large values of the impact parameter b, up to R_p+R_t.

3) Pure electromagnetic dissociation of the incident charge nucleus as a result of its passage through the Coulomb field of the target nucleus, which provides an electromagnetic pulse of short duration, enhanced by the Lorentz contraction factor γ. These interactions occur at impact parameters b greater than the sum of projectile and target nucleus radius (i.e., $b>R_p+R_t$).

2. EXPERIMENT AND RESULTS

2.1. Experimental details

At the Dubna BR-2 synchrophasotron emulsion stacks were irradiated with beams of 4.5 A-GeV ^{16}O and ^{12}C ions and 3.0 A-GeV ^{7}Li ions. The dimensions of the pellicells where $20\ \text{cm}\times 10\ \text{cm}\times 600\ \mu\text{m}$. The composition of the emulsion is given in Table I.

Scanning was carried out along the track,[4] using MBI-9 and Wild binocular microscopes, with magnification 10 $\times 100$.

Totals of 958, 1000, and 968 events were found, giving the interaction mean free paths 12.18 ± 0.33, 14.4 ± 0.33 and 15.2 ± 0.50 cm for ^{16}O, ^{12}C, and ^{7}Li, respectively.

The charged particles (tracks) produced in each interaction are grouped in the following categories.[4]

a) Ionization shower tracks (N_s) with very high velocity $\beta=v/c\geqslant 0.7$ and relative grain density $g^*=g/g_0\leqslant 1.4$, where g_0 is the minimum grain density of the relativistic tracks inside the emulsion. Most of them are mesons, along with some fast hydrogen isotopes.

b) Grey tracks (N_g), which are tracks with range $R\geqslant 3$ mm and $6\geqslant g^*>1.4$. These are mainly knock-out protons from the target nucleus.

c) Black tracks (N_b), which are slow fragments from the target nucleus with range $R<3$ mm and $g^*>6$.

In emulsion experiments, the term heavy tracks is used, and their number is defined as $N_h=N_g+N_b$.

d) The projectile fragments (PFs) are strongly collimated in the forward direction within an angle determined by the approximate relation

$$\sin\theta_{\text{PF}}\leqslant\frac{P_f}{P_i},$$

where P_f is the Fermi momentum[5] and P_i is the incident

 1063-7761/97/040635-06$10.00

TABLE I.

Types of nucleus	$^{1}H_{1}$	$^{12}C_{6}$	$^{14}N_{7}$	$^{16}O_{8}$	$^{80}Br_{35}$	$^{108}Ag_{47}$
Density (N of atoms/10^{22})	3.15	1.412	0.395	0.956	1.028	1.028

momentum. This gives $\theta_{PF} \leqslant 3.0°$ at the Dubna energy. In Heckman and other[4–6] the projection angles of the PFs are measured by finding

$$\tan\ \theta_{PF} = \frac{\Delta y}{\Delta x},$$

where Δy is the deflection of the measured fragment at a longitudinal distance $\Delta x = 1$ cm from the emission point.

The PFs with charge $Z \geqslant 2$ for 708, 1000 and 970 events of ^{16}O, ^{12}C, and ^{7}Li beams respectively were recorded. It is possible to identify the doubly charged fragments ($Z = 2$) by eye. A correct estimation is made by using the δ-ray method.[4]

2.2. Selection of electromagnetic dissociation events

To distinguish between the nuclear interactions and the Coulomb dissociation events, we must consider the expected characteristics of electromagnetic dissociation (ED) events as described in Ref. 7. Generally these ED events are characterized by $N_h = 0$, which means that no slow emitted particles and no lepton pair or β-decay are produced. Also, the incident charge (Z_p) must equal the sum of the fragment charges and the emitted angle θ_{fr} should be $\leqslant 3.0°$ at projectile incident energy 4.5 A-GeV.

These stringent selection criteria are applied to the experimental data of ^{16}O interactions at 4.5 A-GeV and compared with that of ^{16}O interactions at 200 A-GeV.[7,8] It was found that about 6% of the total inelastic interactions are due to the effect of the Coulomb field of the target nucleus on the incident ^{16}O at 4.5 A-GeV, whereas this ratio is about 10% for the same projectile at 200 A-GeV. These results confirm the dependence of the ED cross-section on the incident energy. Table II represents the total number of analyzed events for each beam. The ED events can be divided into two categories, one due to the Coulomb field effect and the other to diffraction on the target surface. This has been corrected for the number of ED events which are not associated with pions, as shown in Table III. Then the angle of the emission fragment ($\theta_{PF} \leqslant 3.0°$) can be divided into two ranges. The lower range of θ_{PF} (0–1.5°) is attributed to the effect of the Coulomb field, and the relatively larger angles up to 3.0° result from the effect of diffraction on the outer surface of the target nucleus.

TABLE III. Fragmentation modes of ^{16}O projectile at 4.5 A- and 200 A-GeV/c.

Fragmentation Mode	Fraction of Frag. Chan. of ^{16}O (4.5) A-GeV %	Fraction of Frag. Chan. of ^{16}O (200) A-GeV %
O→N+H	42.2±6.2	56.0±4.0
→C+2H	1.8±1.3	14.0±2.0
→C+He	5.5±2.2	10.0±2.0
→B+He+H	1.8±1.3	3.6±1.0
→B+3H	2.8±1.6	0.8±0.5
→Be	4.6±2.1	5.5±1.2
→Li	2.8±1.6	2.5±0.8
→4He	1.8±1.3	0.8±0.5
→3He+2H	5.5±2.2	4.4±1.1
→2He+4H	1.8±1.3	3.3±1.0
→He+6H	0.9±0.9	0.8±0.5
Reference	Present Work	Refs. 7 and 8

TABLE II. The $N_h = 0$ events and those due to both Coulomb and diffraction dissociation in ^{16}O-Em. at 4.5 A-GeV/c and ^{16}O-Em at 200 A-GeV. Their separation is deduced from the experimental data.

The projectile and energy	^{16}O (4.5 A-GeV/c)		^{16}O (200 A-GeV)	
	No. of events	%	No. of events	%
Total analysed events	708	-	920	-
$N_h = 0$ events (peripheral coll. + Coul. & diff. diss. + Simulated ED)	98	13.98±1.4	112	12.00±1.13
ED events (Coulomb diss.+ diff. diss.)	-	-	92	10.00±1.00
Pure ED event (Coulomb diss.)	45	6.35±0.94	-	-
Diff. diss. on target (associated with pions)	36	5.08±0.85	-	-
The dissociation into α-fragments only	2 O→4α	0.28±0.19	3 O→4α	0.30±0.06
Diffractive diss. into α-fragments	2 O+Em.→4α+low energy particle	0.28±0.19	2 in all scanned 2934 events	0.07±0.05
Reference	Present work		Refs. 7 and 8	

TABLE IV. The interaction of ^{6}Li, ^{7}Li, ^{12}C and ^{16}O at Dubna energies compared with ^{16}O at higher energies.

Type of projectile and energy (A-GeV)	^{6}Li (4.5)	^{7}Li (3.00)	^{12}C (4.5)
Total no. of analysed inel. int.	968	970	1000
$N_h=0$ events	147	136	98
Fraction %	(15.2±1.25)	(14.02±1.2)	(9.8±1.0)
ED events	70	45	60
Fraction %	(7.2±0.86)	(4.64±0.7)	(6.0±0.77)
Nuclear radius*	2.55 fm	2.42 fm	2.46 fm
Binding energy** BE (in MeV)	32.09	39.25	92
BE/N (MeV)	5.35	5.61	7.7
Lorentz factor (γ)	4.51	3.22	4.51
$E^{\max}$ (MeV)	99.4	72.0	100.5
Duration time τ_d through Ag target component (sec)	0.66×10^{-18}	0.93×10^{-18}	0.66×10^{-18}
References	Ref. 17	Present work	Present work

Type of projectile and energy (A-GeV)	^{16}O (4.5)	^{16}O (60)	^{16}O (200)
Total no. of analysed inel. int.	708	528	920
$N_h=0$ events	98	-	112
Fraction %	(13.98±1.4)	-	(12.00±1.13)
ED events	81	31	92
Fraction %	(11.43±1.27)	(5.87±1.05)	(10.00±1.00)
Nuclear radius*	2.73 fm	2.73 fm	2.73 fm
Binding energy** BE (in MeV)	128	128	128
BE/N (MeV)	8.0	8	8
Lorentz factor (γ)	4.51	64.41	214.71
$E^{\max}$ (MeV)	98.1	1392.3	4641.1
Duration time τ_d through Ag target component (sec)	0.68×10^{-18}	0.50×10^{-19}	0.14×10^{-21}
References	Present work	Refs. 16 and 22	Refs. 7 and 8

*Note.**Ref. 18, **Ref. 19.

The given numbers of ED events for ^{16}O at 200 A-GeV are taken from Refs. 7 and 8. It is clear that these ED events increase with the energy and charge of the incident beam, except for incident ^{7}Li. This exception may arise because the binding energy of the ^{7}Li nucleus is smaller than that of the other nuclei, as illustrated in Table IV. The fraction of ^{16}O dislocation into four clusters of α-particles at 4.5 A-GeV is more than twice that of the splitting of ^{16}O (200 A-GeV) into the same number 4α-particles. One reason for this may be the difference in the time the projectile spends inside the Coulomb field of the target nucleus as shown in the last table; this will be discussed in the next section.

Also, the previous conditions could be applied as selection criteria for $N_h=0$ events of the ^{7}Li interaction at 3.0 A-GeV in order to determine the two easily identified fragments of $Z=1$ and $Z=2$, since the incident projectile has a charge $Z=3$. These events are tabulated in Table IV and compared with other collected data. The difference in the ED event ratios for incident ^{6}Li and ^{7}Li at nearly the same total energy is due to the difference in the binding energy per nucleon. The ^{7}Li nucleus, which contains seven nucleons, is smaller and has higher binding energy per nucleon than the ^{6}Li nucleus. Therefore, ^{6}Li dissociates more readily than ^{7}Li in the case of interaction with the same target and at the same energy per nucleon. For this reason, the larger value of the ED events ratio for ^{6}Li than that for the ^{16}O nucleus at the same energy per nucleon could be anticipated. This is clear from the data given in Table IV.

2.3. Coulomb and diffraction dissociation events

Fragmentation of the projectile can be induced by the strong nuclear interaction (specially at larger impact parameters) or by the electromagnetic field interaction. The basic idea[9] of Coulomb break-up is similar to that for electrodisintegration, but the cross-section is enhanced by Z_t^2 (where Z_t is the target charge number). The virtual photon method of Williams and Weizsäcker[8–12] was used as a suitable way of considering the Coulomb effect in the fragmentation process. The Coulomb field of the target nucleus (in the projectile rest frame) appears as a packet of quasireal photons of short duration and enhanced by the Lorentz contraction factor $\gamma=(1-v^2/c^2)^{-1/2}$.

The collision time[1,13] is roughly

$$t_d=b_{\min}/\gamma c,$$

where $b_{\min}$ is the minimum impact parameter, equal to R_p+R_t. Consequently, the maximum photon frequency is contained in the electromagnetic field

$$\omega_{\max}=c\gamma/b_{\min}.$$

Also, this yields the maximum photon energy

$$E_\gamma^{\max}=hc\gamma/(R_p+R_t).$$

According to this model, the electromagnetic pulse can be sufficiently energetic to excite a giant resonance in the nucleus or to create lepton pairs or pions. From the basic assumptions of the model there is a flux of photons around the nucleus and the photon energy spectrum is computed classically[1,11,13–15] and treated by quantum mechanical calculations distinguished by the multipolarities of the photon spectrum. The values of γ and $E_\gamma^{\max}$ for an Ag target nucleus (the heaviest and most abundant element in the emulsion) for different projectiles and energies are estimated and listed in Table IV. It can be concluded that the values are 4.8, 4.8, and 3.2 for incident ^{16}O (4.5 A-GeV), ^{12}C (4.5 A-GeV) and ^{7}Li (3.0 A-GeV) respectively. The corresponding $E_\gamma^{\max}$ values are less than 100 MeV, which is smaller than the threshold energy for pionization, while for incident ^{16}O (60 A-GeV), it is equal to the value of producing pions having a mass about 140 MeV. But above the threshold for Δ-resonance (1232 MeV), for incident 200 A-GeV, it could produce Δ-resonances, which are subsequently dissociated into protons and pions. The present results for ^{12}C and ^{16}O strongly suggest that no pions are associated with the pure ED events. It is clear from Table III that pions accompanied the separated events due to diffraction dissociation, confirming the

TABLE V. The nuclear mean free paths and inelastic cross-sections compared with the corresponding ED mean free path and σ_{ED} for different projectiles at various energies.

Type of projectile and energy	λ_{int}, cm	λ_{ED}, cm	σ_{int}, mb	σ_{ED}, mb	Reference
$^{7}Li_{3}$ 3.0 *A*-GeV	15.2 ±0.50	351.3 ±49.6	1740.5± 20.0	50.50± 0.6	Present work
$^{12}C_{6}$ 4.5 *A*-GeV	14.4 ±0.33	236.0 ±28.2	1837.2± 20.0	75.11± 0.5	Present work
$^{16}O_{8}$ 4.5 *A*-GeV	12.18±0.33	106.45±10.68	2070.0±140.0	91.2 ± 0.9	Present work
$^{16}O_{8}$ 200 *A*-GeV	12.0 ±0.20	96.0 ± 5.0	2620.0± 50.0	198.96±11.3	Ref. 7

present technique. We conclude that the technique of Ref. 23 can be used only at incident energy smaller than 10 *A*-GeV for light incident nuclei.

2.4. Mean free paths and corresponding cross-sections

Table IV represents the total observed inelastic, peripheral, and electromagnetic interactions. We can say that ED events as a percentage of the total number of events increase with the projectile energy. At the same time, the measured value of λ_{ED} decreases as the incident projectile charge and energy increase. The total inelastic cross-sections σ_{in} are estimated from the relation

$$\sigma_{in}=1/n_{eff}\lambda_{in},$$

where n_{eff} in our type of emulsion is the effective density of emulsion nuclei;

$$n_{eff}=\sum_i n_i\sigma_i \Big/ \sum \sigma_i=3.78\cdot 10^{22} \text{ atoms/cm}^3.$$

However, this number takes the value $3.2\cdot 10^{22}$ atoms/cm^3 for the prepared emulsion in CERN SPS at 200 *A*-GeV[7,14] where λ_{in} is taken as the observed mean free path for each beam. In order to obtain an absolute value for the electromagnetic dissociation cross section σ_{ED}, the measured value of λ_{in} in the emulsion must be converted into an absolute cross-section on the Ag component, which is the heaviest and most abundant element in the emulsion. The estimated σ_{ED} for ED events on the Ag target is then calculated from the relation $\sigma_{ED}=f/\rho\lambda$, where $\rho=1.028\cdot 10^{22}$ atoms/cm^3 is the density of nuclei in the emulsion and $f=0.67$ is a weight factor for this target component reaching unity for the emulsion as a whole.

The values of σ_{ED} as a function of the projectile charge are shown in Fig. 1b for this work and other experiments.[15] This shows that σ_{ED} takes similar values despite the different detectors and reactions. It can be concluded that in spite of the heterogeneity of the emulsion target, it can be used generally to measure ED cross-section.

Figure 1a and 1b illustrates some experimental values of σ_{ED} as a function of projectile charge. The theoretical values of σ_{ED} calculated by the Weizsäcker–Williams (WW) method[11] for relativistic heavy-ion reaction with a ^{197}Au target at Bevalac energies[15] are also included. Here the calculated σ_{ED} values approach large values for both ultrarelativistic energy (100 *A*-GeV) and high projectile charge, as noted in Ref. 15. Also, the experimental and calculated val-

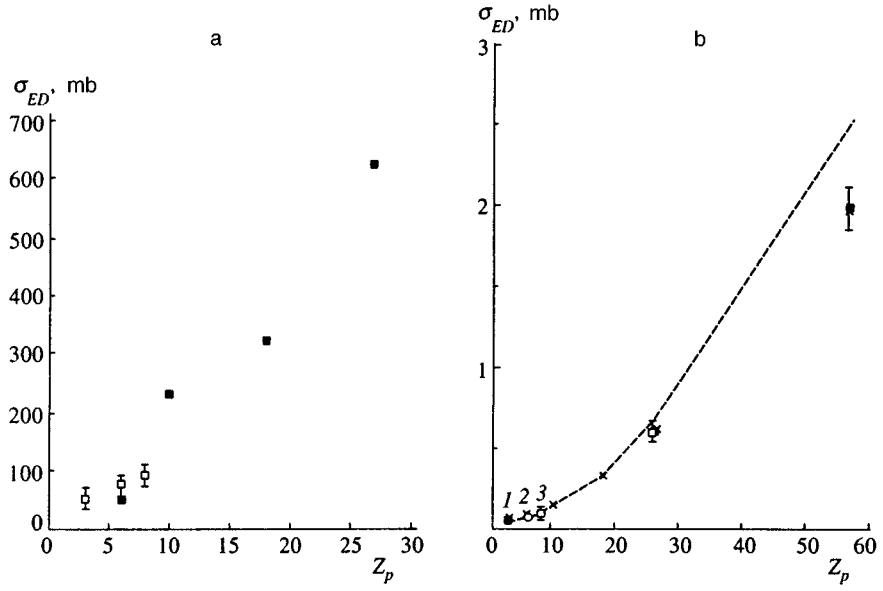

FIG. 1. a—The ED cross-section for the ^{7}Li, ^{12}C and ^{16}O interactions at Dubna energy (3.0–4.5 *A*-GeV) as a function of projectile charge Z_p (■—calculated, □—experimental) b—Our experimental points ■—*1*, ○—*2*, □—*3* for ^{7}Li ^{12}C and ^{16}O respectively at nearly the same energy. The points *x* represent the experimental values of σ_{ED}'s at different projectiles with ^{197}Au target. The corresponding calculated values according to WW-method are indicated by the dashed line.

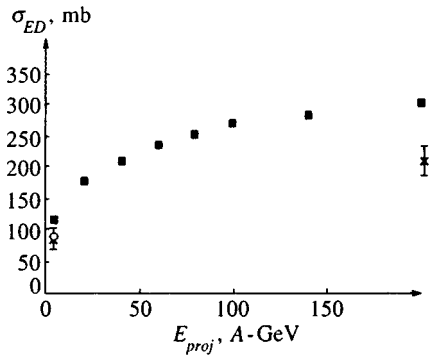

FIG. 2. The calculated ED cross-sections at different energies (■). The points ○ and *x* represent the experimental values of σ_{ED}'s for oxygen and carbon beams at 4.5 *A*-GeV.

ues of the σ_{ED} cross-section for ^{7}Li, ^{12}C and ^{16}O interactions as a function of projectile charge Z_p are shown in the same figure.

From this, it can be inferred that there is fair agreement between the present results for light projectile charge and calculated values at various energies.

The interactions of lithium, carbon and oxygen of the present work at Dubna energies are tabulated in Table IV. For comparison, the corresponding results of oxygen interactions at higher energies extracted from Refs. 7, 16 and 17 are also included. As seen from Table IV and Figs. 1 and 2, the percentage of ED events increases with both energy and incident beam charge.

Table III presents the different modes of carbon and oxygen breakup at 4.5 *A*-GeV inside the emulsion target (mainly due to the electromagnetic field of the Ag target nucleus component). Their numbers, fractions, and observed partial cross-sections are listed. The majority of events lie in the channels producing 1 He and 2 He respectively. This reveals qualitatively that the relative production rates of α-particles in the projectile fragmentation processes are consistent with the values for different projectiles and energies.[20] The largest yield of ED events occurred in the case of emission of one He fragment, associated with complete disintegration of the projectile. This result shows that with decreasing He multiplicity, the surface excitation energy of the projectile spectator increases. Also, the results confirm the same conclusions from the distributions of He fragments produced from nuclear events,[21] which implies the limiting fragmentation behavior of the He-multiplicity distribution.

2.5. Dependence of the electromagnetic dissociation cross-section on target nucleus

Vidovic and Greiner *et al.*[31] have studied the impact-parameter dependence of electromagnetic particle production in ultrarelativistic heavy-ion collisions and directly deduced the equivalent photon method[11,24] from quantum electrodynamics. It is well known that the simple photon method of Williams and Weizsäcker[11] is widely used[2,7,8,15,20–30] by scientists. A model[32] which depends on the WW-method[11] has been proposed for the dissociation of relativistic nuclei in the Coulomb field of heavier nuclei which are easily excited or dissociated by a target nucleus.

In Ref. 33 the decay channels of ^{28}Si projectile dissociation to $p+^{27}$Al, $n+^{27}$Si and $2p+^{26}$Mg using the WW-approximation with the experimental data for photonuclear cross sections were examined, and it was found that there is no evidence for two-photon excitation of the ^{28}Si nucleus. In contract, the authors of Ref. 34 found large values for the neutron cross section for electromagnetic dissociation of ^{197}Au targets by incident ^{20}Ne (1.7 *A*-GeV) and ^{86}K, ^{197}Au, ^{209}Bi beams with 1 *A*-GeV. They explained this observation by the dominant contribution from two-photon gain dipole excitation and interpreted the large cross section as due to the Z_t dependence, since two-photon excitation will behave as Z_t^4 while one-photon excitation shows Z_t^2 behavior (or $Z_t^{1.8}$ behavior for heavy target nucleus), as found by the authors of Ref. 33.

3. CONCLUSIONS

The results confirm the dependence of the electromagnetic dissociation cross section on both incident charge and energy. The measured ED cross section has the same trend as the calculated one based on the Weizsäcker–Williams method, and it increases with increasing of the projectile charge and energy.

For light projectiles, it was found that the largest yield of ED events occurred in the case of one He fragment associated with complete disintegration of the projectile. This shows that with decreasing He fragment multiplicity, the surface excitation energy of the projectile spectator increases.

Calculations by the WW method with data from counter experiments can be used to examine the decay channels for projectile dissociation by the excitation of the projectile nucleus, whereas other experiments interpret the ED of target nucleus in terms of the dominant contribution from two-photon giant dipole excitation.

Despite the heterogenity of the emulsion target, it can be used to measure the ED cross section.

The larger yield of ^{6}Li projectile dissociation than that for incident ^{7}Li may be due to the smaller binding energy per nucleon of the ^{6}Li nucleus.

[1] R. Bhanja *et al.*, Phys. Rev. Lett. **54**, 771 (1985).
[2] C. A. Bertulani and G. Baur, Phys. Rep. **163**, 299 (1988).
[3] H. H. Heckmen and D. E. Greiner, Phys. Rev. Lett. **28**, 926 (1972).
[4] W. H. Barkas, *Nuclear Research Emulsions*, Academic, N.Y. (1963), Vol. 1.
[5] E. J. Moniz *et al.*, Phys. Rev. Lett. **26**, 445 (1971); M. N. Yasin, Ph. D. Thesis, Cairo University (1981).
[6] H. H. Heckmen *et al.*, Phys. Rev. C **17**, 1735 (1978).
[7] G. Baroni *et al.*, Nucl. Phys. A **516**, 673 (1990).
[8] G. Baroni *et al.*, Nucl. Phys. A **531**, 961 (1991). Preprint CERN-PPE/92-08 (1992); Preprint CERN-PPE/91-03 (1992).
[9] X. Artru and G. B. Yodh, Phys. Lett. B **40**, 43 (1972).
[10] T. D. Jackson, *Classical Electrodynamics*, Wiley, N.Y. (1962), p. 524; 2nd ed. (1975), p. 719.
[11] C. F. Weizsäcker, Z. Phys. **88**, 612 (1934); E. J. Williams Phys. Rev. **45**, 729 (1934).
[12] R. Jackle and H. Pilkuhn, Nucl. Phys. A **247**, 521 (1975).
[13] G. Baur and C. Hoffman, Nucl. Phys. A **30**, 247 (1984); C. A. Bertulani

and G. Baur, Nucl. Phys. A **442**, 739 (1985); C. A. Bertulani and G. Baur, Nucl. Phys. A **458**, (1986).
[14] A. Goldberg, Nucl. Phys. A **440**, 636 (1984).
[15] J. C. Hill *et al.*, Phys. Rev. Lett. **60**, 999 (1988);Phys. Rev. C **38**, 1722 (1988);Phys. Rev. C **39**, 524 (1989).
[16] G. Sing and P. L. Jain, Intern. Rept. Aug. 29 (1991).
[17] M. El-Nadi *et al.*, Egypt, J. Phys. No. 1–2 (1993).
[18] C. Douglas and A. Giancoli, *Physics (Principles with Application)*, 3rd ed. Prentice-Hall Int. Ed. (1991). In this book, the data are taken from Brookhaven Nat. Lab. (1990).
[19] E. G. Fuller, Phys. Rep. **127**, 3 (1985); C. W. Dejager *et al.*, Atomic Data and Nucl. Data Tables **14**, 479 (1974).
[20] M. I. Adamovich, Phys. Rev. C **40**, 66 (1989), Preprint El-10838, JINR, Dubna (1977).
[21] D. L. Olsen *et al.*, Phys. Rev. C **24**, 1529 (1981); H. M. Heckmen *et al.*, Phys. Rev. C **37**, 56 (1976).
[22] G. Sing and P. L. Jain, Phys. Rev. C **41**, 999 (1990).
[23] M. Nabil Yasin, J. Fac. Educ. Ain Shams Univ., No. 19, 145 (1994); Preprint ICTP Rep. IC/194/389, Trieste, Italy.
[24] E. J. Williams Proc. R. Soc (London) A **139**, 163 (1933); E. Fermi Z. Phys. **29**, 315 (1924).
[25] G. Baur and L. G. Ferreira, Nucl. Phys. A **518**, 786 (1990).
[26] B. Muller and A. J. Schramm, Phys. Rev. D **42**, 3699 (1990).
[27] E. Papageorgiu, Phys. Lett. B **250**, 155 (1991); M. Dress, J. Ellis, and D. Zeppenfeld, Phys. Lett. B **223**, 454 (1989).
[28] R. N. Cahn and J. D. Jackson, Phys. Rev. D **42**, 3690 (1990).
[29] K. J. Abraham *et al.*, Phys. Lett. B **251**, 186 (1990); C. Hofmann *et al.*, Phys. Lett. B **262**, 210 (1991).
[30] J. Eicher, Phys. Rept. **193**, 165 (1990).
[31] M. Vidovic, M. Greiner, C. Best, and G. Soff, Phys. Rev. C **47**, 2308 (1993).
[32] H. Pilkuhm, Phys. Rev. b **38**, 143 (1972).
[33] J. Sonnadara, Upul Nucl. Phys. A **569**, 149 (1994).
[34] T. Aumann *et al.*, Nucl. Phys. A **569**, 157 (1994).

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

Measurement of the spin and temperature dependence of the $dd\mu$ molecule formation rate in solid and liquid deuterium

V. R. Bom and C. W. E. van Eijk

Delft University of Technology, NL-2629 JB, Delft, Netherlands

D. L. Demin, A. E. Drebushko, V. P. Dzhelepov, V. V. Filchenkov, N. N. Grafov, V. G. Grebinnik, A. D. Konin, D. V. Migachev, A. I. Rudenko, V. T. Sidorov, Yu. G. Zhestkov, and V. G. Zinov

Joint Institute for Nuclear Research, 141980 Dubna, Russia

J. D. Davies

University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K.
(Submitted 24 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1163–1173 (April 1997)

The $dd\mu$ molecule formation rates have been measured from the two hyperfine states of the $d\mu$ atom in the temperature range $T=5-30$ K. Results are consistent with the measurement of the TRIUMF group $T=3$ K and contradict theoretical predictions. This work was performed on the JINR phasotron (Dubna). © *1997 American Institute of Physics.* [S1063-7761(97)00204-7]

1. INTRODUCTION

Muon catalyzed fusion (MCF) in pure deuterium (see the schematic diagram in Fig. 1) has relatively simple kinetics and therefore is an attractive way to check the basic principles of the theory of muonic molecule resonance formation. Significant success has been recently achieved both in theory[1–3] and in the measurements[4–6] of the $dd\mu$ molecule formation rate ($\lambda_{dd\mu}$), including strong spin effects. As can be seen from Fig. 2, measurements of the temperature dependence $\lambda_{dd\mu}(T)$ at $T>20$ K are in excellent agreement with the standard theory of resonance muonic molecule formation. The most impressive consequence of their comparison is the determination of the energy of the weakly bound level in the $dd\mu$ system with an accuracy of $\simeq 0.1$ meV. Note that the latter corresponds to 1% of the relativistic contributions to this energy.

It was thought that measurements at lower temperatures would allow only improved accuracy of the main MCF parameters. However, the recent measurement of the $dd\mu$ molecule formation rate from the spin $F=3/2$ state of the $d\mu$ atom at $T=3$ K shows[7] a large discrepancy with theory. Our goal in this study was to extend the systematic measurement of $dd\mu$ mesomolecule formation rate from the different hyperfine $d\mu$ atom states $\lambda_{3/2}$ and $\lambda_{1/2}$ and the hyperfine transition rate λ_d in the temperature range 5–30 K. Preliminary results have been published.[8]

In this paper we give the results of full analysis, including the determination of the absolute values of the resonant (from the $d\mu$ atom spin state $F=3/2$) and nonresonant ($F=1/2$) $dd\mu$ molecule formation rates. At low deuterium temperature these values were previously measured in liquid deuterium at $T=22$ K (Ref. 6) and $T=23$ K (Ref. 4) and showed a noticeable discrepancy with theory for the value of $\lambda_{1/2}$ (nonresonant).[9] For normalization the authors of Ref. 7 used the value of $\lambda_{1/2}$ obtained in Ref. 4 and then corrected in Ref. 10; they also were not able to measure the deuterium density directly in the experiment.

An important feature of this study is that the measurements with liquid and solid deuterium were performed in the same experiment under conditions having well-defined deuterium density and temperature.

2. EXPERIMENTAL METHOD

The experimental method has been described in detail elsewhere.[6] We measured and analyzed the yield and time distribution of 2.5-MeV neutrons from the $d+d$ fusion reaction at seven temperatures in the range 5–30 K,

$$d\mu+d\rightarrow dd\mu\rightarrow {}^3\mathrm{He}+n. \tag{1}$$

A simplified diagram of the experimental apparatus is shown schematically in Fig. 3. In particular, a specially constructed solid deuterium target (T) of volume 280 cm^3 (Ref. 11) and a total absorption neutron spectrometer[12] (NE-213 provided $n-\gamma$ separation[13]). The spectrometer consisted of two identical parts symmetrically placed around the target with total volume of 22 liters. High neutron detection efficiency (solid angle $\simeq 65\%$ and intrinsic efficiency $\simeq 70\%$ resulted in sufficiently high counting rate yet relatively low random background.

The target was enclosed in a liquid-helium cooling cryostat. Special attention was given to achieving a high uniformity in temperature and density throughout the large target volume. Accordingly, a heat conductor consisting of 500 copper wires 0.4 mm in diameter was placed inside the target, which produced temperature gradients throughout the target not higher than 0.1 K. Temperatures were measured by two helium thermometers, which were placed at different heights inside the target whose temperatures were held constant within 0.2 K.

 1063-7761/97/040641-06$10.00

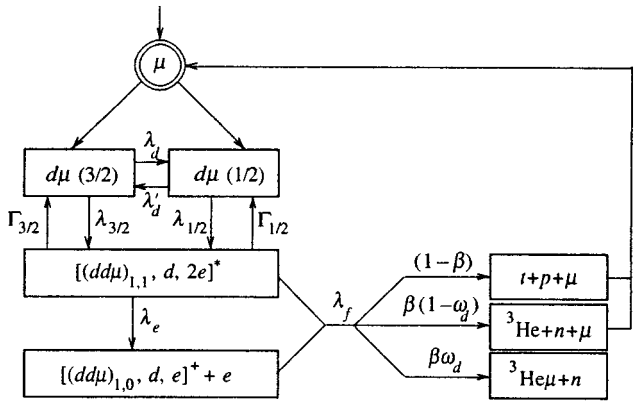

FIG. 1. Scheme of μ-catalyzed processes in pure deuterium.

The deuterium was purified with a palladium filter down to 10^{-1} ppm impurity concentration. The protium content was no higher than 0.5%.

The trigger selected those events for further recording and analysis which corresponded to the appearance of the neutron $[4(N_1+N_2)]$ and electron $[4(N_1+N_2)]$ signals during the 10-μs gate, opened by the muon stop $(1\cdot 2\cdot 3\cdot 4)$ signal. Discrimination against backgrounds originating from the muon stops in the target walls places the requirement on electron times of $t_e > 0.2$ μs after the muon stop ($t_0 = 0$).

As usual, the neutron yield was normalized to the number of electrons from the decay of muons stopped in deuterium. The time spectrum of μ-decay electrons obtained in the run at 19.0 K is shown in Fig. 4. It was analyzed using the expression

$$N(t) = a_1 \exp(-\lambda t) + a_2 v(t) + a_3. \qquad (2)$$

Here the short-lived component $a_2 v(t)$ (dashed curve in Fig. 4) represents muon stops in the target walls; its shape was determined from the measurements with an empty target. The slow exponent corresponds to muons stopped in deuterium with a slope very close to the free muon disappearance rate, ($\lambda_0 = 0.455$ μs^{-1}). The number of events belonging to the slow component (N_e) was used in analysis of the neutron events.

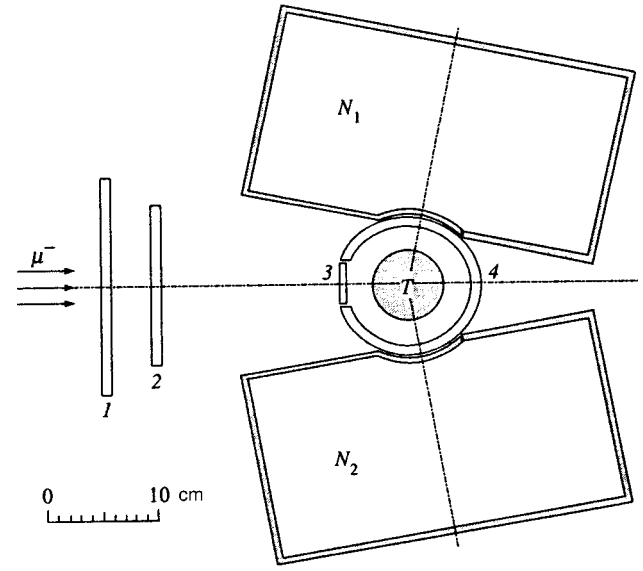

FIG. 3. Simplified scheme of the experimental setup. *1, 2, 3, 4* are scintillation counters; N_1,N_2 are total absorption neutron detectors, T is the deuterium target.

Experimental conditions for nine full runs are given in Table I. As usual, the density is normalized relative to that of liquid hydrogen (LHD, $\phi_0 = 4.25\times 10^{22}$ nuclei/cm^3). The numbers in brackets represent uncertainties in the last figure(s).

The highest statistics were accumulated at the lowest temperature and in runs 4 and 5, which were as close as possible in temperature but in different deuterium phase states. The measurement with helium was made to determine independently the neutron background [a_3 in Eq. (2) above], while the data obtained with an empty target allowed us to check the number of moun stops in the target walls.

Those events were selected for further analysis which satisfied the criteria of having:

1) A neutron in the $n-\gamma$ plot.[6,12]

2) A μ-decay electron in the time interval $t_n + 0.5\mu s \leqslant t_e \leqslant t_n + 2.5$ μs.

Times and amplitudes (recoil proton energy) for those events

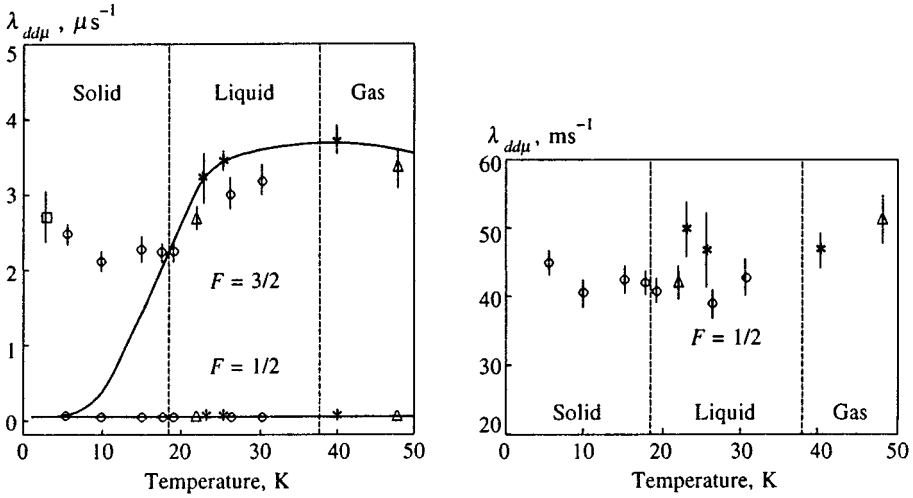

FIG. 2. Dependence $\lambda_{dd\mu}(T)$. Square—Ref. 7; circles—present work; triangles—previous Dubna measurements[6]; stars—Refs. 4 and 5. The line corresponds to the standard theory.[1–3]

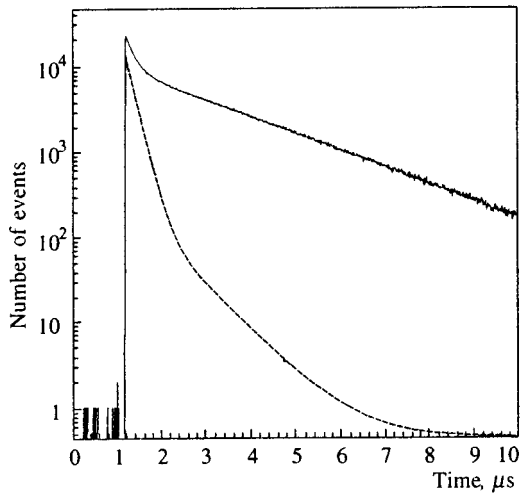

FIG. 4. Time spectrum of the decay electrons. The dashed line represents muon stops in the target walls.

were accumulated separately for each run and for each neutron detector. The final results were obtained from the analysis of such distributions for the first «detected» neutron events. However, the high neutron detection efficiency allowed us to record the «second detected» fusion neutron; their yield and time distributions were then used to verify the normalization procedure and the detection efficiency calculation. Some neutron time distributions are presented in Figs. 5a and 5b and in Fig. 6. We see that the relative yield of the background is low and that the neutron spectrum behavior remains the same when the temperature and the phase state of deuterium are changed.

3. KINETICS OF THE $d-d$ FUSION CYCLE

The $d+d$ muon-catalyzed processes in pure deuterium are shown in Fig. 1. According to standard theory,[1,2] the $d\mu$ atoms are formed with an initial kinetic energy of a few electronvolts in two hyperfine states from which they are quickly thermalized. The thermalization rate is estimated to be[14,15] $\lambda_{\text{therm}} \sim 10^9 \cdot \phi\, \text{s}^{-1}$, which is much higher than the μ molecule formation and the spin-flip rates. The thermalization stage is therefore ignored in the standard theory.

Muonic molecules can be formed either via the nonresonant Auger process, where the energy released under the $dd\mu$ formation is transferred to the conversion electron or via the Vesman resonance mechanism.[16] The rate of nonresonant $dd\mu$ molecule formation depends slightly on the $d\mu$ atom energy ($\epsilon_{d\mu}$) and is equal to $0.03-0.04\ \mu\text{s}^{-1}$ at $T \leqslant 50$ K (Ref. 9). According to the Vesman scheme, the resonant $dd\mu$ formation proceeds via the complex $[(dd\mu),d,2e]^*$, which is in an excited state. This process is characterized by a set of resonances whose positions are determined by the spin states of the $d\mu$ atom ($F=3/2,1/2$) and of the $dd\mu$ molecules ($S=3/2,1/2$), as well as by the rotation states of the initial D_2 molecule (K_i) and the complex (K_f). The transitions having

$$F=3/2 \rightarrow S=1/2, \quad K_i=0 \rightarrow K_f=1\,, \quad K_i=1 \rightarrow K_f=2 \tag{3}$$

dominate at the lowest temperatures.[17]

Normally, the $dd\mu$ formation is ignored during thermalization, but is considered for Maxwell's distributions of the $d\mu$ atom thermal energies. To obtain the values of $\lambda_{dd\mu}(T)$ for a given temperature, one therefore integrates the function[17] $\lambda_{dd\mu}(\epsilon_{d\mu})$ over the Maxwellian $W(\epsilon_{d\mu};T)$. This procedure was used in Ref. 18 to give the dependence of $\lambda_{3/2}(T)$ shown in Fig. 2. The Maxwell distributions for $T=5$ K and 20 K are shown in Fig. 7, together with the resonance closest in energy. As can be seen from this figure, the thermal energy distribution for $T=5$ K does not overlap this resonance. Only the nonresonant $dd\mu$ formation is therefore expected to contribute at this temperature. From Fig. 7 it follows also that at $T=20$ K $d\mu$ atoms spend a small part of their ''Maxwell cycle'' in the resonance region.

TABLE I. Parameters of the exposures.

Number	Target filling	Temperature, K	Density (LHD)	N_e, 10^6
1	Deuterium, solid	5.5(3)	1.43(4)	4.49(18)
2	Deuterium, solid	9.9(2)	1.43(4)	2.983(14)
3	Deuterium, solid	15.1(2)	1.42(4)	2.842(14)
4	Deuterium, solid	17.7(2)	1.40(4)	5.777(19)
5	Deuterium, liquid	19.0(2)	1.31(4)	4.598(17)
6	Deuterium, liquid	26.3(3)	1.19(4)	2.786(13)
7	Deuterium, liquid	30.5(3)	1.08(4)	1.738(12)
8	Helium	14.7(2)	0.37(1)	1.487(9)
9	Vacuum			

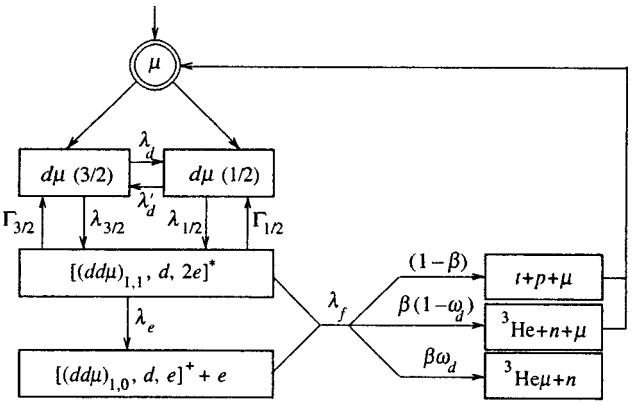

FIG. 5. Time spectra of the «first detected» catalysis neutron. The deuterium temperature: a—19.0 K (liquid); b—5.5 K (solid). Lines correspond to the function (4) with an optimal parameters found from the fit.

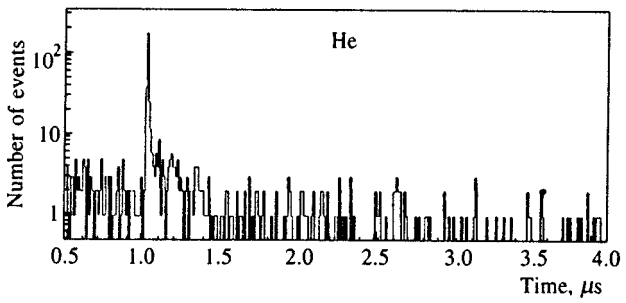

FIG. 6. Time spectrum of the background neutron events (target is filled with helium).

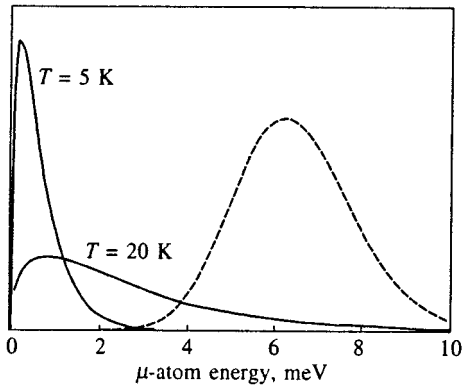

FIG. 7. Maxwell energy distribution of $d\mu$ atoms for $T=5$ K and $T=20$ K. The closest to zero resonance in the $dd\mu$ formation is shown by the dashed line.

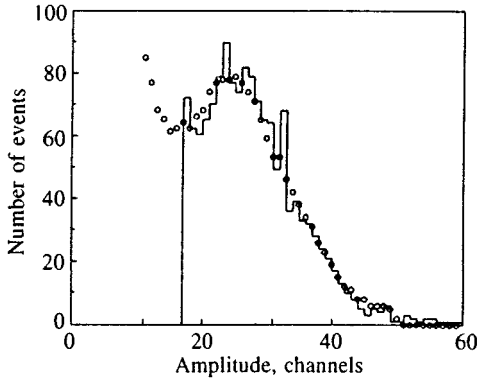

FIG. 8. Energy spectrum of protons recoiling from neutrons plotted for detector N1 with the target at 19.0 K (histogram). Circles are the corresponding calculations. The number of events is given in arbitrary units.

We conclude that the measured value of $\lambda_{dd\mu}$ should therefore be compared with a calculated effective value which includes:

1) a contribution during the thermalization stage, no higher than a few percent;[1,2]

2) integration over the Maxwell distribution. Note that the competition between the spin-flip and scattering processes was disregarded (see Ref. 18);

3) only $\simeq 1/4$ part of the $dd\mu$ molecules undergoes fusion in competition with the back decay of the complex.[1]

4. ANALYSIS

A set of the differential equations corresponds to the scheme of the $d-d$ fusion cycle shown in Fig. 1. When thermalization and the $d+d$ fusion rates are sufficiently high, it has an exact solution[1] for the form of the neutron time distribution:[1,19]

$$F_n(t)=b_f \exp(-\lambda_f t)+b_s \exp(-\lambda_s t). \tag{4}$$

For the fast exponent its slope λ_f is determined mainly by the spin-flip rate λ_d and its amplitude b_f is determined by the value of $\lambda_{3/2}$. The amplitude of the slow component b_s is close to the value of $\lambda_{1/2}$.

The parameters of the function (4) were found from the fit of the time distributions of the first detected neutrons. These spectra were convoluted with a Gaussian resolution function, to account for the finite time resolution (σ). The value of σ and the time zero (t_0) were optimized for each run. The analysis showed that the time zero stability during the data collection was better than 1 ns. The background due to muon stops in the target walls was approximated by an exponent with the slope $\lambda_f^b=5\ \mu s^{-1}$. Accidental neutron events were fitted as an exponent with $\lambda_s^b=\lambda_0$ for $t_n \geqslant t_0$ and with a constant value for $t_n \leqslant t_0$.

At the next stage of the analysis the absolute values of the steady state $dd\mu$-molecule formation rate were obtained from

$$\lambda_{ss}=\frac{\lambda_s}{\phi\beta_s}\left[\frac{N_n^s}{N_e}\,\epsilon_n f_t\right]. \tag{5}$$

Here the expression in the brackets means the absolute neutron yield for the steady state of the $d-d$ fusion cycle, λ_s is the slope of the slow exponent in (4), and β_s is the partial probability of the reaction (1). To a good approximation (better than 1%) $\beta_s=\beta_{nr}=0.53$, where β_{nr} corresponds to the nonresonant $dd\mu$ formation.

In the expression for the absolute neutron yield N_n^s is the number of neutron events in the ''slow'' component of the time spectrum (4), N_e is the number of electrons indicated in Table I, f_t allows for the finite time interval for detection of a fusion neutron followed by a μ-decay electron, and ϵ_n is the neutron detection efficiency. The latter was calculated using two ''independent'' Monte Carlo codes. One code[20] was written specially for our experimental $d-d$ program and the other used the standard package GEANT[21] and the low-energy neutron cross sections therein. The results of the two codes coincide within 3–5%. To determine the efficiency loss due to the finite threshold, the calculated recoil proton energy spectrum was reconciled with the experimental distribution. This procedure was repeated for data of each run and the example for $T=19.0$ K is given in Fig. 8. The spikes in the spectrum are due to ADC differential nonlinearity.

The value of λ_{ss}, as well as the slope λ_s and the ratio b_f/b_s (4), are found from the fit used for the numerical solution of the set of differential equations referred to above. For the partial probability of the reaction (1) $\beta_r=0.58$ was used for the resonant $dd\mu$ formation ($F=3/2$) and $\beta_{nr}=0.53$ was used for the nonresonant formation[22] ($F=1/2$).

TABLE II. Experimental results.

	The rates of the $dd\mu$ molecule formation and $d\mu$ atom hyperfine transition rate, μs^{-1}			
Temperature, K	$\lambda_{1/2}$	$\lambda_{3/2}$	$\lambda_{3/2}/\lambda_{1/2}$	λ_d
5.5	0.0448(18)	2.48(13)	55.3(1.8)	31.7(1.0)
9.9	0.0403(20)	2.11(14)	52.3(2.3)	29.3(1.2)
15.1	0.0424(20)	2.27(16)	53.5(2.4)	32.5(1.5)
17.7	0.0419(18)	2.24(11)	53.4(1.7)	32.8(1.2)
19.0	0.0407(21)	2.27(14)	55.8(1.9)	30.2(1.0)
26.3	0.0389(20)	3.03(20)	77.8(2.4)	36.1(1.4)
30.5	0.0428(26)	3.20(21)	74.8(2.4)	32.0(1.1)

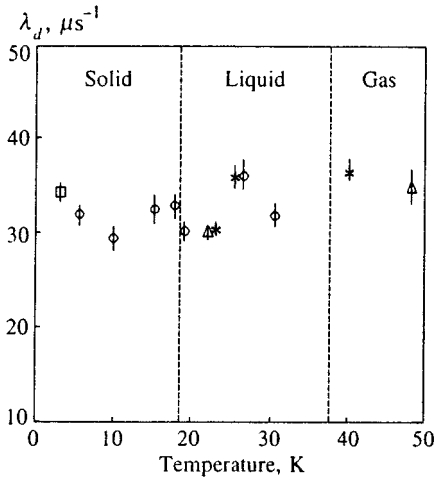

FIG. 9. Experimental data on the $d\mu$ atom spin-flip rate, λ_d. Square—Ref. 7; circles—present work; triangles—previous Dubna measurements;[6] stars—Refs. 4 and 5.

This procedure then gave the values of $\lambda_{3/2}$, $\lambda_{1/2}$, and λ_d. There were a few small corrections (few percent), e.g., for the loss due to $n-\gamma$ separation.

5. RESULTS AND DISCUSSION

The experimental results are presented in Table II.

The values in brackets are the errors due only to statistics, the fit, and the corrections and do not include systematic uncertainties from ϕ (3%) and ϵ_n (8%). Our results are shown in Figs. 2 and 9, together with the data of other authors. Again, our data are given without systematic errors in order to show more clearly their dependence on temperature.

The data for the ''second detected'' neutrons (N_1-N_2 and N_2-N_1) were also analyzed. Their time distribution relative to the ''first detected'' neutrons is shown in Fig. 10. The data are summed over all exposures with solid deuterium. The curve in this figure corresponds to exponents with the fast and slow slopes found above in the analysis of the ''first'' neutrons. The measured spectra are in satisfactory agreement with the predicted spectra.

The value of $\lambda_{1/2}^{(2)}$ was found from the analysis of the second neutrons normalized to the number of ''first'' neutrons:

$$\lambda_{1/2}^{2}=0.041(3)\ \ \mu s^{-1}\quad \text{(statistical error only)}.$$

This value agrees with that obtained for the ''first'' neutrons.

As can be seen from Fig. 2, our data for $\lambda_{1/2}$ are in agreement with our previous measurement using liquid deuterium at $T=22$ K and with the results of the PSI group at $T=23$ K; the latter initially gave $\lambda_{1/2}=0.0500(34)\times(22)\ \mu s^{-1}$ (Ref. 4) and then $\lambda_{1/2}=0.045(5)\ \mu s^{-1}$ (Ref. 10). The experimental results are in excess of theoretical prediction,[9] $\lambda_{1/2}^{(th)}\sim 0.03\ \mu s^{-1}$.

Of course, our main result is the measurements of the $dd\mu$ molecule formation rate from the upper $d\mu$ atom spin state for the lowest deuterium temperatures. Together with the pioneering result of the TRIUMF group,[7] they sharply contradict the ''standard'' theory, according to which only the nonresonant $dd\mu$ formation from the $d\mu$ atom spin states can contribute at the lowest temperatures. Possible mechanisms to explain it are considered in Refs. 18, 22, and 23.

According to Ref. 22, the $d\mu$ atoms moving in solid deuterium have insufficient time to fully thermalize, because they lose energy only in inelastic interactions with the lattice excitation. Significant $dd\mu$ formation therefore occurs at higher than thermal energies. This effect can explain the experimental data qualitatively[18] but quantitative agreement with the experiment is achieved only for few definite values of the inelastic cross sections. This mechanism could be investigated by repeating the experiment with enhanced protium.

Another possible explanation involves transitions with negative $d\mu$ atom resonance energy for $dd\mu$ formation[23] with the transfer of the released energy to lattice excitation. The transition $K_i=1\rightarrow K_f=0$ is appropriate for this scheme. The liquid and solid deuterium in this experiment were held at room temperature ortho-para ratio because equilibration is so slow. The experiment should be repeated with catalyzed $p\rightarrow\sigma$ to, among other things, check this mechanism, because with pure ortho-deuterium[23] only transitions with positive resonance energies are possible (3).

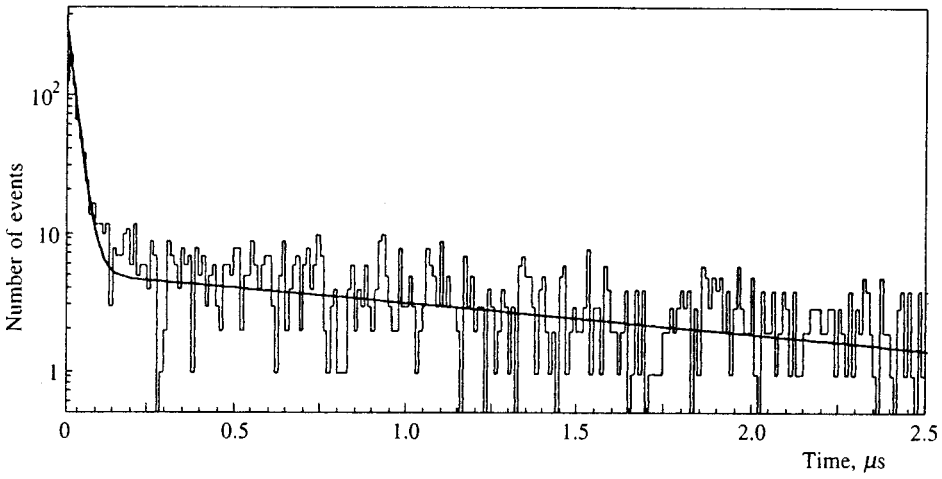

FIG. 10. Time distribution of the second detected neutron accumulated for all exposures with solid deuterium. Line is the function (4) with the exponent slopes obtained from the analysis of the ''first detected neutrons.''

Of course, both mechanisms can explain the experimental results, but, as was pointed out in Ref. 18, the data for solid and liquid deuterium are nearly identical. Perhaps the problem is more complicated and needs a more complete consideration. Incidentally, the enhanced rate is independent of the structure of the solid deuterium lattice, i.e., the TRIUMF group forms fcc solid deuterium directly from the gas phase, whereas our solid deuterium has hcp structure since it comes from the liquid.

Finally, the experimental data on λ_d are given in Fig. 9. Our results are in agreement with the previous measurements, both in the solid and liquid deuterium. At the same time, full set of the data for $T \leqslant 30$ K does not show such a sharp difference from the results for gaseous deuterium as was manifested in the first measurement in liquid deuterium.

We wish to thank E. P. Krasnoperov for developing the solid-deuterium target. We also thank M. M. Petrovsky' and A. P. Kustov for assistance with the tests and the runs.

The work was performed under the assistance of the Russian Fund for Fundamental Research and INTAS (Brussels).

[1] L. I. Menshikov, L. I. Ponomarev, T. A. Strizh, and M. P. Faifman, Zh. Éksp. Teor. Fiz. **92**, 1173 (1987) [Sov. Phys. JETP **65**, 656 (1987)].
[2] M. P. Faifman, Muon Cat. Fusion **2**, 247 (1988). M. P. Faifman, L. I. Menshikov, and T. A. Strizh, Muon Cat. Fusion mcf **4**, 1 (1989).
[3] A. Scrinzi, Muon Cat. Fusion **5/6**, 179 (1990).
[4] N. Nagele, W. H. Breunlich, M. Cargnelli *et al.*, Nucl. Phys. A **493**, 397 (1989).
[5] J. Zmescal, W. H. Breunlich, M. Cargnelli *et al.*, Muon Cat. Fusion **1**, 109 (1987). W. H. Breunlich, M. Cargnelli, P. Kammel *et al.*, Muon Cat. Fusion **5/6**, 149 (1987).
[6] V. P. Dzhelepov, V. G. Zinov, S. A. Ivanovsky' *et al.*, Zh. Éksp. Teor. Fiz. **101**, 1105 (1992) [Sov. Phys. JETP **74**, 589 (1992)], Muon Cat. Fusion **7**, 387 (1992).
[7] P. Kammel, in *Muonic Atoms and Molecules, Proc. of the Int. Workshop at the Centro Stefano Franchim on Monte Verita*, ed. by L. Schaller, C. Petitjean, Birkhauser Verlag, CH-4010, Basel, Boston, Berlin (1993), p. 111. P. E. Knowles, J. M. Bailey, G. A. Beer *et al.*, Hyp. Int. **101/102**, 21 (1996).
[8] D. L. Demin, V. P. Dzhelepov, V. V. Filchenkov *et al.*, Hyp. Int. **101/102**, 13 (1996).
[9] M. P. Faifman, Muon Cat. Fusion **4**, 341 (1989).
[10] A. Scrinzi, P. Kammel, J. Zmeskal *et al.*, Phys. Rev. A **47**, 4691 (1993).
[11] D. L. Demin, V. P. Dzhelepov, N. N. Grafov *et al.*, Hyp. Int. **101/102**, 583 (1996).
[12] V. P. Dzhelepov, V. V. Filchenkov, A. D. Konin *et al.*, Nucl. Instr. and Meth. A **269**, 634 (1988).
[13] V. G. Zinov, E. Lonsky, and A. I. Rudenko, Prib. Tech. Exp. No. **1**, 91 (1991).
[14] A. Adamchak and V. S. Melezhik, Muon Cat. Fusion **5/6**, 303 (1990).
[15] J. B. Kraiman, W. H. Breunlich, M. Cargnelli *et al.*, Muon Cat. Fusion **5/6**, 43 (1990), Phys. Rev. Lett. **63**, 1942 (1989).
[16] E. A. Vesman, JETP Lett. **5**, 91 (1967).
[17] M. P. Faifman, T. A. Strizh, E. A. G. Armour *et al.*, Hyp. Int. **101/102**, 179 (1996).
[18] V. V. Filchenkov, Hyp. Int. **101/102**, 37 (1996).
[19] V. V. Filchenkov, Communications JINR E1-89-57, Dubna (1989).
[20] V. V. Filchenkov and L. Marczis, Communications JINR E13-88-56, Dubna (1988).
[21] CERN Program Library, W5013.
[22] A. Adamchak, Hyp. Int. **101/102**, 113 (1996).
[23] L. I. Menshikov and V. V. Filchenkov, Hyp. Int. **101/102**, 207 (1996).

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

Resonance fluorescence of a two-level atom excited by a superposition of coherent states, and the instability of the average atomic dipole moment

S. Ya. Kilin and V. N. Shatokhin

B. I. Stepanov Institute of Physics, Belarus Academy of Sciences, 220072 Minsk, Republic of Belarus
(Submitted 21 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1174–1189 (April 1997)

We find the evolution of average atomic variables in the resonance fluorescence of a two-level atom excited by a superposition of coherent states shifted in phase by π. A new effect is predicted, the quantum instability of the average atomic dipole moment, with a strong correlation between atom and field being the reason. We propose different ways of verifying the effect in experiments involving high-Q optical and microwave cavities. © *1997 American Institute of Physics.* [S1063-7761(97)00304-1]

1. INTRODUCTION

The interest in the so-called nonclassical states of optical fields (squeezed fields, superposition fields, and the like, which requires using the language of quantum theory for their description), which resurged at the end of the 1970s, is still evident today.[1,2] In many respects the interest is due to the important applications that such fields have in extremely precise measurements, optical communications, and quantum computers. The study of these applications has led to fundamental questions, e.g., questions concerning the interaction between fields in nonclassical states and atoms and molecules.

Resonance fluorescence, a phenomenon of QED in which the quantum nature of optical fields manifests itself most vividly, reveals many aspects of such interactions. It has been intensively studied both theoretically and experimentally (a review of the work done prior to 1981 can be found in Ref. 3). The standard model for describing resonance fluorescence uses a two-level atom interacting with a set of electromagnetic modes of free space acting as a reservoir for the atom. In describing the excitation of the atom it is usually assumed that one mode is initially coherent. However, if we assume that excitation is done by nonclassical light, new features of resonance fluorescence not observed in excitation by coherent light manifest themselves. For instance, Gardiner[4] discussed the decay of a two-level atom into a wide-band vacuum and showed that one component of the atomic polarization decays at a rate much lower than the rate of spontaneous decay into an ordinary vacuum. Carmichael *et al.*[5] examined the problem of the spectrum of resonance fluorescence into a squeezed vacuum and discovered a strong dependence of the width of the central peak of the fluorescence spectrum on the phase of the driving field. Smart and Swain[6,7] reported the existence, for a narrow band of Rabi frequencies, of unusually shaped fluorescence spectra of a two-level atom interacting with a squeezed vacuum. They concluded that the anomalous profiles of the fluorescence spectra are a specific feature of the interaction between atomic systems and nonclassical fields. The statistical properties of the resonance fluorescence excited by the nonclassical light from a second-harmonic generator and a degenerate parametric amplifier in the low-intensity limit were discussed by Lyublinskaya and Vyas.[8]

The new results obtained by the authors of Refs. 4–8 are based on the traditional approach to studies of resonance fluorescence. In this approach the interaction Hamiltonian is written in the dipole and rotating-wave approximations. In some cases a fairly full description of the atom–field interaction can be achieved by using the quantum master equation for the density matrix averaged over the field states (the reduced atomic density matrix). The effectiveness of this method is due to the use of a wide-band reservoir spectrum in deriving the equation, an approach known as the Markov approximation.[3–14] In this approximation the quantum regression theorem is true, which makes it possible, among other things, to study the behavior of many-time atomic correlation functions. In examining resonance fluorescence, the authors of Ref. 13 developed a method for deriving a chain of coupled equations for the operators whose averages are the spectral field correlation functions of various orders.

A characteristic feature of this research in the resonance fluorescence excited by nonclassical light is the perturbative approach to studying the interaction between nonclassical light and an atom. The solution of the problem for arbitrary intensities of the nonclassical light is complex because one is forced to calculate all higher-order moments of the atom–field correlators. Note that for arbitrary light intensities the problem has been solved only in the special case of coherent radiation,[9,10,12,14] the reason being that a coherent state, which is an eigenstate of the annihilation operator, does not change because of one-photon absorption processes and the state of the driving field remains coherent. This is not the case for other states of the exciting field, and the traditional approaches require generalization.

In this paper we examine the excitation of fluorescence by nonclassical light that is a quantum superposition of two coherent states with identical amplitudes, but shifted in phase (SCS). When the amplitude is large, these nonclassical states are of the ''Schrödinger's cat'' type,[2,15–22] which makes them highly important for experimental verification of the principles of quantum mechanics. Many methods of generating SCS have been proposed: by a nonlinear wave process,[15] by continuous photodetection,[16] by negative feedback in a Kerr medium,[17] by four-wave mixing,[18] and by nonresonant

 1063-7761/97/040647-09$10.00

interactions of Rydberg atoms and a field in a microwave cavity.[19]

Buzek *et al.*[20] found that the nonclassical properties of SCS, such as squeezing and the sub-Poisson photon statistics, emerge because of quantum interference of the coherent states. Quantum interference not only explains the above properties but also changes substantially the nature of atom–field interactions in resonance fluorescence. Below we show that for a certain threshold amplitude of the initial coherent state the average dipole moment of an atom interacting with an SCS becomes unstable. Moreover, the Rabi oscillations of the average dipole moment are suppressed for all values of the amplitude. To clarify the dynamics of the atom we use the Heisenberg equation method. This method makes it possible to interpret the emerging effect as the result of quantum interference and correlated atom–field dynamics in a much fuller way than does the averaging method[21] employed earlier.

The plan of the paper is as follows. In Sec. 2 we discuss some properties of SCS. A model for the interaction with the reservoir in the Heisenberg picture is set up in Sec. 3. In Sec. 4 we discuss the dynamics of atomic averages. There we also show that the discovered instability in the average dipole moment of the transition is due to the quantum interference of the probability amplitudes of the initial field state, and we suggest a way of experimentally verifying this effect.

2. NONCLASSICAL PROPERTIES OF SUPERPOSITIONS OF COHERENT STATES

Let us take a discrete superposition of coherent states with identical amplitudes:

$$|\Psi\rangle = A^{1/2}\sum_{k=1}^{n} e^{i\phi_k}|\alpha e^{i\vartheta_k}\rangle, \tag{1}$$

where A is the normalization constant, and the phases ϕ_k and ϑ_k can be arbitrary. The quantum interference in the phase space is either constructive or destructive depending on the values of these phases.[2,20] The state (1) is a special case of the idea, put forward by Jansky and Vinogradov[22], of a continuous superposition around a circumference, a superposition to be generated for molecular vibrations. Let us restrict our discussion to the case where $n=2$, $\phi_1=\vartheta_1=0$, $\phi_2=\phi$, and $\vartheta_2=\pi$. The associated state

$$|\chi\rangle = \{2[1+\exp(-2|\alpha|^2)\cos\ \phi]\}^{-1/2}(|\alpha\rangle + |-\alpha\rangle e^{i\phi})$$
$$= N^{-1/2}(|\alpha\rangle + |-\alpha\rangle e^{i\phi}) \tag{2}$$

is the simplest example of an SCS. It is an eigenstate not of the photon annihilation operator, as it is in the case of a coherent state, but of the square of that operator. Notwithstanding the apparent simplicity, this state exhibits various nonclassical properties, which emerge because of the quantum interference between the coherent states. Most methods proposed for generating SCS deal with states of type (2). We see that the definition of the states contains an additional phase factor $e^{i\phi}$, which affects the quantum statistical properties of the superposition.

The case $\phi=0$ corresponds to even-parity $(+)$ coherent states and the case $\phi=\pi$ to odd-parity $(-)$ coherent states,

$$\{2[1\pm\exp(-2|\alpha|^2)]\}^{-1/2}(|\alpha\rangle \pm |-\alpha\rangle), \tag{3}$$

while $\phi=\pi/2$ corresponds to Yurke–Stoler states

$$2^{-1/2}(|\alpha\rangle + i|-\alpha\rangle). \tag{4}$$

As shown in a Refs. 2 and 20, notwithstanding that the differences are only in the values of the phase factor, the nonclassical properties of the above three states differ considerably. For example, even-parity states can contain only an even number of photons, odd-parity states an odd number of photons, and in Yurke–Stoler states the photon statistics is of the Poisson type. Even-parity coherent states have super-Poisson photon statistics and display second-order squeezing, while odd-parity states are characterized by sub-Poisson statistics and display no second-order squeezing. As for Yurke–Stoler states, they display second- and fourth-order squeezing.

Superposition quantum states exhibit other properties inherent in nonclassical states of light. In particular, the interference of the probability amplitudes of the states $|\alpha\rangle$ and $|-\alpha\rangle$ in their superposition makes the Wigner quasiprobability function

$$W(\beta) = \frac{1}{\pi N}\{e^{-2|\beta-\alpha|^2} + e^{-2|\beta+\alpha|^2}$$
$$+ 2e^{-2|\beta|^2}\cos[\phi + 4\ \mathrm{Im}(\beta\alpha^*)]\} \tag{5}$$

of state $|\chi\rangle$ (Ref. 2) negative in the quantum-interference region. The presence of quantum interference of the $|\chi\rangle$ states manifests itself in another unexpected way in the Glauber–Sudarshan distribution function $P(\beta)$, which has the form (see the Appendix)

$$P(\beta) = N^{-1}\Big\{\delta(\beta-\alpha) + \delta(\beta+\alpha)$$
$$+ \frac{e^{-2|\alpha|^2}}{|\beta|}\delta(\varphi_\alpha - \varphi_\beta + \pi/2)(e^{+i\phi}\delta_{AC}(|\beta| - i|\alpha|)$$
$$+ e^{-i\phi}\delta_{AC}(|\beta| + i|\alpha|))\Big\}. \tag{6}$$

The first two terms on the right-hand side coincide, to within a normalization constant, with the P-functions of the coherent states $|\alpha\rangle$ and $|-\alpha\rangle$, and the two remaining terms represent the interference part of the P-function and characterize the purely quantum properties of the states. The last two terms involve the singular function

$$\delta_{AC}(x-z) = e^{-z\partial/\partial x}\delta(x) = \lim_{\varepsilon\to 0}\frac{1}{\sqrt{\pi\varepsilon}}e^{-(x+z)^2/\varepsilon}, \tag{7}$$

which is even more singular than Dirac's delta-function. Since integrating this function with a function $F(x)$ of a real variable x,

$$\int_{-\infty}^{\infty} dx\,\delta_{AC}(x-z)F(x) = F(z), \tag{8}$$

amounts to analytically continuing $F(x)$ into the complex z plane (z in (8) is a complex-valued quantity), $\delta_{AC}(x-z)$ has

become known as the generalized function of analytic continuation.[21] Equation (6) shows that when a superposition state is averaged with the P-function,[1)] the interference terms lead to the replacement of the absolute value of the complex amplitude, $|\beta|$, in the averaged expression by the imaginary quantity $i|\alpha|$. Below we show that this property of the P-function (6), which reflects a property of quantum interference, results in the instability of the average dipole moment of an atom excited by a superposition of coherent states.

3. A MODEL OF INTERACTION WITH A RESERVOIR. HEISENBERG PICTURE

Let us examine a quantum system consisting of a single two-level atom and a set of field modes comprising a reservoir. To simplify our reasoning we assume that the field is inside a spherical cavity (the quantization volume is a sphere of radius R), with the atom at the center of the sphere. For such a geometry the spectrum of the eigenfrequencies of the field modes with which the atom interacts is nearly equidistant[23] with a mode separation equal to $c\pi/R$.

Initially the atom is unexcited and is in state $|1\rangle$; the field modes, with the exception of the resonant mode, are in the vacuum state. The resonant mode, labeled r, is excited to the superposition state $|\chi\rangle$ (Eq. (2)), which for the sake of convenience will be denoted by $|\alpha_+\rangle$:

$$|\alpha_+\rangle=(N_+)^{-1/2}(|\alpha\rangle+|-\alpha\rangle e^{i\phi}), \tag{9}$$

$$N_+=2(1+\exp(-2|\alpha|^2)\cos\ \phi). \tag{10}$$

We will also need the state $|\alpha_-\rangle$

$$|\alpha_-\rangle=(N_-)^{-1/2}(|\alpha\rangle-|-\alpha\rangle e^{i\phi}), \tag{11}$$

$$N_-=2(1-\exp(-2|\alpha|^2)\cos\ \phi), \tag{12}$$

which has an additional phase π in comparison to $|\alpha_+\rangle$. At $t=0$ the atom and field begin to interact. The total system Hamiltonian in the electric-dipole and rotating-wave approximations has the form

$$\hat H=\hat H_{0F}+\hat H_{0A}+\hat H_{IAF}, \tag{13}$$

where

$$\hat H_{0F}=\hbar\sum_k\ \omega_k\hat a_k^+\hat a_k,$$

$$\hat H_{0A}=\hbar\omega_{21}\hat\sigma_+\hat\sigma_-,$$

$$\hat H_{IAF}=\hbar\sum_k\ (\hat p_k^{(+)}\hat a_k^+ +\hat p_k^{(-)}\hat a_k)$$

are, respectively the Hamiltonian operators of the free field, the free atom, and the atom–field interaction. Here $\hat a_k$ and $\hat a_k^+$ are the Bose operators of photon annihilation and creation in mode k, $\hat\sigma_+$ and $\hat\sigma_-$ are the Pauli spin operators, which obey the commutation relations

$$\{\hat\sigma_+,\hat\sigma_-\}=1,\quad [\hat\sigma_+,\hat\sigma_-]=\hat\sigma_z,$$

with $\hat\sigma_z$ the atomic inversion operator, and $\hat p_k^{(+)}=g_k\hat\sigma_-$ and $\hat p_k^{(-)}=g_k^*\hat\sigma_+$ are coupling constants. In accordance with the above conditions, at $t=0$ the density operator for the atom–field system can be written as

$$\hat\rho(0)=\hat\rho_A(0)\hat\rho_F(0)\equiv|\psi_+\rangle\langle\psi_+|,$$
$$|\psi_+\rangle=|1\rangle_A|\alpha_+\rangle\prod_{k\neq r}\ |0\rangle_k\equiv|1\rangle|\alpha_+\rangle|\{0\}\rangle. \tag{14}$$

The Heisenberg equations for the system's operators are

$$\dot{\hat a}_k=-\frac{i}{\hbar}[\hat a_k,\hat H]=-i\omega_k\hat a_k-ig_k\hat\sigma_-, \tag{15.1}$$

$$\dot{\hat\sigma}_-=-i\omega_{21}\hat\sigma_-+i\sum_k\ g_k^*\hat a_k\hat\sigma_z, \tag{15.2}$$

$$\dot{\hat\sigma}_z=2i\sum_k\ (\hat p_k^{(+)}\hat a_k^+-\hat p_k^{(-)}\hat a_k). \tag{15.3}$$

Plugging the solution of Eq. (15.1),

$$\hat a_k(t)=\hat a_k(0)e^{-i\omega_k t}-ig_k\int_0^t d\tau\ \exp(-i\omega_k(t-\tau))\hat\sigma_-(\tau) \tag{16}$$

into Eqs. (15.2) and (15.3), we arrive at the following system of equations from which the Heisenberg field operators $\hat a_k(t)$ have been eliminated:

$$\dot{\hat\sigma}_-=-i\omega_{21}\hat\sigma_-+i\sum\ g_k^*\hat a_k(0)\exp(-i\omega_k t)\hat\sigma_z$$
$$+\int_0^t d\tau\sum_k\ |g_k|^2\exp(-i\omega_k(t-\tau))\hat\sigma_-(\tau)\hat\sigma_z(t), \tag{17.1}$$

$$\dot{\hat\sigma}_z=2i\sum_k\ (\hat p_k^{(+)}a_k^+(0)\exp(i\omega_k t)-\hat p_k^{(-)}\hat a_k(0)$$
$$\times\exp(-i\omega_k t))-2\int_0^t d\tau\sum_k\ |g_k|^2\exp(-i\omega_k(t-\tau))$$
$$\times\{\hat\sigma_+(\tau)\hat\sigma_-(t)+\hat\sigma_-(\tau)\hat\sigma_+(t)\}. \tag{17.2}$$

The standard approach to solving the above system of equations is to go to the free-space limit $R\to\infty$. In the one-dimensional case this amounts to the following replacement (see, e.g., Ref. 14):

$$\sum_k\ |g_k|^2\cdots\approx\int_0^\infty d\omega|g(\omega)|^2\rho(\omega)\cdots, \tag{18}$$

where $\rho(\omega)$ is the mode density, equal to $R/2\pi c$ in the case of the equidistant spectrum of a spherical cavity. Evaluation of the integrals in the Markov approximation, which amounts to replacing the correlation function of atomic perturbations,

$$K(t-\tau)=\sum_k\ |g_k|^2\exp(it(\omega_{21}-\omega_k)(t-\tau)),$$

by the delta function $\Gamma\delta(t-\tau)$, reduces the system of equations (17) to

$$\dot{\hat{\sigma}}_- = -\Gamma\hat{\sigma}_- + i\sum_k g_k^*\hat{\sigma}_z\hat{a}_k(0)\exp(i(\omega_{21}-\omega_k)t), \tag{19.1}$$

$$\dot{\hat{\sigma}}_z = 2i\sum_k [\hat{p}_k^{(+)}(t)\hat{a}_k^+(0)\exp(-i(\omega_{21}-\omega_k)t)] - 2i\sum_k [\hat{p}_k^{(-)}(t)\hat{a}_k(0)\exp(i(\omega_{21}-\omega_k)t)] - 2\Gamma, \tag{19.2}$$

where $\Gamma = \pi|g(\omega_{21})|^2\rho(\omega_{21})$ is the rate of spontaneous atomic decay, which coincides with the Einstein coefficient $A/2$. Using the commutation relations[12]

$$[\hat{a}_k(0),\ \hat{q}(t)] = ig_k\left[\int_0^t d\tau\ \exp(i\omega_k\tau)\hat{\sigma}_-(\tau),\hat{q}(t)\right], \tag{20}$$

which follow from Eq. (16), and the commutativity of the simultaneous field and atom operators, $\hat{a}(t)$and $\hat{q}(t)$, we can perform normal ordering of the field operators in Eqs. (19). The result is the following final system of equations for the atomic operators:

$$\dot{\hat{\sigma}}_- = -\Gamma\hat{\sigma}_- + i\hat{\sigma}_z\hat{L}_-(t), \tag{21.1}$$

$$\dot{\hat{\sigma}}_z = -2\Gamma(1+\hat{\sigma}_z) + 2i[\hat{L}_+(t)\hat{\sigma}_- - \hat{\sigma}_+\hat{L}_-(t)], \tag{21.2}$$

where

$$\hat{L}_-(t) = [\hat{L}_+(t)]^+ = \sum_k g_k^*\hat{a}_k(0)\exp(i(\omega_{21}-\omega_k)t)$$

are free-field operators, which act as operators of multiplicative noise in the Heisenberg equations (21).

The above system of operator equations cannot be solved exactly, but in some cases the averaged equation can be integrated. Obviously, this is determined by the initial state of the reservoir. In this paper we consider only one of these possibilities: the superposition of two coherent states, (2) or (9), as the initial state for the excited mode. But before we solve Eqs. (21), let us discuss the limits of applicability of the Markov approximation. As noted in the Introduction, this approximation leads to equivalent results in two physically distinct situations: the problem with a source of classical field, and the problem where the quantized field mode is initially prepared as a coherent state with a finite amplitude equal to the amplitude of the classical field. Since the presence of a source leads to temporal divergences in the average numbers of photons in the reservoir modes,[14] there emerges a contradiction between the final results (an infinite energy) and the initial conditions. To resolve this contradiction we assume that the coherent mode is heavily populated, so that in the quantum system interacting with the reservoir there is time for a steady state to set in before the coherent mode becomes depleted. In this sense the Markov approximation is applicable for all times. If, however, the problem is approached more rigorously, then for the time limit of the applicability of the Markov approximation one must take the time in the course of which the total calculated energy $E(t)$ in the reservoir modes reaches the initial value $E_0 = \hbar\omega_0|\alpha|^2$. This condition can be shown to be equivalent to the following inequality:

$$t \leqslant 2\pi\rho(\omega_0), \tag{22}$$

where $\rho(\omega_0)$ is the mode density at the frequency ω_0. For modes interacting with an atom at the center of a spherical cavity the density is $1/\Delta$, and the right-hand side of (22) is the cavity's ''revival'' time (see, e.g., Ref. 24) $T_R = 2R/c$, coinciding with the crossing time of the cavity.

4. INSTABILITY IN ATOMIC AVERAGES CAUSED BY QUANTUM INTERFERENCE

The system of quantum stochastic equations (21) has been described (see, e.g., Ref. 12) for the coherent state of an initially excited mode. The averaged equations in this case form a closed system of equations for three atomic averages: $\langle\hat{\sigma}_-\rangle$, $\langle\hat{\sigma}_+\rangle$, and $\langle\hat{\sigma}_z\rangle$. In the initial superposition state (9), there can be no reduction to only three atomic averages in the averaging of Eqs. (21) over the initial state, since $|\alpha_+\rangle$ is not an eigenstate of the annihilation operator. Indeed, let us set up an equation for the average matrix element

$$\langle\hat{\sigma}_-(t)\rangle = \mathrm{Tr}_{A+F}(\hat{\sigma}_-(t)\hat{\rho}_A(0)\hat{\rho}_F(0)) = \langle\psi_+|\hat{\sigma}_-(t)|\psi_+\rangle.$$

To this end we average the left- and right-hand sides of the operator equation (21.1) over the initial density operator. Because of the inhomogeneous term $i\hat{\sigma}_z\hat{L}_-(t)$ we have

$$\left\langle\psi_+\left|i\hat{\sigma}_z(t)\sum_k g_k^*\hat{a}_k(0)\exp(i(\omega_{21}-\omega_k)t)\right|\psi_+\right\rangle$$
$$= ig_r^*\langle\alpha_+|\langle 1|\langle\{0\}|\hat{\sigma}_z(t)|\{0\}\rangle|1\rangle\hat{a}_r(0)|\alpha_+\rangle$$
$$= i\alpha_+ g^*\langle\alpha_+|\langle 1|\langle\{0\}|\hat{\sigma}_z(t)|\{0\}\rangle|1\rangle|\alpha_-\rangle$$
$$= i\alpha_+ g^*\langle\psi_+|\hat{\sigma}_z|\psi_-\rangle,$$

where the states

$$|\psi_+\rangle = |1\rangle|\alpha_+\rangle|\{0\}\rangle \text{ and } |\psi_-\rangle = |1\rangle|\alpha_-\rangle|\{0\}\rangle \tag{23}$$

are transformed into each other by the resonant-field annihilation operator $\hat{a}$:

$$\hat{a}|\alpha_+\rangle = \alpha_+|\alpha_-\rangle, \quad \hat{a}|\alpha_-\rangle = \alpha_-|\alpha_+\rangle, \tag{24}$$

where $\alpha_+ = \alpha\sqrt{N_-/N_+}$, $\alpha_- = \alpha\sqrt{N_+/N_-}$, and the states $|\alpha_\pm\rangle$ and the normalization constants $N_\pm$ are specified by Eqs. (9)–(12). Thus, new variables related to the appearance of the new photon state $|\alpha_-\rangle$ emerge. By setting up equations for these variables we arrive at two independent systems of equations with six equations in each:

$$\langle\psi_+|\dot{\hat{\sigma}}_-|\psi_+\rangle = -\Gamma\langle\psi_+|\hat{\sigma}_-|\psi_+\rangle + i\alpha_+ g^*\langle\psi_+|\hat{\sigma}_z|\psi_-\rangle, \tag{25.1}$$

$$\langle\psi_+|\dot{\hat{\sigma}}_z|\psi_-\rangle = -2\Gamma(\langle\psi_+|\psi_-\rangle + \langle\psi_+|\hat{\sigma}_z|\psi_-\rangle) + 2i(\alpha_+^* g\langle\psi_-|\hat{\sigma}_-|\psi_-\rangle - \alpha_- g^*\langle\psi_+|\hat{\sigma}_+|\psi_+\rangle), \tag{25.2}$$

$$\langle\psi_-|\dot{\hat{\sigma}}_-|\psi_-\rangle=-\Gamma\langle\psi_-|\hat{\sigma}_-|\psi_-\rangle+i\alpha_-g^*\langle\psi_-|\hat{\sigma}_z|\psi_+\rangle, \tag{25.3}$$

$$\langle\psi_+|\dot{\hat{\sigma}}_+|\psi_+\rangle=(\langle\psi_+|\hat{\sigma}_-|\psi_+\rangle)^*, \tag{25.4}$$

$$\langle\psi_-|\dot{\hat{\sigma}}_z|\psi_+\rangle=(\langle\psi_+|\hat{\sigma}_z|\psi_-\rangle)^*, \tag{25.5}$$

$$\langle\psi_-|\dot{\hat{\sigma}}_+|\psi_-\rangle=(\langle\psi_-|\hat{\sigma}_-|\psi_-\rangle)^*, \tag{25.6}$$

and

$$\begin{aligned}\langle\psi_+|\dot{\hat{\sigma}}_z|\psi_+\rangle=&-2\Gamma(1+\langle\psi_+|\hat{\sigma}_z|\psi_+\rangle)\\&+2i(\alpha_+^*g\langle\psi_-|\hat{\sigma}_-|\psi_+\rangle\\&-\alpha_+g^*\langle\psi_+|\hat{\sigma}_+|\psi_-\rangle),\end{aligned} \tag{26.1}$$

$$\langle\psi_-|\dot{\hat{\sigma}}_-|\psi_+\rangle=-\Gamma\langle\psi_-|\hat{\sigma}_-|\psi_+\rangle+i\alpha_+g^*\langle\psi_-|\hat{\sigma}_z|\psi_-\rangle, \tag{26.2}$$

$$\begin{aligned}\langle\psi_-|\dot{\hat{\sigma}}_z|\psi_-\rangle=&-2\Gamma(1+\langle\psi_-|\hat{\sigma}_z|\psi_-\rangle)\\&+2i(\alpha_-^*g\langle\psi_+|\hat{\sigma}_-|\psi_-\rangle\\&-\alpha_-g^*\langle\psi_-|\hat{\sigma}_+|\psi_+\rangle),\end{aligned} \tag{26.3}$$

$$\langle\psi_+|\dot{\hat{\sigma}}_-|\psi_-\rangle=-\Gamma\langle\psi_+|\hat{\sigma}_-|\psi_-\rangle+i\alpha_-g^*\langle\psi_+|\hat{\sigma}_z|\psi_+\rangle, \tag{26.4}$$

$$\langle\psi_+|\dot{\hat{\sigma}}_+|\psi_-\rangle=(\langle\psi_-|\hat{\sigma}_-|\psi_+\rangle)^*, \tag{26.5}$$

$$\langle\psi_-|\dot{\hat{\sigma}}_+|\psi_+\rangle=(\langle\psi_+|\hat{\sigma}_-|\psi_-\rangle)^*. \tag{26.6}$$

The structure of these systems of equations is such that the average value of the dipole moment operator, $\langle\hat{\sigma}_-(t)\rangle\equiv\langle\psi_+|\hat{\sigma}_-(t)|\psi_+\rangle$, is related not to the average population difference $\langle\hat{\sigma}_z(t)\rangle\equiv\langle\psi_+|\hat{\sigma}_z(t)|\psi_+\rangle$ but to the interference matrix element $\langle\psi_+|\hat{\sigma}_z(t)|\psi_-\rangle$, equal to the correlation between the atomic inversion operator $\hat{\sigma}_z(t)$ at time t and the field operator $\hat{a}(0)$ at $t=0$,

$$\begin{aligned}\alpha_+\langle\psi_+|\hat{\sigma}_z(t)|\psi_-\rangle&\equiv\langle\hat{\sigma}_z(t)\hat{a}(0)\rangle\\&\equiv\langle\psi_+|\hat{\sigma}_z(t)\hat{a}(0)|\psi_+\rangle,\end{aligned}$$

and to the correlation between the number of photons at $t=0$ and the dipole moment operator at time t,

$$\langle\hat{a}^+(0)\hat{\sigma}_-(0)\hat{a}(0)\rangle=|\alpha_+|^2\langle\psi_-|\hat{\sigma}_-(t)|\psi_-\rangle.$$

In a similar way, the average value of the inversion operator is related to the interference matrix element $\langle\psi_+|\hat{\sigma}_-(t)|\psi_-\rangle$, which determines that the atom–field correlation $\langle\hat{\sigma}_-(t)\hat{a}(0)\rangle$, and to the correlation between the number of photons at $t=0$ and the atomic inversion at time t:

$$\langle\hat{a}^+(0)\hat{\sigma}_z(t)\hat{a}(0)\rangle=|\alpha_+|^2\langle\psi_-|\hat{\sigma}_z(t)|\psi_-\rangle.$$

Obviously, this structure of the equations, i.e., the formation of a closed system for the atomic $\langle\hat{A}(t)\rangle$ and atom–field $\langle\hat{B}(t)\hat{a}(0)\rangle$ and $\langle\hat{a}^+(0)\hat{B}(t)\hat{a}(0)\rangle$, stems from the fact that the initial field state is an eigenstate of the square of the annihilation operator. In the general case of an arbitrary initial state of a selected field mode, the averaged equations (21) form an infinite chain of equations for the atom–field normally ordered correlation functions $\langle(\hat{a}^+(0))^mB(t)\times(\hat{a}(0))^n\rangle$. Note that this chain reduces to a closed system of a finite number of equations for two classes of states. The first class consists of SCS discretely distributed around a circumference,

$$\hat{a}^n|\xi\rangle=\alpha^n|\xi\rangle,\quad n=0,1,\ldots,$$

and, as a degerenate case of SCS, ordinary coherent states. The second class consists of finite superpositions of Fock states, $|\zeta\rangle=\Sigma_{m=0}^n c_m|m\rangle$.

Solving the system of equations (25) and (26) can easily be reduced to evaluating integrals if we employ the Laplace transform. Under the given initial conditions,

$$\begin{aligned}\langle\psi_+|\hat{\sigma}_-(0)|\psi_+\rangle=\langle\psi_+|\hat{\sigma}_-(0)|\psi_-\rangle=\langle\psi_-|\hat{\sigma}_-(0)|\psi_+\rangle\\=\langle\psi_-|\hat{\sigma}_-(0)|\psi_-\rangle=0,\end{aligned} \tag{27}$$

$$\langle\psi_+|\hat{\sigma}_z(0)|\psi_+\rangle=-1,\quad\langle\psi_+|\hat{\sigma}_z(0)|\psi_-\rangle=-\langle\psi_+|\psi_-\rangle,$$

the Laplace transforms of the solutions are

$$\begin{aligned}\langle\psi_\pm|\tilde{\sigma}_z(p)|\psi_\pm\rangle=&-\frac{1}{2p}\left[1+\frac{Q_+(p)}{2Q_-(p)}\left(1-\frac{N_\mp}{N_\pm}\right)\right.\\&\left.+\frac{Q_-(p)}{2Q_+(p)}\left(1+\frac{N_\mp}{N_\pm}\right)\right],\end{aligned} \tag{28}$$

$$\begin{aligned}\langle\psi_\pm|\tilde{\sigma}_-(p)|\psi_\pm\rangle=&\pm\alpha g^*\frac{1+2\Gamma/p}{N_\pm Q_-(p)}\\&\times\exp(-2|\alpha|^2)\sin\phi,\end{aligned} \tag{29}$$

where

$$Q_\pm(p)=(p+\Gamma)(p+2\Gamma)\pm4|\alpha g|^2 \tag{30}$$

are polynomials that determine the characteristic exponents in the time dependence of the atomic averages. As Eq. (29) implies, the time dependence of the average dipole moment,

$$\begin{aligned}\langle\hat{\sigma}_-(t)\rangle\equiv\langle\psi_+|\hat{\sigma}_-(t)|\psi_+\rangle=\alpha g^*(N_+)^{-1}\exp(-2|\alpha|^2)\\\times\sin\phi\ (s_1e^{p_1t}+s_2e^{p_2t}+s_0)\end{aligned} \tag{31}$$

is determined by the roots of the polynomial $Q_-(p)$,

$$p_{1,2}=-\frac{3\Gamma}{2}\pm\sqrt{4|\alpha g|^2+\frac{\Gamma^2}{4}}. \tag{32}$$

These differ from the roots of the polynomial $Q_+(p)$, which describe Rabi oscillations for the coherent initial state, in the sign in front of the absolute value of the Rabi frequency, $|\alpha g|^2\to-|\alpha g|^2$. The coefficients s_i in Eq. (31) are

$$s_{1,2}=\frac{1+2\Gamma/p_{1,2}}{p_{1,2}^2-p_1p_2},\quad s_0=\frac{\Gamma}{\Gamma^2-2|\alpha g|^2}. \tag{33}$$

The obtained solution for the time dependence of the average dipole moment $\langle\hat{\sigma}_-(t)\rangle$ differs drastically from the case of

an initially coherent excited mode. Since the roots p_1 and p_2 are real quantities, under the chosen initial conditions there is not a single value of the field amplitude for which the dipole moment oscillates. More than that, for $|\alpha g|^2 > \Gamma^2/2$ the root p_1 is positive, with the result that the average dipole moment is unstable.

The effect is the strongest at $\phi = \pi/2$, while at $\phi = 0$ or $\phi = \pi$, i.e., for even- or odd-parity coherent states, the average dipole moment is not excited at all. The reason is that at $\phi = \pi n$ the SCS are orthogonal and the average field amplitude at $t = 0$, or $\langle \hat{a}(0) \rangle$, is zero. In view of this and the chosen initial conditions, all the initial values of the averages in the system of equations (25) vanish, and so do the inhomogeneous terms in this system. Hence at $\phi = \pi n$ the solution of the system (25) is zero. At $\phi \neq \pi n$, in (25) there appear a nonzero average amplitude of the initial field state and nonzero inhomogeneous terms $\langle \psi_+ | \psi_- \rangle$ caused by the quantum interference of the coherent states $|\alpha\rangle$ and $|-\alpha\rangle$, which form the initial superposition state of the field. At $\phi = \pi n$ the contributions of the interference terms to the average field amplitude reduce each other (destructive interference), while at $\phi = \pi/2$ they amplify each other (constructive interference). In the process of interacting with the atom, the initial field quantum interference manifests itself in the time dependence of the average atomic dipole moment.

The emergence of interference terms is characteristic of averaging over states that have no classical analogs. The classical feature of a coherent state consists, in particular, in the fact that it does not change when a finite number of photons in this state are annihilated, since $\hat{a}^k |\alpha\rangle = \alpha^k |\alpha\rangle$. An SCS exhibits both purely quantum properties and properties that make it resemble a coherent state. Indeed, annihilation of a single photon transforms the SCS into a new state, in which the average field amplitude is opposite, to within a small correction, in sign to the average field strength in the initial state, i.e., in the process of the interaction with the atom the phase of the field spontaneously changes by π. On the other hand, the coherent state $|a\rangle$ and the states $|\alpha_\pm\rangle$ satisfy the equation $\hat{a}^2 |\xi\rangle = \alpha^2 |\xi\rangle$. Together these two features of SCS determine the dynamics of atomic averages.

All we have said above about the features of SCS is vividly illustrated by the type of the Glauber quasiprobability function of the state $|\alpha_+\rangle$ (see Eq. (6)). In comparison to other quasiprobability functions, knowing the P-function of the initial field state, $P(\beta,0)$, is especially important in studying the dynamics of the interaction of the atom and the quantized field: with allowance for the normally ordered nature of the Heisenberg equations (21), the solution of these equations for an arbitrary initial state of the field mode can be obtained by averaging, with the P-function of the initial field state, $P(\beta,0)$, the solution obtained for the initial coherent state $|\beta\rangle$ (Ref. 21):

$$\langle \hat{\sigma}_\mu(t) \rangle = \int d^2\beta P(\beta,0) \langle \beta | \widehat{\sigma_\mu}(t) | \beta \rangle, \quad \mu = \pm, z, \quad (34)$$

where

$$\widehat{\sigma_\mu}(t) = \mathrm{Tr}_A(\rho_A(0) \langle \{0\} | \hat{\sigma}_\mu(t) | \{0\} \rangle).$$

By using the well-known atomic averages $\langle \beta | \widehat{\sigma_\mu}(t) | \beta \rangle$ with an initial coherent field state (see, e.g., Ref. 12) we can find the evolution of atomic averages for all initial values of the field. For an atom that initially was in the ground state,

$$\langle \beta | \widehat{\sigma_-}(t) | \beta \rangle = \beta g^* (\tilde{s}_1 e^{\tilde{p}_1 t} + \tilde{s}_2 e^{\tilde{p}_2 t} + \tilde{s}_0), \quad (35)$$

where

$$\tilde{p}_{1,2} = -\frac{3\Gamma}{2} \pm i \sqrt{4|\beta g|^2 - \frac{\Gamma^2}{4}},$$

$$\tilde{s}_{1,2} = \frac{1 + 2\Gamma/\tilde{p}_{1,2}}{\tilde{p}^{\,2}_{1,2} - \tilde{p}_1 \tilde{p}_2}, \quad \tilde{s}_0 = \frac{\Gamma}{\Gamma^2 + 2|\beta g|^2}.$$

Upon averaging (35) with the P-function (6) we arrive at precisely expression (31). Here we see that dynamic instability in the dipole moment emerges because of the interference part of the P-function (6), whose action on an analytic function of the complex variable $\beta = |\beta| e^{i\varphi_\beta}$ reduces, as noted earlier, to the substitutions

$$|\beta| \to \pm i |\alpha|, \quad \varphi_\beta \to \varphi_\alpha + \pi/2,$$

i.e., the real absolute value of the field amplitude becomes pure imaginary and the phase increases by 90°.

The above ideas, which demonstrate the link between the instability of the dipole moment of the atom and the quantum interference in the initial field state, can be augmented by an interpretation involving the quantum dynamics of the field and atomic states. Let us examine the evolution of the reduced density operator $\hat{\pi}(t)$ of the system consisting of the atom and the resonant mode, the operator being averaged over the vacuum states of the other reservoir modes. For the initial state of the resonant mode we take the coherent state $|\beta\rangle$. To make the situation more transparent we ignore the atom's spontaneous decay ($\Gamma = 0$). Now it can easily be shown that in this case $\hat{\pi}(t)$ can be represented by the following dyadic product of pure states:

$$\hat{\pi}(t) = |\Psi(t)\rangle \langle \Psi(t)|, \quad (36)$$

where

$$|\Psi(t)\rangle = [\cos(|\beta g^*| t) |1\rangle + i \sin(|\beta g^*| t) |2\rangle] |\beta\rangle \equiv |\theta(\beta,t)\rangle |\beta\rangle. \quad (37)$$

Obviously, the density operator (36) can be represented by the product of atom and field density operators, with the state of the field remaining coherent with the passage of time. In other words, the states of the atom and field are independent.

Let us now assume that the SCS (9) is the initial state of the resonant mode. We denote the desired density operator by $\hat{\pi}^*(t)$. This operator can be obtained by integrating the expression (36) with the Glauber–Sudarshan function (6) (cf. (34)). After integration we obtain

$$\hat{\pi}^*(t)=A(|\alpha\rangle|\theta(\alpha,t)\rangle\langle\theta(\alpha,t)|\langle\alpha|+|-\alpha\rangle|\theta(\alpha,t)\rangle \times\langle\theta(\alpha,t)|\langle-\alpha|+|-\alpha\rangle|\tilde{\theta}_+(\alpha,t)\rangle\langle\tilde{\theta}_-(\alpha,t)| \times\langle\alpha|e^{i\phi}+|\alpha\rangle|\tilde{\theta}_-(\alpha,t)\rangle\langle\tilde{\theta}_+(\alpha,t)|\langle-\alpha|e^{-i\phi}), \tag{38}$$

where

$$|\tilde{\theta}_\pm(\alpha,t)\rangle=\cosh(|\alpha g^*|t)|1\rangle\pm\sinh(|\alpha g^*|t)|2\rangle.$$

In our case, just as for a coherent state, the field state does not vary in time, since

$$\mathrm{Tr}_A(\hat{\pi}^*(t))=|\alpha_+\rangle\langle\alpha_+|.$$

But now the density operator cannot be factorized: the states of the atom and the field become correlated. Indeed, on the basis of (38) we can easily show that the field state

$$\hat{\pi}_F(t)|_{|1\rangle\langle1|}=A(\cos^2(|\alpha g^*|t)(|\alpha\rangle\langle\alpha|+|-\alpha\rangle\langle-\alpha|)) +\cosh^2(|\alpha g^*|t)(|-\alpha\rangle\langle\alpha|e^{i\phi}+|\alpha\rangle \times\langle-\alpha|e^{-i\phi}), \tag{39}$$

corresponds to the atom in the ground state, and the state

$$\hat{\pi}_F(t)|_{|2\rangle\langle2|}=A(\sin^2(|\alpha g^*|t)(|\alpha\rangle\langle\alpha|+|-\alpha\rangle\langle-\alpha|)) -\sinh^2(|\alpha g^*|t)(|-\alpha\rangle\langle\alpha|e^{i\phi}+|\alpha\rangle \times\langle-\alpha|e^{-i\phi}), \tag{40}$$

to the atom in the excited state. Thus, this approximation clearly demonstrates that in the interaction of the atom with a superposition field the initially independent states of the atom and the field become ''entangled.'' Such entanglement also occurs in the interaction of a two-level atom and a field in a cavity (the Jaynes–Cummings model). Studies of the time evolution of the Wigner function in the ground and excited atomic states has revealed the existence of strong correlation between field and atom in such a quantum system. The phases of the characteristic interference structure in the Wigner functions of the ground and excited states are opposite, i.e., in the process of evolution the field follows the state of the atom, with the atom in the ground state corresponding to the field state $|\alpha_+\rangle$, and the atom in the excited state to the field state $|\alpha_-\rangle$. The emerging positive quantum feedback leads to instability in the average dipole moment.

The exponential growth of the average dipole moment can be observed in experiments that study the time dependence of the coherent component of scattered light, which, as is known (see, e.g., Ref. 9), is proportional to the square of the absolute value of the matrix element of the average dipole moment. However, the presence of the factor $\exp(-2|\alpha|^2)$, which ''destroys'' the interference part, and the time limit on the applicability of the model make the observation of the effect in resonance fluorescence problematic. For instance, on the basis of (35) we can easily estimate the time t it takes the matrix element of the average dipole moment to assume the measured values. This time proves to be longer than T_R but shorter than $2T_R$, which means it exceeds the time limit on the applicability of the results obtained in the present work. One could attempt to balance the exponentially small factor by an ensemble of atoms. However, this would require studying in detail the problem of collective scattering with allowance for interatomic correlations induced by the field and for the geometric factor. One way of avoiding these difficulties in observing the effect in free space is to use a continuous source of superposition light[2)] (which has yet to be developed), similar to a laser. The second way amounts to the following. At the initial stage of intracavity interactions, when the response of the emitted field and the attenuation of the field in the cavity are fairly weak, the dynamics of the atom is similar to the dynamics of resonance fluorescence in free space. Hence the predicted effect of quantum instability can also be observed for a single atom interacting with a resonant mode prepared in SCS, e.g., by the method suggested by Brune *et al.*[19] The existing experimental methods make it possibly to verify the effect in both the microwave range[25,26] and the optical range.[27] In microwave cavities the following values of the parameters have been attained:[25,26] the atom–field coupling constant $g=2\pi\times17$ kHz, the cavity attenuation rate $k=2\pi\times6.7$ kHz, and the spontaneous decay rate $\Gamma=2\pi\times5$ kHz. At such values of the values of the parameters, at an SCS amplitude $\alpha=3$, and for an interaction time that obeys the relationship $gt_{\mathrm{int}}=2\pi\times0.3$, the dipole moment exponentially increases to 0.26×10^{-6}, instead of the full period of Rabi oscillations as is the case with a coherent initial state of a field of the same amplitude. In an optical cavity,[27] where the values of the respective parameters are $g=2\pi\times7.2$ MHz, $k=2\pi\times0.7$ MHz, $\Gamma=2\pi\times2.5$ MHz, $\alpha=3$, and $gt_{\mathrm{int}}=2\pi\times0.3$, the dipole moment exponentially increases to 1.2×10^{-4}, instead of the full period of Rabi oscillations.

5. CONCLUSION

We have studied the interaction of a single two-level atom with a reservoir that has an equidistant spectrum whose resonant mode initially is excited to a superposition coherent mode, while the other modes are in a vacuum state. We have established the evolution of the averages of the atomic variables by the method of Heisenberg equations and by the method of averaging the solutions found for a coherent initial state with the Glauber quasiprobability function of the superposition initial state. The dynamics of the atom being excited by a superposition of two coherent fields differs significantly from the case of excitation by a coherent field. The differences are due to the quantum interference of the amplitudes of the probability of finding the field in states $|\alpha\rangle$ and $|-\alpha\rangle$, with the evolution of the average dipole moment being determined entirely by this quantity (mathematically this means that the Glauber P-functional has the appropriate interference terms, which determine the evolution of the dipole moment). The nature of the quantum interference depends on the phase shift ϕ between the coherent states. When the phase shift is zero or π, the average dipole moment is not excited, since the contributions of the interference terms extinguish each other, and when $\phi=\pi/2$, the average dipole moment is at its maximum because of mutual amplification of the interference contributions. The main result of interference here is that for Rabi frequencies that obey the inequality

$|\alpha g|>\Gamma/\sqrt{2}$ the average dipole moment becomes unstable. The physical reason for the quantum instability is that the correlation between atom and field in the interaction is strong. This results in the onset of positive feedback in the system.

To observe the effect in experiments we have proposed using a microwave or optical cavity in which field states in the form of a superposition of coherent fields are generated using well-known techniques. The average dipole moment of an atom placed in such a cavity must exponentially grow in the course of the interaction time, instead of the full period of Rabi oscillations, as is the case with a coherent initial state.

Partial financial support for this work was provided by the Belarus Fund for Fundamental Research (Grants Nos. F18-203 and MP96-38) and the International Soros Program for Education in the Exact Sciences. One of the authors (V.N.Sh.) is grateful to A. P. Nizovtsev, T. M. Maevskaya, and D. S. Mogilevtsev for their useful comments.

APPENDIX

The Glauber function $P(\beta)$ of an arbitrary field state is defined as the two-dimensional Fourier transform of the normally ordered characteristic function $\chi_N(\lambda)=\mathrm{Tr}(\hat{\rho}\ \exp(\lambda\hat{a}^+)\exp(-\lambda^*\hat{a}))$ (Ref. 1):

$$P(\beta)=\frac{1}{\pi^2}\int d^2\lambda\,\chi_N(\lambda)\exp(\lambda^*\beta-\lambda\beta^*). \tag{A1}$$

If the field mode is in the pure state $|\chi\rangle$ defined by Eq. (2), the characteristic function consists of four terms, two of which coincide, to within a normalization constant, with the characteristic functions of the coherent states $|\alpha\rangle$ and $|-\alpha\rangle$, while the other two correspond to the interference terms:

$$\chi_N^{\mathrm{coh}}(\lambda)=\frac{1}{N}[\exp(\lambda\alpha^*-\lambda^*\alpha)+\exp(-\lambda\alpha^*+\lambda^*\alpha)], \tag{A2}$$

$$\chi_N^{\mathrm{int}}(\lambda)=\frac{1}{N}[\exp(-\lambda\alpha^*-\lambda^*\alpha-2|\alpha|^2-i\phi)+\exp(\lambda\alpha^*+\lambda^*\alpha-2|\alpha|^2+i\phi)]. \tag{A3}$$

The total quasiprobability function is

$$P(\beta)=P_{\mathrm{coh}}(\beta)+P_{\mathrm{int}}(\beta). \tag{A4}$$

The Fourier transform of $\chi_N^{\mathrm{coh}}(\lambda)$ yields the coherent part of the quasiprobability and consists of two delta functions with singularities at the points $\beta=\alpha$ and $\beta=-\alpha$:

$$P_{\mathrm{coh}}(\beta)=\frac{1}{\pi^2}\int d^2\lambda\,\chi_N^{\mathrm{coh}}(\lambda)\exp(\lambda^*\beta-\lambda\beta^*)=N_+^{-1}(\delta(\beta-\alpha)+\delta(\beta+\alpha)). \tag{A5}$$

To calculate the interference part of the P-function we use the integral representation of integer-order Bessel functions,[28]

$$J_k(z)=\frac{1}{2\pi}\int_0^{2\pi}dt\ \exp(i(kt-z\ \sin\ t)), \tag{A6}$$

and the expression for the generating function of the modified Bessel functions,

$$\exp(z\ \cos\ t)=\sum_{k=-\infty}^{\infty}I_k(z)\exp(ikt). \tag{A7}$$

Combining (A6) and (A7), we arrive at the following expression for the Fourier transform $P_{\mathrm{int}}(\beta)$ of the function $\chi_N^{\mathrm{int}}(\lambda)$:

$$P_{\mathrm{int}}(\beta)=\frac{2e^{-2|\alpha|^2}}{N_+\pi}\sum_{k=-\infty}^{\infty}\exp(ik(\varphi_\alpha-\varphi_\beta))\times\int_0^{\infty}|\lambda|d|\lambda|(I_k(2|\lambda||\alpha|)J_k(2|\lambda||\beta|)\times\exp(i\phi)+I_k(-2|\lambda||\alpha|)J_k(2|\lambda||\beta|)\times\exp(-i\phi)). \tag{A8}$$

Using the identity that relates Bessel functions of real and imaginary arguments and is obtained by expanding these functions in power series,

$$\exp\left(\pm iy\frac{\partial}{\partial x}\right)J_k(x)\Big|_{x=0}=(\pm i)^kI_k(y), \tag{A9}$$

and a representation of the delta function derived from the Fourier–Bessel formula,[1]

$$\delta(x-y)=\int_0^{\infty}sxJ_k(sx)J_k(sy)ds, \tag{A10}$$

we arrive at the following formula for $P_{\mathrm{int}}(\beta)$:

$$\begin{aligned}P_{\mathrm{int}}(\beta)&=\frac{\exp(-2|\alpha|^2)}{N\pi}\sum_{k=-\infty}^{\infty}i^k\exp(ik(\varphi_\alpha-\varphi_\beta))\\&\times\left[\exp\left(i\phi-i|\alpha|\frac{\partial}{\partial x}\right)\frac{\delta(|\beta|-x)}{|\beta|}+\exp\left(-i\phi+i|\alpha|\frac{\partial}{\partial x}\right)\frac{\delta(|\beta|-x)}{|\beta|}\right]\Big|_{x=0}\\&=\frac{\exp(-2|\alpha|^2)}{N}\delta\left(\varphi_\alpha-\varphi_\beta+\frac{\pi}{2}\right)\left[\exp\left(i\phi-i|\alpha|\frac{\partial}{\partial x}\right)+\exp\left(-i\phi+i|\alpha|\frac{\partial}{\partial x}\right)\right]\frac{\delta(|\beta|-x)}{|\beta|}\Big|_{x=0}\\&=\frac{\exp(-2|\alpha|^2)}{N|\beta|}\delta\left(\varphi_\alpha-\varphi_\beta+\frac{\pi}{2}\right)\\&\times(\exp(i\phi)\,\delta_{AC}(|\beta|-i|\alpha|)+\exp(-i\phi)\,\delta_{AC}(|\beta|+i|\alpha|)).\end{aligned} \tag{A11}$$

[1)]To avoid misunderstanding we note that in view of the peculiar action of the analytic continuation function, all operations of complex conjugation in the averaged function must be performed prior to integration rather than after integration.

[2)]The existing lasing modes make it possible, at best, to obtain pulsed sources (see, e.g., Ref. 17) of macroscopic quantum superpositions of coherent states.

[1] S. Ya. Kilin, *Quantum Optics: Fields and Their Detection* [in Russian], Navuka i tékhnika, Minsk (1990).

[2] V. Buzek and P. L. Knight, in *Progress in Optics*, Vol. 34, E. Wolf (ed.), North-Holland, Amsterdam (1995).

[3] B. R. Mollow, in *Progress in Optics*, Vol. 19, E. Wolf (ed.), North-Holland, Amsterdam (1981).

[4] C. W. Gardiner, Phys. Rev. Lett. **56**, 1917 (1986).

[5] H. J. Carmichael, A. S. Lane, and D. F. Walls, J. Mod. Opt. **34**, 2539 (1987).

[6] S. Smart and S. Swain, Phys. Rev. A **48**, R50 (1993).

[7] S. Swain, Phys. Rev. Lett. **73**, 1493 (1994).

[8] I. E. Lyublinskaya and R. Vyas, Phys. Rev. A **48**, 3966 (1993).

[9] B. R. Mallow, Phys. Rev. **188**, 1969 (1969).

[10] H. J. Carmichael and D. F. Walls, J. Phys. B **9**, 1199 (1976).

[11] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 2nd ed., Springer, Berlin (1985), Chap. 10.

[12] S. Ya. Kilin, Preprint No. 152, Institute of Physics of the Belarus Academy of Sciences, Minsk (1978).

[13] P. A. Apanasevich and S. Ya. Kilin, Zh. Prikl. Spektrosk. **24**, 738 (1976); Phys. Lett. A **62**, 83 (1977).

[14] G. S. Agarwal, *Quantum Statistical Theories of Spontaneous Emission and Their Relation to Other Approaches*, Springer, Berlin (1974).

[15] B. Yurke and D. Stoler, Phys. Rev. Lett. **57**, 13 (1991).

[16] T. Ogawa, M. Ueda, and N. Imoto, Phys. Rev. A **43**, 6458 (1991).

[17] J. J. Slosser and G. J. Milburn, Phys. Rev. Lett. **75**, 418 (1995).

[18] S. Ya. Kilin and D. S. Mogilevtsev, Phys. Lett. A **198**, 85 (1995).

[19] M. Brune, P. Nussenzveig, F. Schmidt-Kaler, F. Bernadot, A. Maali, J. M. Raimond, and S. Haroche, Phys. Rev. Lett. **72**, 3339 (1994).

[20] V. Buzek, A. Vidiella Barranco, and P. L. Knight, Phys. Rev. A **45**, 6570 (1992).

[21] S. Ya. Kilin and V. N. Shatokhin, Phys. Rev. Lett. **76**, 1051 (1996).

[22] J. Jansky and A. V. Vinogradov, Phys. Rev. Lett. **64**, 2771 (1990).

[23] J. A. Stratton, *Electromagnetic Theory*, McGraw-Hill, New York (1941).

[24] N. B. Narozhny, J. J. Sanchez-Mondragon, and J. H. Eberly, Phys. Rev. A **22**, 226 (1981).

[25] G. Raithel, C. Wagner, H. Walther, L. M. Narducci, and M. O. Scully, in: *Advances in Atomic, Molecular, and Optical Physics*, Suppl. 2, P. Berman (ed.), Academic Press, New York (1994), p. 57.

[26] S. Haroche and J. M. Raimond, in *Advances in Atomic, Molecular, and Optical Physics*, Suppl. 2, P. Berman (ed.), Academic Press, New York (1994), p. 122.

[27] J. Kimble, in *Advances in Atomic, Molecular, and Optical Physics*, Suppl. 2, P. Berman (ed.), Academic Press, New York (1994), p. 203.

[28] *Handbook of Mathematical Functions*, M. Abramowitz and I. A. Stegun (ed.), National Bureau of Standards Applied Mathematics Series 55, Washington, D.C. (1964), Chap. 9.

Translated by Eugene Yankovsky

Dirac particle with an anomalous magnetic moment in a circularly polarized wave and in constant longitudinal magnetic and electric fields

O. Ya. Savchenko

Institute of Chemical Kinetics and Combustion, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 25 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1190–1193 (April 1997)

The wave function of an uncharged Dirac particle with an anomalous magnetic moment is calculated for the case where a circularly polarized wave propagating along constant magnetic and electric fields is present. © *1997 American Institute of Physics.* [S1063-7761(97)00404-6]

The wave function of a Dirac particle in an electromagnetic wave and a constant magnetic field directed along the wave was determined in Ref. 1 for the case where the particle has a charge but no anomalous magnetic moment. In the present paper the wave function is determined for the case where the Dirac particle has no charge but has an anomalous magnetic moment. A constant longitudinal electric field is also present.

The wave function of an uncharged Dirac particle with anomalous magnetic moment μ is determined by the following modified Dirac equation:[2]

$$(\gamma_L \hat{k}_i + k_0)\psi = \frac{i\mu}{2\hbar c} F_{ij}\gamma_{ij}\psi,$$
$$\hat{k}_i = \frac{\partial}{\partial x_i}, \quad k_0 = \frac{m_0 c}{\hbar}, \quad F_{ij} = \frac{\partial \Phi_j}{\partial x_i} - \frac{\partial \Phi_i}{\partial x_j}. \tag{1}$$

Here m_0 is the particle's mass, γ_i are the Dirac matrices,[2,3] and the Φ_i are the components of the field 4-vector, which in a right-handed circularly polarized wave and longitudinal magnetic and electric fields H and E are given by

$$\Phi_x = \Phi \sin \omega\eta - \frac{1}{2} Hy, \quad \Phi_y = \Phi \cos \omega\eta + \frac{1}{2} Hx,$$
$$\Phi_z = 0, \quad \Phi_t = iEz, \tag{2}$$

where $\omega\Phi$ is the field strength of the wave, $\omega c/2\pi$ is the wave's frequency, and $\eta = ct + z$. We seek a solution of Eq. (1) based on the divisor

$$\Gamma = \frac{1}{4}(1 + i\gamma_{12})(1 + \gamma_4) \tag{3}$$

in the form of the sum

$$\psi = \exp(ik_0 ct)(\psi_1 + \psi_2\gamma_1 + \psi_3\gamma_3 + \psi_4\gamma_{31})\Gamma. \tag{4}$$

The components ψ_i of this sum depend only on η, and in a rotating system of coordinates in which one axis is directed along the electric field, of the wave and the other along the magnetic field they can be found by solving the following system of equations:

$$g_\Phi e^{-i\omega\eta}(\psi_2 - i\psi_4) + \left(-g_E + \frac{1}{k_0}\frac{\partial}{\partial\eta}\right)\psi_3 + \left(2 - g_H - \frac{i}{k_0}\frac{\partial}{\partial\eta}\right)\psi_1 = 0,$$
$$-g_\Phi e^{i\omega\eta}(\psi_1 + i\psi_3) + \left(g_E + \frac{1}{k_0}\frac{\partial}{\partial\eta}\right)\psi_4 + \left(g_H + \frac{i}{k_0}\frac{\partial}{\partial\eta}\right)\psi_2 = 0,$$
$$g_\Phi e^{-i\omega\eta}(\psi_4 + i\psi_2) + \left(g_E + \frac{1}{k_0}\frac{\partial}{\partial\eta}\right)\psi_1 + \left(-g_H + \frac{i}{k_0}\frac{\partial}{\partial\eta}\right)\psi_3 = 0,$$
$$-g_\Phi e^{i\omega\eta}(\psi_3 - i\psi_1) + \left(-g_E + \frac{1}{k_0}\frac{\partial}{\partial\eta}\right)\psi_2 + \left(2 + g_H - \frac{i}{k_0}\frac{\partial}{\partial\eta}\right)\psi_4 = 0, \tag{5}$$

where g_Φ, g_E, and g_H are the ratios of $\mu\omega\Phi$, μE, and μH to $m_0 c^2$, the particle's self-energy. The following functions comprise the solution of this system of equations:

$$\psi_1 = i(a_1^+ + a_1^-)\exp[i(\lambda - \omega/2)\eta],$$
$$\psi_2 = i(a_2^+ + a_2^-)\exp[i(\lambda + \omega/2)\eta],$$
$$\psi_3 = (a_1^+ - a_1^-)\exp[i(\lambda - \omega/2)\eta],$$
$$\psi_4 = (a_2^+ - a_2^-)\exp[i(\lambda + \omega/2)\eta]. \tag{6}$$

The constants $a_1^\pm$ and $a_2^\pm$ and the eigenvalues λ are given by

$$a_1^+ = \sqrt[4]{[\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} \pm 1]/[\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} \mp 1]},$$
$$a_2^- = \mp\sqrt[4]{[\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} \mp 1]/[\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} \pm 1]}, \tag{7}$$

$$a_1^- = (1 - g_H + ig_E)a_1^+, \quad a_2^+ = (1 + g_H + ig_E)a_2^-,$$

$$\frac{\lambda}{k_0} = \frac{1}{2}(g_H^2 + g_E^2) \pm \sqrt{g_\Phi^2 + (g_\omega - g_H)^2},$$

where g_ω is the ratio of the photon energy $\hbar\omega c$ to twice particle self-energy. The components of the velocity $v/c = i\langle\psi^*|\gamma_\alpha|\psi\rangle/\langle\psi^*|\gamma_t|\psi\rangle$ are

$$\begin{aligned}
\frac{v_e}{c} &= \pm\frac{1}{2}\frac{g_E}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta}, \\
\frac{v_h}{c} &= \mp\frac{1}{2}\frac{g_H}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta}, \\
\frac{v_z}{c} &= \frac{\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2}}{1 \pm g_H + \Delta} - 1,
\end{aligned} \tag{8}$$

the components of the spin $i\langle\psi^*|\gamma_{\alpha\beta}|\psi\rangle/\langle\psi^*|\gamma_t|\psi\rangle$ are

$$\begin{aligned}
s_e &= \pm\frac{g_H g_E}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta}, \\
s_h &= \pm\frac{1 - (1/2)(g_H^2 - g_E^2)}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta}, \\
s_z &= \frac{\pm 1 + g_H\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2}}{1 \pm g_H + \Delta},
\end{aligned} \tag{9}$$

and the components $d_\alpha = \langle\psi^*|\gamma_{\alpha t}|\psi\rangle/\langle\psi^*|\gamma_t|\psi\rangle$ are

$$\begin{aligned}
d_e &= \pm\frac{1}{2}\frac{g_H^2 - g_E^2}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta}, \\
d_h &= \pm\frac{g_H g_E}{g_\omega - g_H}\ \frac{g_\Phi}{1 \pm g_H + \Delta} = s_e, \\
d_z &= \frac{g_H\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} \pm (1/2)(g_H^2 + g_E^2)}{1 \pm g_H + \Delta},
\end{aligned} \tag{10}$$

where $\Delta = (1 + (1/2)g_H^2 + (1/2)g_E^2)\sqrt{1 + [g_\Phi/(g_\omega - g_H)]^2} - 1$, and the labels e and h indicate that the respective components are directed along the electric and magnetic fields in the wave. The electric polarization is equal to the electric dipole moment $\mu\mathbf{d}$, and the magnetic polarization to the magnetic dipole moment $\mu\mathbf{s}$. When the wave's frequency ν tends to the resonant frequency $\nu_0 = 2\mu H/h$, the electric polarization of the state tends (according to (10)) to

$$P_e = \pm\frac{\mu^3(H^2 - E^2)}{2(m_0c^2)^2}, \quad P_h = \pm\frac{\mu^3 HE}{(m_0c^2)^2}, \quad P_z = \frac{\mu^2 H}{m_0 c}, \tag{11}$$

and the magnetic polarization tends (according to (9)) to

$$\begin{aligned}
M_e &= \pm\frac{\mu^3 HE}{(m_0c^2)^2}, \quad M_z = \frac{\mu^2 H}{m_0 c^2}, \\
M_h &= \pm\mu\left[1 - \frac{1}{2}\frac{\mu^2(H^2 - E^2)}{(m_0c^2)^2}\right].
\end{aligned} \tag{12}$$

The smaller the value of $|\nu_0 - \nu|$ compared to $|\mu\Phi\omega/h|$ (the extent of broadening of the resonance by the wave's field), the more accurate the description of the electric and magnetic polarizations of the state near resonance by Eqs. (11) and (12).

Hence in ordinary magnetic fields, for which $|\mu H/m_0c^2|$ and $|\mu E/m_0c^2|$ are sure to be less than unity, the anomalous magnetic moments of uncharged Dirac particles (e.g., neutrons) pointing in the same direction as the constant magnetic field are reoriented by a resonant circularly polarized wave in the direction of the magnetic field of the wave, while those pointing in the opposite direction are reoriented in the opposite direction, provided that the broadening of the resonance by the wave's field is much greater than the broadening caused by the thermal motion of the particles and the nonuniformity of the constant magnetic field. In the general case, the transverse electric polarizability α_e and the transverse magnetic polarizability α_h, equal to the ratio of the electric and magnetic transverse polarizations to the wave's field strength, are given by

$$\begin{aligned}
\alpha_e &= \pm\frac{\mu^2}{h(\nu - \nu_0)}\ \frac{(g_H - ig_E)^2}{1 \pm g_H + \Delta}, \\
\alpha_h &= \pm\frac{\mu^2}{h(\nu - \nu_0)}\ \frac{2 - (g_H - ig_E)^2}{1 \pm g_H + \Delta}.
\end{aligned} \tag{13}$$

Here the real and imaginary parts of α_e determine the components of the electric polarizability along the electric field in the wave and perpendicular to that field, respectively, and the real and imaginary parts of α_h determine the components of the magnetic polarizability along the magnetic field in the wave and perpendicular to that field, respectively. The imaginary parts of α_e and α_h are equal in magnitude but have opposite signs. Hence the energy of a state in the field of the wave does not change, and the refractive index of a medium consisting of uncharged Dirac particles with anomalous magnetic moments is a real quantity. If the medium is tenuous, its refractive index is given by

$$n = 1 + 2\pi N(\alpha_e + \alpha_h) = 1 \pm \frac{2\mu^2 N}{\hbar(\nu - \nu_0)}\ \frac{1}{1 \pm g_H + \Delta}, \tag{14}$$

where N is the number of particles per unit volume. In the limit of small g_H and g_E, Eq. (14) yields a refractive index equal to that of a medium consisting of classical particles with magnetic moment μ and mechanical angular momentum $\hbar/2$.

[1] P. J. Redmond, Math. Phys. **6**, 1163 (1965).
[2] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Springer, Berlin (1958).
[3] A. Sommerfeld, *Wave Mechanics*, 6th ed., Methuen, London (1944).

Translated by Eugene Yankovsky

Emergence of a stabilization regime in quantum systems subject to a strong laser field and the Kramers–Henneberger approximation

E. A. Volkova, A. M. Popov, O. V. Smirnova, and O. V. Tikhonova

D. V. Skobel'tsin Scientific Research Institute for Nuclear Physics, M. V. Lomonosov State University at Moscow, 119899 Moscow, Russia

(Submitted 29 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1194–1206 (April 1997)

We use the Kramers–Henneberger approximation to investigate the phenomenon of stabilization of quantum systems with short-range potentials in a strong electromagnetic field. We identify the physical limits imposed on our investigation by the use of this approximation, and the ranges of parameter values of the system and external field that lead to stabilization. We analyze the dependence of the stabilization threshold on the frequency of the laser light. The conclusions obtained from our analytic investigation are confirmed by numerical calculations.
 [S1063-7761(97)00504-0]

1. INTRODUCTION

By studying the effect of intense laser light on the dynamics of quantum systems, we can observe a number of new properties associated specifically with strong fields. Among these are above-threshold ionization and the phenomenon of stabilization of atoms and molecules in a strong field.[1–3] By stabilization we mean a decrease in the probability of ionization of the system W_i as the intensity P of the incident laser light increases.[1–4] This can be explained qualitatively by the fact that for large values of the electromagnetic field intensity the oscillation amplitude of an electron driven by this field significantly exceeds the characteristic size of the original atomic system. Therefore, the electron may be treated as practically free, in which case absorption of energy from the field is impossible (as is emission), and hence the ionization probability decreases.[5,6] Figure 1 illustrates the paradox inherent in the concept of stabilization: it is the weak, nearly-free-electronlike, coupling between the atomic subsystem and the laser field at the maximum of the laser pulse that in the final analysis leads to the system's resistance to ionization. One manifestation of this is a significant recovery of the population of the original atomic state towards the end of the laser pulse. The data shown in Fig. 2 are an example of stabilization of the system near the maximum of the laser pulse. In terms of the electron density at the maximum of the laser pulse (as opposed to the leading edge), the ionization is manifested in the detachment of a packet of the electron density $|\psi(x)|^2$ which ''leaks to infinity.'' Figure 2a shows that there is no ''breakaway'' packet in the distribution $|\psi(x)|^2$ at the maximum of the laser pulse (as opposed to the leading edge). Two symmetric packets appear at a considerable distance from the central dichotomous maximum because of ionization at the leading edge of the pulse. The behavior of $|\psi(x)|^2$ at the end of the laser pulse (Fig. 2b) indicates a second time interval in which ionization occurs, brought about by the trailing edge of the pulse.

It is noteworthy that analytic investigation of the dynamics of the ''atom+field'' system using a basis of initial atomic states is hindered in the case of strong fields due to the inapplicability of perturbation theory. By converting to the Kramers system,[7] we can use a basis of new states that describe the unified ''atom+strong electromagnetic field'' system; these states exist during the time the laser pulse acts.

In this new basis of Kramers-Henneberger states, the dynamics of the system are described by the equation[8–10]

$$i\hbar\,\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\,\frac{\partial^2\psi}{\partial x^2}+V(x+a_e\cos\,\omega t), \tag{1}$$

where

$$a_e=\frac{e\varepsilon_0}{m\omega^2} \tag{2}$$

is the amplitude of the electron's oscillatory motion caused by a wave field with amplitude ε_0 and frequency ω and V is the initial atomic potential. By expanding the time dependence of the potential in a Fourier series and separating the zeroth harmonic V^{KH}, the so-called Kramers–Henneberger potential, we obtain another way to write Eq (1):

$$i\hbar\,\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m}\,\frac{\partial^2\psi}{\partial x^2}+V^{\mathrm{KH}}(x,a_e)+\delta V, \tag{3}$$

where

$$\delta V=\sum_{n\neq 0}V_n(x,a_e)e^{in\omega t}, \tag{4}$$

and $V_n(x,a_e)$ is the nth harmonic of the Fourier expansion of the oscillating potential.

An advantage of this choice of basis is that for strong fields the perturbation δV is small and in this system of coordinates the problem can be solved using perturbation theory (in contrast to the original basis). The emergence of a regime of stabilization when the laser light intensity approaches atomic values is related to the fact that the perturbation δV becomes small, and the real behavior of the system differs negligibly from the solution to the problem with the DC Kramers–Henneberger potential $V^{\mathrm{KH}}(x,a_e)$.[8] In this case, increasing the laser light frequency ω obviously leads to a decrease in the matrix element $\langle V_n e^{in\omega t}\rangle$, and thus the ionization probability. Therefore, stabilization in the

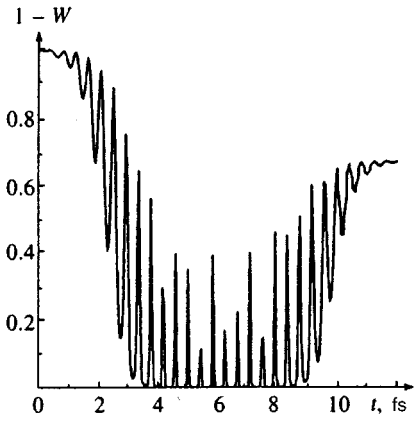

FIG. 1. Probability of remaining in the ground state during a laser pulse for a one-dimensional model of a negative hydrogen atom obtained by the method of direct numerical integration of the time-dependent Schroedinger equation (see Ref. 4): the intensity of the laser light $P = 5 \times 10^{15}$ W/cm^2, the energy of the field quanta $\hbar\omega = 5$ eV. The laser pulse has a Gaussian shape, with a duration of 4 fs based on half-height and a signal maximum that is reached at $t = 6$ fs.

Kramers–Henneberger basis is often described as high-frequency stabilization.[6] Note also that in the strong-field region the dependence of the harmonic V_n on the laser field is estimated[11] to be $V_n \sim (a/a_e)\, V_0$, where a is the half-width of the initial atomic potential and V_0 is its depth. Thus, at sufficiently large laser light intensities the condition $a_e \gg a$ is satisfied and the harmonics become small; however, more and more terms in the sum (4) are found to be close in order of magnitude, and consequently must be included in solving the problem.[12] The smallness of the harmonics for $a_e \gg a$ is in good agreement with the traditional representation of the stabilization conditions, in that this is the case where the amplitude of oscillations of the electron is large, allowing us to assume that it is practically free and to neglect the effect of the original atomic potential. However, in the low-frequency range the condition $a_e \gg a$ is easily satisfied for sufficiently small values of the laser light intensity. Does this imply that we might observe stabilization at considerably smaller intensities in the low-frequency regime? What is the criterion for emergence of a stabilization regime? If we define the stabilization threshold P^* as the value of laser intensity above which we observe a decrease in W_i with increasing P, and use $a_e \sim a$ as our threshold condition, we find that $P^* \sim \omega^4$. However, if we define the stabilization threshold from the condition $ka_e \sim 1$ as in Refs. 3, 6, we find that $P^* \sim \omega^3$. Both these asymptotic expressions lead us to conclude that low frequency stabilization should be observable at very low intensities. However, these conclusions are found to contradict certain physical notions. Although treatment of the system dynamics using Kramers–Henneberger basis states involves a mathematically correct transformation from one system of coordinates to another, there must be physical reasons for the impossibility of using this method in the low-frequency regime.

Pont and Gaurila[6] calculated the lifetime of an atom as a function of light intensity for various values of the laser frequency. And, in fact, they did observe a rapid decrease in the threshold intensity for stabilization with decreasing optical frequency, although they did not analyze the specific form of the dependence. It is also significant that the stabilization threshold is defined only for high optical frequencies, $\hbar\omega \geqslant 27.2$ eV.

Volkova and Popov[13] the authors also established the existence of frequency-dependent stabilization for ionization of a system with short-range potentials in the range $\hbar\omega = 1$–5 eV. They showed that both the dependences mentioned above qualitatively describe the results of numerical experiments. As no one has investigated the phenomenon of stabilization in the range of still lower frequencies $\hbar\omega \leqslant 1$ eV, the questions formulated above have not yet been answered either.

The main advantage of the Kramers–Henneberger scheme is that it allows the problem of interaction of an atom with a strong field to be treated perturbatively. On the other hand, it is the use of perturbation theory rather than the transformation to the Kramers system that limits the region of admissible system and laser parameters. Therefore, it is natural to define the physical limits of applicability of the Kramers–Henneberger method as the range of parameters for which perturbation theory is valid, i.e., the kth-order ionization probability per unit time calculated by perturbation theory (in the basis of Kramers–Henneberger states) exceeds the $(k+1)$th-order ionization probability. In, particular the following condition should hold:

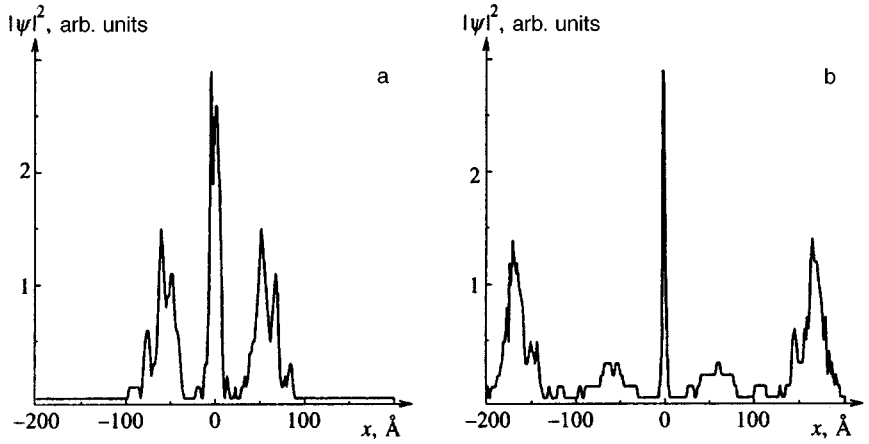

FIG. 2. Spatial distribution of the probability density during laser excitation for the same system as in Fig. 1, at two times: a—at the maximum of the laser pulse, and b—as the pulse decays.

$$\frac{\dot{W}_n^{\mathrm{II}}}{\dot{W}_n^{\mathrm{I}}} \ll 1 \tag{5}$$

for any value of n, where $\dot{W}_n^{\mathrm{I}}$, $\dot{W}_n^{\mathrm{II}}$ are the ionization rates calculated in first and second order perturbation theory corresponding to a transition from the ground state to the continuum in which the energy of the electron changes by $n\hbar\omega$. It is also necessary to require that the dipole approximation be applicable, since the Kramers–Henneberger formalism was developed within its framework.

In this paper, we will employ the Kramers–Henneberger method, determining its applicability from condition (5), to identify ranges of parameter values that lead to stabilization of the system. We will determine conditions for the appearance of a stabilization regime and the frequency dependence of the stabilization threshold, with the low-frequency case discussed as a special case. We will also discuss the question of how stabilization in the Kramers–Henneberger regime relates to interference stabilization.[14] Our theoretical conclusions are confirmed by the results of numerical experiments.

2. NUMERICAL CALCULATIONS OF THE DYNAMICS OF THE SYSTEM

We first numerically integrated the time-dependent Schroedinger equation directly in the two different bases, i.e., the basis of states of the original atomic potential and the basis of Kramers–Henneberger states. The first problem takes the form

$$i\hbar\,\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m}\,\frac{\partial^2\psi(x,t)}{\partial x^2} + [V(x) - ex\varepsilon(t)\cos\,\omega t]\psi(x,t),$$
$$\psi(x,t=0) = \varphi_1(x). \tag{6}$$

Here

$$\varepsilon(t) = \varepsilon_0\,\exp\left(-\frac{(t-t_0)^2}{2\tau^2}\right)$$

is the envelope of the laser pulse, τ is its half-width, and the time t_0 corresponds to the maximum of the pulse; $\varphi_1(x)$ is the wave function of the ground state of the atomic potential. In the calculations we set $t_0 = 3\tau$. For $V(x)$ we chose a rectangular well with various values of the depth V_0 and width $2a$, and also the smoother short-range potential taken from Ref. 4 and corresponding to a one-dimensional model of a negative hydrogen ion.

In the Kramers system of coordinates the problem takes the form

$$\begin{cases} i\hbar\,\dfrac{\partial\psi(x,t)}{\partial t} = -\dfrac{\hbar^2}{2m}\,\dfrac{\partial^2\psi(x,t)}{\partial x^2} + V(x+a_e\,\cos\,\omega t)\,\psi(x,t), \\ \psi(x,t=0) = \varphi_{\mathrm{KH}}^{(1)}(x). \end{cases} \tag{7}$$

The difference between this problem and Eq. (6) lies in the representation of both the atomic potential and the field as a single potential oscillating in the Kramers coordinate system, and in the choice of the lower Kramers–Henneberger state with 100% occupancy as an initial condition. Both of these problems allow us to compute the wave function of the system $\psi(x,t)$, which can be used to obtain the population $|C_k(t)|^2$ of the various states $\varphi_k(x)$ in the basis of atomic states and $\varphi_{\mathrm{KH}}^{(k)}$ in the basis of Kramers–Henneberger states, and also the ionization probability W_i of the system:

$$C_k(t) = \int \psi(x,t)\,\varphi_k(x)\,dx, \tag{8}$$

or

$$C_k(t) = \int \psi(x,t)\,\varphi_{\mathrm{KH}}^{(k)}(x)\,dx, \tag{8a}$$

$$W_i = 1 - \sum_k |C_k(t)|^2. \tag{9}$$

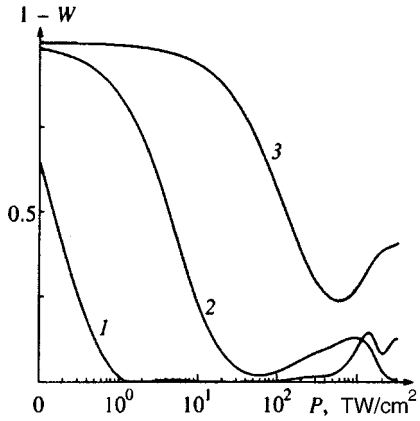

FIG. 3. Dependence of the quantity $1 - W_i$ on the intensity of a laser pulse with duration $2\tau = 20$ fs for a one-dimensional H^- ion and various values of the quantum energy: $\hbar\omega = 1.2$ (*1*), 2.5 (*2*), and 5 eV (*3*).

Figure 3 shows the quantity $1 - W_i$ calculated in this way for the case of the potential from Ref. 4 as a function of the intensity of the laser pulse for various values of the laser frequency. The data indicate more intense ionization at lower frequencies. The portions of the curves with positive derivative correspond to an ionization probability that decreases with increasing intensity of the laser emission, i.e., stabilization. These data allow us to conclude that the threshold for stabilization decreases as we enter the low-frequency range. However, even when the photon energy satisfies $\hbar\omega = 1.2$ eV, the probability of ionization is practically 100% over a rather wide interval of intensity values. In this case the curve develops a horizontal portion (a ''shelf''), which leads to ambiguity in determining the stabilization threshold. If we nevertheless define the threshold for stabilization as before, then the rate of decrease of this quantity as a function of frequency will be less in the low-frequency range.

The threshold for stabilization can also be defined by solving the problem (7), where it corresponds to that value of field intensity at which the lifetime of the bound Kramers–Henneberger state is a minimum given that the system is in this state initially. Both these approaches give very similar dependences of the stabilization threshold P^* on the field frequency (see Fig. 4); these results are practically independent of the shape of the initial atomic potential when its

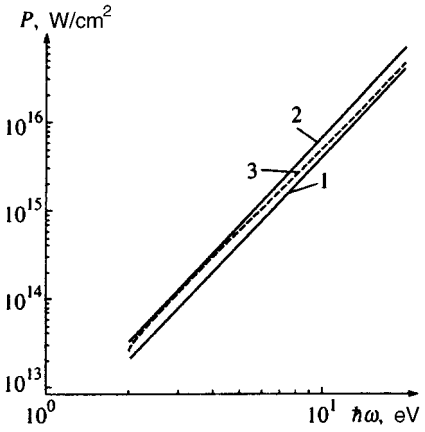

FIG. 4. Dependence of the stabilization threshold on photon energy obtained by finding the minimum lifetime of the Kramers–Henneberger ground state (1) and the lifetime calculated by finding the minimum of the quantity $1 - W_i$ (2). Curves *1*, *2* are for a rectangular potential well with depth $V_0 = 3.33$ eV and width $2a = 3$ Å, while *3* is for the potential used to model the one-dimensional negative hydrogen ion (see Ref. 4).

depth and width are fixed. However, approximation of the data obtained here by a power-law function $P^*(\omega)$ gives values of the exponent in the range from 3 to 4, which differ very little. Thus, for curves *1*, *2*, *3* the exponents turn out to equal 3.31, 3.57, and 3.23 respectively. Calculations for a rectangular well of depth 12 eV and width 3.33 Å show an abrupt change in the slope of the linear function $P^*(\hbar\omega)$ when plotted on a log-log scale as we enter the region of low frequencies (Fig. 5.) These data indicate that the function $P^*(\hbar\omega)$ can be approximated by a power law with an exponent equal to 3.7 for $\hbar\omega \geqslant 10$ eV and 2.0 for $\hbar\omega < 10$ eV. The abrupt change in slope of the curve indicates a relative increase in the ionization threshold as we enter the low-frequency range. This change in the frequency dependence indicates that at low frequencies stabilization takes place at relatively high laser intensities. Analysis of the dependence of the stabilization threshold on the depth of the original rectangular potential for fixed photon energy (Fig. 6) leads to an analogous conclusion. As long as the field frequency remains relatively high compared to the ionization potential of the original state, the quantity P^* is a weak function of the well depth V_0. However, for large values of V_0 the photon

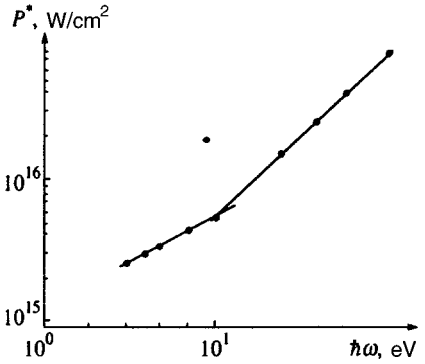

FIG. 5. Dependence of the stabilization threshold on photon energy for a rectangular potential well with depth $V_0 = 12$ eV and width $2a = 3.33$ Å.

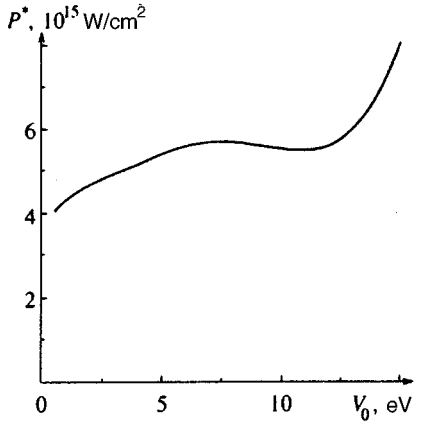

FIG. 6. Dependence of the stabilization threshold on the depth of a rectangular potential well V_0 for fixed values of the well width $2a = 3.33$ Å and field quantum energy $\hbar\omega = 10$ eV.

energy $\hbar\omega$ becomes comparable to or smaller than the binding energy, so that this frequency becomes ''small'' for a given value of V_0. Hence the stabilization threshold increases sharply.

Let us now try to estimate the stabilization threshold $P^*(\omega)$ from analytical considerations.

3. ESTIMATING THE STABILIZATION THRESHOLD

The emergence of a regime of stabilization is connected with low probability of ionization per unit time $\dot{W}_{\mathrm{KH}}$ from the lower state. When we approximate the continuum state by a plane wave, we can write the following:

$$\dot{W}_{\mathrm{KH}} \propto \left| \int \varphi_{\mathrm{KH}}^{(1)} V_n e^{ikx} dx \right|^2. \tag{10}$$

This quantity will necessarily be small if many periods of the plane wave are contained within the characteristic width of the Kramers–Henneberger potential. Since the characteristic width of the Kramers–Henneberger potential is $2a_e$, where a_e is the amplitude of oscillation of an electron acted on by the electromagnetic wave, the threshold condition has the form

$$ka_e \sim 1. \tag{11}$$

Relating the wave vector k with the energy of an electron in the continuum

$$\frac{k^2\hbar^2}{2m} = \hbar\omega - I^{\mathrm{KH}}, \tag{12}$$

we obtain

$$P^* = \omega^3 \frac{mc}{16\pi e^2} \frac{1}{1 - I^{\mathrm{KH}}(a_e)/\hbar\omega}, \tag{13}$$

where $I^{\mathrm{KH}}(a_e)$ is the ionization potential of the state under discussion which depends on the amplitude of oscillation of the electron a_e, which is given in terms of the laser intensity at the threshold P^* by the following expression:

$$a_e^2 = \frac{8\pi e^2}{cm^2\omega^4} P^*. \tag{14}$$

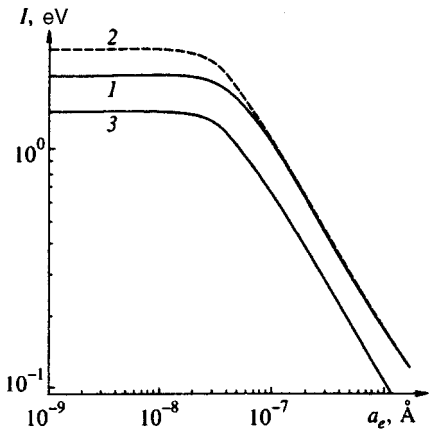

FIG. 7. Dependence of the ionization potential of the Kramers–Henneberger ground state on the oscillation amplitude of an electron for various depth parameters V_0 and width parameters $2a$ of the initial atomic potential: *1*—$V_0 = 3.33$ eV, $2a = 3$ Å, *2*—$V_0 = 5$ eV, $2a = 2$ Å, *3*—$V_0 = 3.33$ eV, $2a = 2$ Å.

Thus, in order to solve Eq. (13) it is necessary to know the form of the function $I^{\mathrm{KH}}(a_e)$.

The required dependence was investigated via numerical calculations for rectangular potentials with various parameters (Fig. 7). It turned out that in the range of values of a_e larger than the half-width of the atomic potential, the quantity I^{KH} could be approximated with good accuracy by a power-law function of a_e:

$$I^{\mathrm{KH}} \propto a_e^{-0.75}. \tag{15}$$

Using the function $I^{\mathrm{KH}}(a_e)$ shown in Fig. 7, we solved Eq. (13) and obtained the function $P^*(\omega)$. The data obtained for $V_0 = 3.33$ eV and $2a = 3$ Å are plotted in Fig. 8 on a log–log scale. The presence on the graph of two straight-line segments with markedly differing slopes ($P^* \propto \omega^{2.8}$, in the high-frequency range and $P^* \propto \omega^{1.5}$ in the low-frequency range) is in good agreement with the results of numerical integration of the Schroedinger equation, and indicates a relative increase in the stabilization threshold at low frequencies. In this case, the condition $\hbar\omega > I_0 \approx I^{\mathrm{KH}}$ is satisfied in the range of frequencies for which $P^* \propto \omega^{2.8}$ (here I_0 is the ionization potential of the initial atom). Further decreasing the frequency ω leads to multiphoton ionization out of the atomic potential; however, in the Kramers–Henneberger representation the ionization remains single-photon up to values $\hbar\omega = 0.1$ eV. Thus, at the stabilization threshold the condition

$$\hbar\omega \geqslant I^{\mathrm{KH}},$$

is always satisfied; this corresponds to the Gavrila–Kaminski condition[8] for applicability of the Kramers–Henneberger approximation. Note, however, that this condition does not characterize the stabilization effect as a high-frequency phenomenon, since it turns out to hold at very small frequencies and does not constitute a frequency bound on the applicability of this approximation from below.

The question of when the Kramers–Henneberger approximation can be used will be discussed in more detail in the next section.

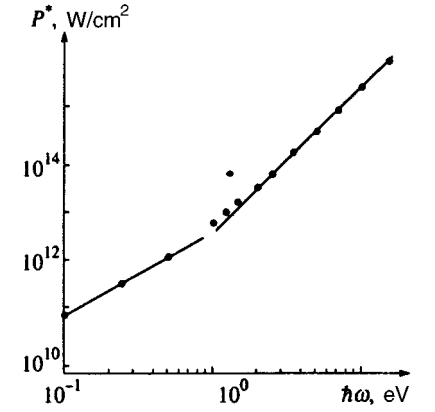

FIG. 8. Stabilization threshold vs. energy of the field quanta plotted using Eq. (13), taking into account the dependence of the ionization potential for the Kramers–Henneberger ground state on the amplitude of oscillation of the electron.

4. ANALYSIS OF WHEN THE KRAMERS–HENNEBERGER METHOD CAN BE USED AND THE EXISTENCE OF A STABILIZATION EFFECT

In light of the problems mentioned above, which arise when the process of ionization of a quantum system is analyzed in the low-frequency limit, we must address the problem of delineating the region of applicability of the Kramers–Henneberger treatment, starting from Eq. (5), and also of investigating the resulting region with regard to dependence of the ionization probability on the laser intensity. Our investigation allowed us to establish ranges of values of P and ω within which stabilization is observed.

Condition (5) implies that for ionization from the Kramers–Henneberger ground state, in which the electron energy changes by $n\hbar\omega$, the probability per unit time $\dot{W}_n^{\mathrm{II}}$ for ionization in second-order perturbation theory is much smaller than the first-order quantity $\dot{W}_n^{\mathrm{I}}$ for any value of n.

The ionization probability per unit time from the Kramers–Henneberger ground state $\psi_{\mathrm{KH}}^{(1)}(x)$ given by first-order perturbation theory has the form

$$\dot{W}_n^{\mathrm{I}} = \frac{2\pi}{\hbar^2} |V_{0k}^{(n)}|^2 \frac{m}{\hbar k_n}. \tag{16}$$

Here

$$V_{0k}^{(n)} = \int \varphi_{\mathrm{KH}}^{(1)}(x) V_n(x, a_e) \exp(ik_n x)\, dx \tag{17}$$

is the matrix element that connects the initial state with the continuum state characterized by a wave vector k_n, where k_n is defined by the condition

$$\frac{\hbar^2 k_n^2}{2m} = n\hbar\omega - |I^{\mathrm{KH}}| \approx n\hbar\omega. \tag{18}$$

Calculating the matrix element (17), we obtain[12]

$$\dot{W}_n^{\mathrm{I}} \sim \pi B \frac{V_0}{\hbar} \frac{J_n^2(k_n a_e)\mathrm{sinc}^2(k_n a)}{n^2 k_n a_e (1 + a/a_e)}, \tag{19}$$

where $B = 2mV_0a^2/\hbar^2$ is the Born parameter, and V_0 and a are the depth and half-width of the original atomic potential.

Analogous computations to second order in perturbation theory and use of the pole approximation yields the following expression for $\dot{W}_n^{\mathrm{II}}$:

$$\dot{W}_n^{\mathrm{II}} \sim \pi B \frac{V_0}{\hbar} \sum_j \frac{B^2}{(n-j)^2 j^2} \times \frac{\mathrm{sinc}^2((k_n-k_j)a) J_{n-j}^2((k_n-k_j)a_e)}{(k_j a)^2 k_n a_e (1+a/a_e)} \times J_j^2(k_j a_e)\mathrm{sinc}^2(k_j a). \tag{20}$$

Here the summation ranges over intermediate states with label j.

Using Eqs. (19) and (20), we can rewrite condition (5) in the form

$$\frac{\dot{W}_n^{\mathrm{II}}}{\dot{W}_n^{\mathrm{I}}} \sim \sum_j \frac{n^2 B^2}{(n-j)^2 j^2} \frac{\mathrm{sinc}^2(k_j a)\mathrm{sinc}^2((k_n-k_j)a)}{(k_j a)^2\ \mathrm{sinc}^2(k_n a)} \times \frac{J_{n-j}^2((k_n-k_j)a_e) J_j^2(k_j a_e)}{J_n^2(k_n a_e)} \ll 1. \tag{21}$$

The ranges of values of P and ω that satisfy this relation are bounded, while bounds on the applicability of the Kramers–Henneberger method exist in the sense of Eq. (5).

By searching over all j for the maximum value of $\dot{W}_n^{\mathrm{II}}/\dot{W}_n^{\mathrm{I}}$ and using the asymptotic forms of the Bessel functions for various relations between the argument and index[16] along with known expansions of the function sinc x, we obtain the region of values of P and ω that satisfy (21):

for $ka<1$, i.e., $\omega < \left(\frac{a_{\mathrm{at}}}{a}\right)^2 \omega_{\mathrm{at}}$:

$$\frac{P}{P_{\mathrm{at}}} > \frac{n^5}{(n\mp 1)^5} \left(\frac{V_0}{I_{\mathrm{at}}}\right)^4 \left(\frac{a}{a_{\mathrm{at}}}\right)^4 \left(\frac{\omega}{\omega_{\mathrm{at}}}\right), \tag{22}$$

for $ka>1$, i.e., $\omega > \left(\frac{a_{\mathrm{at}}}{a}\right)^2 \omega_{\mathrm{at}}$:

$$\frac{P}{P_{\mathrm{at}}} > \frac{n^7}{(n\mp 1)^7} \left(\frac{V_0}{I_{\mathrm{at}}}\right)^4 \left(\frac{\omega_{\mathrm{at}}}{\omega}\right). \tag{23}$$

The ''+'' sign in the denominator corresponds to $n=1$; P_{at}, I_{at}, a_{at}, and ω_{at} are values of the atomic intensity, energy, length, and frequency. Since Eqs. (22) and (23) should be valid for any n, it is sufficient to require that they holed for those n that maximize the expressions on the right sides of (22) and (23). It is easy to verify that this occurs for $n = 2$. The resulting bounds for the region of applicability of the Kramers–Henneberger approach (for $n = 2$) are plotted on a log–log scale in Fig. 9 (straight lines *1*, *2*). We add to these the condition for applicability of the dipole approximation $v \ll c$, giving the expression

$$\frac{P}{P_{\mathrm{at}}} < (137)^2 \left(\frac{\omega}{\omega_{\mathrm{at}}}\right)^2 \tag{24a}$$

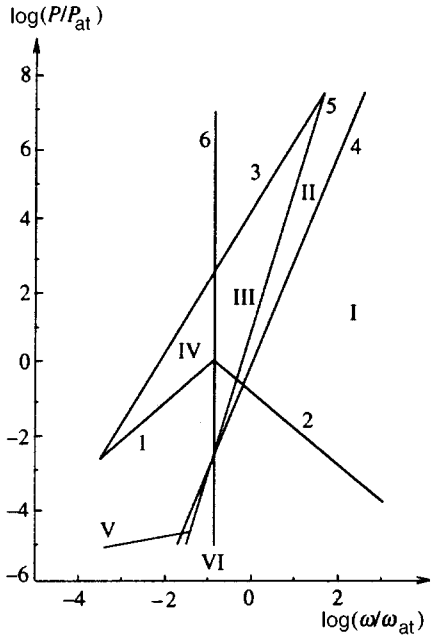

FIG. 9. Bounds on the region of applicability of the Kramers–Henneberger method and subregions with different frequency dependences $\dot{W}_i(P)$, the straight lines *1*, *2* are bounds on the applicability of the Kramers–Henneberger method obtained from the condition $\dot{W}_n^{\mathrm{II}}/\dot{W}_n^{\mathrm{I}} = 1$; *3*—the condition $e\varepsilon_0/m\omega = c$; *4*—the condition $ka_e = 1$; *5*—the condition $a = a_e$; *6*—the condition $ka = 1$. The parameters of the potential were $V_0 = 5$ eV, $2a = 2$ Å.

The straight line *3* shown in Fig. 9 representing this expression constitutes an additional restriction on the admissible region. Use of more rigorous criteria for applicability of the nonrelativistic approximation, based on comparing the relativistic drift velocity of an electron due to the Lorentz force with the rate of spreading of the electron wave packet,[17] leads to the relation

$$\frac{P}{P_{\mathrm{at}}} < (23)^2 \left(\frac{\omega}{\omega_{\mathrm{at}}}\right)^2. \tag{24b}$$

Note that this condition does not qualitatively change the picture under discussion; however, it is a stronger constraint on the range of applicability of the Kramers–Henneberger approximation at low frequencies than condition (24a). The resulting bounds on the applicability of this approximation define the ranges of values of P and ω we are looking for; in particular, they bound the range of frequencies where it can be used from below. It is also worth noting that our analysis is correct when (18) holds, which in the limit is equivalent to the conditions $\hbar\omega > I_0$ or $ka > 1$. This condition ''cuts off'' a portion of the admissible region, whose dimensions depend on the parameters of the original atomic potential.

Using the expression for the total ionization probability per unit time

$$\dot{W}_{\mathrm{tot}}^{\mathrm{I}} \cong \sum \frac{2\pi}{\hbar} B V_0 \frac{J_n^2(k_n a_e)\mathrm{sinc}^2(k_n a)}{n^2 k_n a_e (1+a/a_e)}, \tag{25}$$

let us investigate the dependence of $\dot{W}_{\mathrm{tot}}^{\mathrm{I}}$ on the intensity of the laser pulse. In this case the required region is divided up into subregions by the straight lines:

$$4-ka_e=1 \quad \text{or} \quad \frac{P}{P_{\text{at}}}=\left(\frac{\omega}{\omega_{\text{at}}}\right)^3,$$

$$5-a_e=a \quad \text{or} \quad \frac{P}{P_{\text{at}}}=\left(\frac{a}{a_{\text{at}}}\right)^2\left(\frac{\omega}{\omega_{\text{at}}}\right)^4,$$

$$6-ka=1 \quad \text{or} \quad \frac{\omega}{\omega_{\text{at}}}=\left(\frac{a}{a_{\text{at}}}\right)^2$$

$$=\text{const for a given potential.}$$

Here $k=\sqrt{2m\hbar\omega/\hbar^2}$.

The parameters of the potential V_0 and a, along with the straight lines *4*, *5*, and *6*, lead to a division of the initial region into as many as six subregions. In each subregion the quantity $\dot{W}^{\text{I}}_{\text{tot}}$ depends on P in a different way. By examining the functions $\dot{W}^{\text{I}}_{\text{tot}}(P)$ in each of the subregions, we can trace the values of parameters P and ω that lead to a stabilization regime (see Fig. 9):

$$\text{I}-\dot{W}^{\text{I}}_{\text{tot}}\propto P, \quad \text{IV}-\dot{W}^{\text{I}}_{\text{tot}}\propto\frac{1}{P},$$

$$\text{II}-\dot{W}^{\text{I}}_{\text{tot}}\propto\frac{1}{\sqrt{P}}, \quad \text{V}-\dot{W}^{\text{I}}_{\text{tot}}\propto\sqrt{P},$$

$$\text{III}-\dot{W}^{\text{I}}_{\text{tot}}\propto\frac{1}{\sqrt{P}}, \quad \text{VI}-\dot{W}^{\text{I}}_{\text{tot}}\propto P.$$

Thus, stabilization occurs in subregions II, III, and IV, since in these regions increasing P leads to a decrease in the ionization probability. However, the relative positions of the straight lines $ka_e = 1$ and $a_e = a$ may change depending on the parameter a of the atomic potential. This implies that stabilization can begin as soon as the condition $ka_e = 1$ is fulfilled. However, in this case $a_e < a$ is possible, i.e., the double-well Kramers–Henneberger potential has not yet formed. Thus, the frequency dependence of the stabilization threshold can be either $P^* \propto \omega^3$ or $P^* \propto \omega^4$, which is also observed in the numerical calculations. At low frequencies the function $P^*(\omega)$ becomes linear, which agrees qualitatively with the data shown in Figs. 5 and 8. The fact that the stabilization is possible even before the two-well Kramers–Henneberger potential forms is extremely significant in the context of comparing mechanisms for interference stabilization[14] with this kind of stabilization. It is well known that interference stabilization arises because of destructive interference between probability amplitudes for transitions from two closely spaced states in the discrete spectrum of the atomic potential to the continuum. A typical example of a system in which stabilization of this kind is possible is a Rydberg atom. However, the Kramers–Henneberger potential can also have closely-spaced levels with overlapping widths when the condition $a_e \gg a$ holds. Therefore, we might expect the mechanism of interference stabilization to operate in our system, and that the Kramers–Henneberger formalism simply gives an alternate method for describing it. This point of view was advanced by the authors of Ref. 15. However, in a number of cases our calculations show that stabilization is achieved even when $a_e < a$ holds, when the Kramers–Henneberger potential is still a single well with a single isolated bound state. This indicates that the mechanisms of interference stabilization and stabilization in the Kramers–Henneberger regime are different. In the limit $a_e \gg a$, interference stabilization can contribute and compete with Kramers–Henneberger stabilization. However, it is also noteworthy that there is a range of parameters for which stabilization does not arise in general even though $a_e \gg a$.

5. CONCLUSION

Thus, our analysis of the conditions for stabilization lead us to the following conclusions.

1. Bounds on the applicability of the Kramers–Henneberger method can be identified which show that the Kramers–Henneberger approach cannot be used at low-frequencies.

2. In the range of applicability of the Kramers–Henneberger method, stabilization can occur only for certain subsets of the system parameters, i.e., the atomic-potential parameters and field frequency, and the dependence of the stabilization threshold on frequency is different for different ranges of these parameters.

3. Since formation of a two-well Kramers–Henneberger potential is not a necessary condition for stabilization, the mechanism of stabilization in this regime does not uniquely reduce to the interference mechanism.

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Grant No. 95-02-06258).

[1] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms* (Springer-Verlag, 1993); N. B. Delone and V. P. Krainov, Usp. Fiz. Nauk **165**, 1295 (1995).
[2] K. Burnett, V. C. Reed, and P. L. Knight, J. Phys. B **26**, 561 (1993).
[3] M. V. Fedorov, *The Electron in a Strong Optical Field* (Nauka, Moscow, 1991).
[4] A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **5**, 1184 (1995).
[5] M. Pont, N. R. Walet, M. Gavrila, and C. W. McCurdy, Phys. Rev. Lett. **61**, 939 (1988).
[6] M. Pont and M. Gavrila, Phys. Rev. Lett. **65**, 2362 (1990).
[7] H. A. Kramers, *Collected Papers* (North-Holland, Amsterdam, 1956).
[8] M. Gavrila and J. Z. Kaminski, Phys. Rev. Lett. **52**, 613 (1984).
[9] M. J. Offerhaus, M. Gavrila, and J. Z. Kaminski, Phys. Lett. A **112**, 151 (1985).
[10] J. van de Ree, M. Gavrila, and J. Z. Kaminski, Phys. Rev. A **37**, 4536 (1988).
[11] E. A. Volkova, A. M. Popov, and O. V. Smirnova, Zh. Éksp. Teor. Fiz. **109**, 138 (1996) [JETP **82**, 72 (1996)].
[12] E. A. Volkova, A. M. Popov, and O. V. Smirnova, Zh. Éksp. Teor. Fiz. **106**, 1360 (1994) [JETP **79**, 736 (1994)].
[13] E. A. Volkova and A. M. Popov, Zh. Eksp. Teor. Fiz. **105**, 592 (1994) [JETP **78**, 315 (1994)].
[14] M. V. Fedorov and A. M. Movsesian, J. Opt. Soc. Amer. B **5**, 850 (1988); M. V. Fedorov, M. Yu. Ivanov, and A. M. Movsesian, J. Phys. B **23**, 2245 (1990).
[15] J. B. Watson, C. H. Keitel, P. L. Knight, and K. Burnett, Phys. Rev. A **52**, 4025 (1995).
[16] M. Abramowitz and I. Stegun, (Eds.) *Handbook of Mathematical Functions* (U.S. Govt. Printing Office, Washington, DC, 1972).
[17] and M. V. Fedorov, Laser Phys. **3**, 265 (1993).

Translated by Frank J. Crowne

Decay of transient nutation in two-level spin systems

G. G. Fedoruk

A. N. Sevchenko Research Institute of Applied Physics Problems, 220064 Minsk, Republic of Belarus
(Submitted 9 August 1996)
Zh. Éksp. Teor. Fiz. **111**, 1207–1213 (April 1997)

This paper reports the results of an experimental study of the decay of transient NMR nutations in a two-level spin system with homogeneous line broadening. The NMR nutation signals in glycerin were studied for $10 \leqslant \omega_1 T_2 \leqslant 150$, where $\omega_1 = \gamma H_1$, with γ the gyromagnetic ratio and H_1 the amplitude of the magnetic component of the radio-frequency field, and T_2 is the transverse relaxation time. It is found that in a high-power field ($\omega_1 T_2 \gg 1$) the nutation decay rate is independent of ω_1 and is quantitatively described by Bloch's model. The data is compared with the data on non-Bloch (ω_1-dependent) EPR-nutation decay in quartz (R. Boscaino, F. M. Gelardi, and J. P. Corb, Phys. Rev. B **48**, 7077 (1993)). © *1997 American Institute of Physics.* [S1063-7761(97)00604-5]

1. INTRODUCTION

Lately the decay of transient nutations in two-level spin systems has attracted attention in connection with verifying the validity of Bloch equations in the high-power driving-field limit.[1,2] Other phenomena such as free-induction decay, echo, and burning out of dips, which constitute the basis of coherence resonant spectroscopy, have been studied for a similar reason (see Refs. 1–3 and the works cited therein).

Time-dependent oscillations (Rabi oscillations or nutations) constitute the simplest coherent dynamic effect in the spectroscopy of quantum systems used in NMR, EPR, and optical resonance.[1,4,5] This effect can be observed within a time interval that is short compared to the relaxation time and reflects a situation in which a new stationary state is established after the sudden (nonadiabatic) turn-on of a resonant interaction of an intense electromagnetic field and the equilibrium quantum system. The oscillation frequency is the measure of the interaction between the quantum system and the field, while the decay of the oscillations provides information about dephasing and relaxation processes taking place in the presence of an electromagnetic field. These processes are usually described by Bloch's model by introducing phenomenological transverse and longitudinal relaxation times, T_2 and T_1.

The first transient nutations were observed nearly 50 years ago in NMR experiments[6] and 20 years ago in optical-resonance and EPR experiments, and then in multiphoton resonances (see the literature cited in Ref. 1). Nevertheless, up to now the validity of Bloch's model in the high-power driving-field limit of the nutation mode has not been substantiated by experiments. The difficulties lie in registering a transient signal during a pulse of the driving field, and imply a narrow range of admissible field values and limited possibilities of observing undistorted decaying nutations, a process masked by the inhomogeneities in the driving field.

The fact that the observed decay rate for the nutation signal is higher than the value predicted by Bloch's model was usually related to the inhomogeneities in the driving field in the bulk of the sample.[4–7] Sometimes, to overcome this limitation and to study relaxation in the presence of a driving field, more complex modifications of nutation experiments (rotary echo[4,5] and two-pulse delayed nutation[8]), which allow for reversal of the dephasing caused by the inhomogeneities of the driving field, are employed. However, the decay time for these signals exceeds the transverse relaxation time determining the decay of nutations.[5]

Only recently have results been obtained in experiments that quantitatively investigated the properties of decay of transient nutations not distorted by the inhomogeneities of a variable driving field.[1] Such studies, involving two-photon EPR[1] in quartz, for spin systems with inhomogeneous line broadening, revealed a discrepancy between the experimental data and the predictions of Bloch's model: the measured nutation decay rate proved to be higher and linearly dependent on the amplitude of the driving field. The reason for such behavior has yet to be established. The problem is important because the theory of resonant interaction of field and matter is based on Bloch equations. Boscaino *et al.*[1] assumed that the anomalous (non-Bloch) decay is a property of truly homogeneous quantum systems, whose effective spectral linewidth is characterized by a term dependent on the amplitude of the driving field, in addition to being characterized by the internal time T_2. The present paper is an attempt to verify the validity of Bloch's model in the high-power driving field limit by experimentally studying the decay of transient (time-dependent) nutations in a two-level spin system with homogeneous line broadening. The experiments involved studying the NMR for protons in glycerin.

2. THE PREDICTIONS OF BLOCH'S MODEL

The solution of the Bloch equations in Torrey's high-power driving field approximation[4,6] ($\omega_1 \gg 1/T_2,\ 1/T_1$) for the v-component of transient nutation (the absorption signal) has the form

$$v = v_0 \frac{\omega_1}{\Omega} \exp\left[-\left\{ \frac{1}{T_2} - \frac{1}{2}\left(\frac{1}{T_2} - \frac{1}{T_1} \right)\left(\frac{\omega_1}{\Omega} \right)^2 \right\} t \right] \sin \Omega t, \tag{1}$$

where $t > 0$,

$$\Omega = \sqrt{\omega_1^2 + (\omega_0 - \omega)^2}, \tag{2}$$

 1063-7761/97/040665-04$10.00

v_0 is the equilibrium value of the population difference, $\omega_1=\gamma H_1$ is the Rabi frequency, γ is the gyromagnetic ratio, ω and H_1 are the frequency and amplitude of the magnetic component of the radio-frequency (rf) field, and ω_0 is the resonant frequency of the quantum system.

For resonant ($\omega=\omega_0$) excitation of the spin system with $T_1=T_2$, the detected signal must (according to (1)) be described by the following relationship:

$$v=v_0\ \exp\left(-\frac{t}{T_2}\right)\sin\ \omega_1 t. \tag{3}$$

We see that nutation decay is determined solely by the time T_2 and is independent of the strength of the rf field acting on the system.

Because of its simplicity, this case is convenient for checking the predictions of Bloch's model in the high-power driving field limit. On the other hand, by studying this situation experimentally it is also possible to check the hypothesis of Boscaino *et al.*[1] concerning the dependence of nutation decay on the amplitude of the field acting on the system. According to this hypothesis, the decay time τ of nutation of a homogeneous line must depend on the field strength in the following manner:

$$\frac{1}{\tau}=\frac{1}{T_2}+2\beta\omega_1, \tag{4}$$

with β a dimensionless parameter equal to 0.021–0.033 (according to the EPR data of Boscaino *et al.*[1] on parametric centers in quartz), and τ the time of an e-fold decrease in the amplitude of the exponentially decaying nutations. The goal of the experimental studies described below was to establish the conditions sufficient for checking both Bloch's model and the above-mentioned hypothesis.

3. EXPERIMENT

The choice of NMR was determined by the relaxation times, which are longer than the relaxation times in EPR and optical resonance. The experiments were done using an NMR spectrometer with a frequency 14.4 MHz at room temperature. Transient nutations were formed by pulsed variations of the amplitude of the rf field. The modulator ensured the formation of rf pulses with an attenuation in the off state of at least 70 dB. The pulses were fed to a power amplifier with a high-Q circuit that corrected the spectral composition of a pulse with a rise time of about 7 μs. After the amplifier the pulses were fed through an attenuator to an rf bridge. Detection of the signals in the course of an rf pulse required good matching of the rf channel (60–80 dB). After the amplifier the signal of the bridge unbalance caused by the interaction of the rf field and the spin system was sent to a phase-sensitive detector with a time constant of about 3 μs.

The effect of inhomogeneities in the rf field was reduced by reducing the filling factor of the pick-up coil to 0.0025 via a reduction of the sample volume to 3 mm^3. The loss of signal-to-noise ratio caused by the decrease in sample volume was partially offset by coherent digital signal integration. On the whole, a compromise between the homogeneity of the driving field, the duration of the experiment, and the attainable signal-to-noise ratio was chosen. The inhomogeneity of the field H_1 in the sample was monitored by measuring the shape of the two-pulse delayed-nutation signal[8] and by observing the nonexponential decay of nutations.

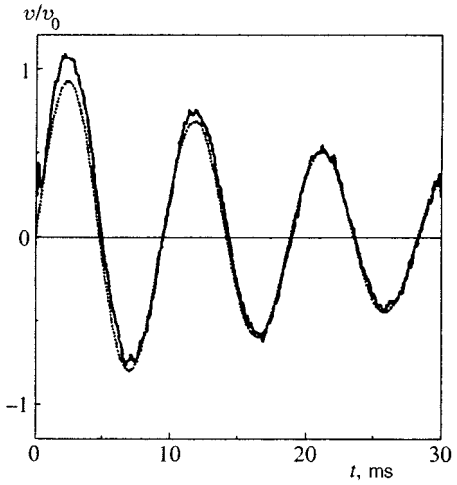

FIG. 1. The transient nutation signal detected at $\omega_1=2\pi\times 107$ Hz and $\omega-\omega_0=0$. The dotted curve was obtained from Eq. (3) with $T_2=32$ ms.

As is known,[4,6] the spatial inhomogeneity of a polarizing magnetic field, ΔH, can transform the nutations into a form described by a zero-order Bessel function, and for this reason it may lead to additional decay of the nutations. In the experiment described here, $\gamma\Delta H<1/T_2$, and inhomogeneous broadening was negligible. This was monitored by the time of decay of the free-induction signal and by the shape of the initial region of nutations.

Protons in glycerin ($T_1\approx T_2$) were selected for the two-level system with homogeneous broadening. The relaxation parameters of this system have been thoroughly studied,[4,9] and the first nutations in this system were observed by Torrey.[6] At room temperature and at the operating frequency ($\omega_0/2\pi=14$ MHz) for protons in glycerin, $T_1=3.2\times 10^{-2}$s (Ref. 9). The uniformity of the external fields achieved in the experiments described below made it possible to detect, in the course of the above time interval, nutation signals that were essentially undistorted by inhomogeneities. Here, at maximum rf power, the nutation frequency amounted to about $2\pi\times 0.8$ kHz and was found to exceed $1/T_2$ by a factor larger than 150, which to a considerable extent corresponds to the high-power driving-field limit ($\omega_1 T_2\gg 1$).

4. RESULTS AND DISCUSSION

The transient nutation signals were detected when the spin system was under resonant excitation ($\omega=\omega_0$). Accurate tuning to resonance was monitored by measuring the frequency of the observed nutations, with the minimum frequency corresponding to resonance. The length of an rf pulse was 30 ms, the pulse repetition period was 500 ms, and the signal was averaged 128 times. A typical oscillogram of the detected signal is depicted in Fig. 1. The rf pulse is turned on at $t=0$. Clearly, Eq. (3) provides a good description of the observed signal at $\omega_1=2\pi\times 107$ Hz and $T_2=32$ ms (the dotted curve). The fact that the top of the initial section of

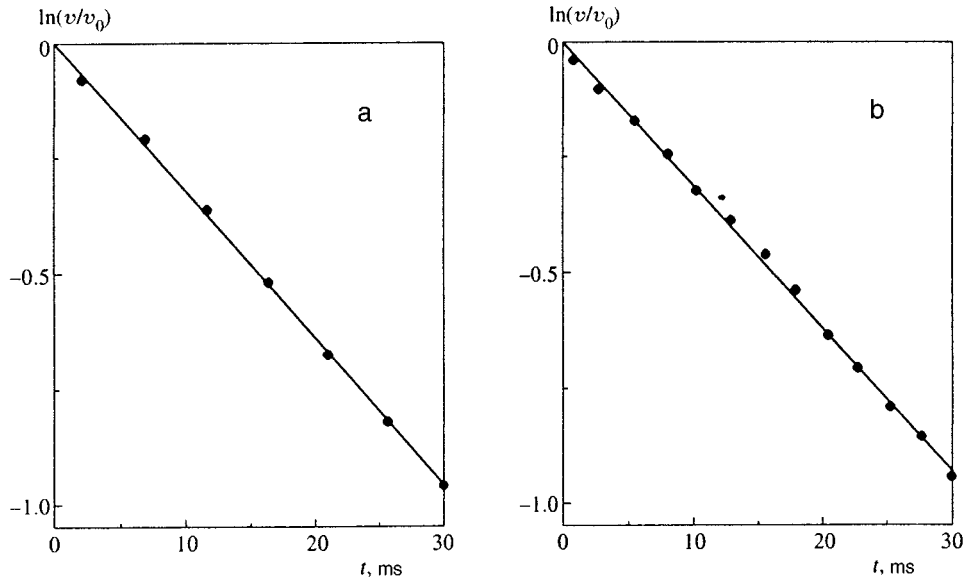

FIG. 2. Time dependence of the nutation signal amplitude at (a) $\omega_1=2\pi\times107$ Hz and (b) $\omega_1=2\pi\times426$ Hz. The solid lines represent the approximation of the experimental data by an exponential function.

the detected signal is shifted in relation to the theoretical value is due to the drift of the zero line in the course of the rf pulse, the drift being caused by imperfect matching of the rf bridge. This shift was found to increase with rf power, but it had no effect on the amplitude properties of the nutations and was taken into account in processing the data.

Figure 2 depicts the decay of nutations for two different values of the amplitude of the driving rf field. The data show that in both cases the decay follows an exponential law with a characteristic time τ close to the value of T_2 for the system.

The dependence of the decay time for transient nutations on the amplitude of the driving rf field is depicted in Fig. 3. The data show that the nutation decay time for $\omega_1T_2<20$ increases somewhat (by 10%) with H_1. This is caused by temporal instability in the polarizing magnetic field, which changes the nutation frequency (see Eq. (2)) due to variations in $\omega_0-\omega$ in the course of the experiment. Averaging signals with distinct frequencies leads to the above-mentioned decrease in τ during the coherent integration process. As H_1 grows, the effect of the instability of the polarizing magnetic field diminishes (in accordance with (2)), and vanishes when ω_1T_2 becomes greater than 20. As Fig. 3 shows, in the $20<\omega_1T_2<106$ interval the nutation decay time remains constant at 33.0 ± 1.5 ms. This coincides with the known relaxation time T_2 for protons in glycerin,[9] which suggests that within this range of rf-field amplitude the field's inhomogeneity is negligible in this experiment. The interval corresponds to a high-power driving field, in view of which it appears that within experimental error (5%) nutation decay is independent of the driving field amplitude and can by described by Bloch's model. A further increase in $H_1(\omega_1T_2>106)$ drives the rate of nutation decay up (the corresponding data are not depicted in Fig. 3). The decay becomes exponential and results from inhomogeneities in the rf field.

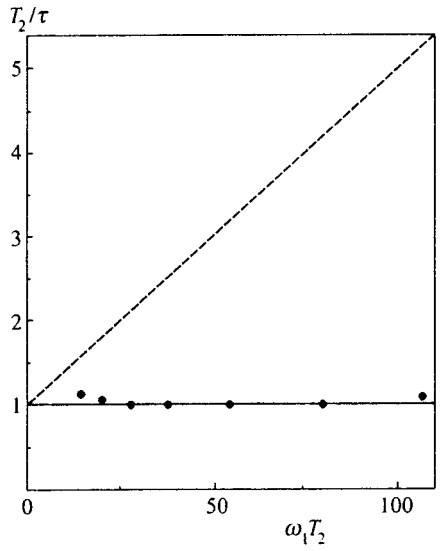

FIG. 3. Dependence of the nutation decay time on the Rabi frequency. The points represent the experimental data, the solid line represents Bloch's model, and the dashed line was obtained from Eq. (4) with $\beta=0.021$.

It is worth noting that the nutation decay time measured by Torrey[6] in a 1.5 cm^3 sample of glycerin was about 11 ms and resulted from inhomogeneities in the rf field.

According to Boscaino *et al.*,[1] the decay time of nutations of a homogeneous line depends on the amplitude of the driving electromagnetic field. The predicted ω_1-dependence of the nutation decay time (Eq. (4)) is depicted in Fig. 3 by a dashed line. If we assume that for two-level quantum systems this dependence is universal, comparison of the EPR data on paramagnetic defects in quartz and the NMR data on protons in glycerin shows that there is no correlation between the two sets. If Eq. (4) is valid, the parameter β for proton nutation decay must be smaller than this quantity for EPR nutations in quartz by a factor of at least 100. In this case determining the small ($<10^{-4}$) but finite value of β requires increasing the accuracy in measuring the parameters of proton nutation decay.

On the other hand, the fact that the decay time of proton nutations coincides with T_2 suggests that for the given quan-

tum system $\beta=0$. This means that the decay in two-level spin systems with $T_2\approx T_1$ is described by Bloch's model, with the result that the purported ω_1-dependence of the effective width of a homogeneous spectral line is not corroborated by experiments. In this situation, the reason for the observed anomalous (non-Bloch) decay of EPR nutations in quartz can be established, in particular, by studying the ω_1-dependence of nutation decay in this object in the single-photon mode (to eliminate features of the two-photon mode from nutation studies). It is also important to establish the features of the nutation process due to the dynamics of quantum systems in the presence of a high-power driving field for distinct T_2 and T_1.

5. CONCLUSION

NMR studies of protons in glycerin ($T_2\approx T_1$) have shown that the rate of decay of transient nutations is independent (to within 5%) of the amplitude of the driving electromagnetic field up to $\omega_1 T_2\approx 110$. This suggests that for homogeneously broadened two-level spin systems Bloch's model is valid in the high-power driving field limit ($\omega_1 T_2 \gg 1$). The data obtained in these experiments do not corroborate the hypothesis of Boscaino *et al.*[1] about the dependence of the effective width of a homogeneous spectral line on the amplitude H_1 of the driving variable field, the reason being that the value of β, which reflects the extent of the action of H_1 on the nutation decay rate, proved to be smaller than the value of β measured for EPR nutations on quartz by a factor no less than 100. This suggests that non-Bloch nutation decay is related to the features of quantum systems at $T_2 \neq T_1$, in view of which the problem of interpreting the ω_1-dependence of the nutation decay rate for quartz[1] remains unresolved and requires further experimental investigation.

The author is grateful to I. Z. Rutkovskiĭ and V. S. Kuz'min for discussing the results. Partial financial support for this work was provided by the Belarus Fund for Fundamental Research.

[1] R. Boscaino, F. M. Gelardi, and J. P. Corb, Phys. Rev. B **48**, 7077 (1993).
[2] V. S. Kuz'min, Zh. Prikl. Spektrosk. **63**, 229 (1996).
[3] V. S. Malinovskiĭ, Zh. Éksp. Teor. Fiz. **108**, 1907 (1995) [JETP **81**, 1041 (1995)].
[4] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford (1961).
[5] R. L. Shoemaker, in *Laser and Coherence Spectroscopy*, J. I. Steinfeld (ed.), Plenum Press, New York (1978), p. 197.
[6] H. C. Torrey, Phys. Rev. **76**, 1059 (1949).
[7] R. E. Orlowski and A. H. Zewail, J. Chem. Phys. **70**, 1390 (1979).
[8] V. S. Kuz'min, A. P. Saĭko, and G. G. Fedoruk, Zh. Éksp. Teor. Fiz. **99**, 215 (1991) [Sov. Phys. JETP **72**, 121 (1991)].
[9] A. A. Vashman and I. S. Pronin, *Nuclear Magnetic Relaxation Spectroscopy* [in Russian], Énergoatomizdat, Moscow (1986), p. 100.

Translated by Eugene Yankovsky

Elastic scattering of low-energy electrons by a uranium atom: reliability of theoretical predictions of based on a model description

D. P. Grechukhin and A. V. Lomonosov*)

Russian ''Kurchatov Institute'' Science Center, 123182 Moscow, Russia
(Submitted 5 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1214–1228 (April 1997)

We used the method of phase functions to solve the radial relativistic Dirac equation and nonrelativistic Schroedinger equation. With these solutions, we investigated the elastic scattering of slow electrons by a uranium atom, and obtained numerical values for the total cross section and elastic scattering phases. In order to check the correctness of the results found from the method of phase functions, in all cases we also solved the Dirac and Schroedinger equations by direct numerical integration. Several types of polarization and exchange potentials were used to simulate the scattering process. We conclude that the strong dependence of the cross section for elastic scattering of an electron by uranium on the shape of the effective potential of the latter at small kinetic energies ($E_k<5$ eV) makes it impossible to predict the presence or absence of a Ramsauer effect reliably. © *1997 American Institute of Physics.* [S1063-7761(97)00704-X]

1. INTRODUCTION

The Ramsauer effect (i.e., the presence of a deep minimum in the elastic scattering cross section of slow electrons by atoms) was first observed experimentally for atoms of certain inert gases, for instance argon. In Ref. 1, Kudrin and Drozdov argued that this effect could be exploited to make an MHD generator based on a nonequilibrium two-temperature uranium plasma. In this device, a spectrum of electrons would be generated with kinetic energies E_k in the energy neighborhood of the Ramsauer minimum cross section in order to create a medium with high electrical conductivity. These authors calculated the cross section for elastic scattering of slow electrons by a uranium atom for the range of electron kinetic energies $0<E_k<40$ eV using a nonrelativistic version of the method of phase functions.[2,3] The model of the interaction of an electron with the uranium atom used in Ref. 1 was based on the Thomas–Fermi model,[4] including the polarization potential[5] $V_p(r)$ arising from the electric dipole moment induced by the slow electron. The experimental determination of the elastic scattering cross section for uranium at kinetic energies less than 10 eV is a complicated problem, and we do not know of anyone who has solved it. In this range of energies E_k practical theoretical calculations of the scattering cross section are all single-particle calculations in which the interaction between the scattered electron and the multielectron system (a heavy atom) is approximated by a spherically symmetric potential $V(r)$ that decreases faster than r^{-1} as $r\to\infty$:

$$V(r)=V_a(r)+V_p(r)+V_{\rm ex}(r), \tag{1}$$

where $V_a(r)$ is the average atomic potential, determined by the density distribution $\rho(r)$ of electrons in the atom, $V_p(r)$ is the polarization potential, and $V_{\rm ex}(r)$ is the exchange potential.

All these components of the interaction are in turn modeled by various radial dependences, whose parameters are chosen based on certain atomic characteristics. In this situation, questions arises regarding the reliability of theoretical predictions of the behavior of the total elastic cross section (especially in the range of kinetic energies $E_k<10$ eV of interest to us, where s, p, and d partial waves all contribute to the total cross section and the Ramsauer effect is possible) and the stability of this quantity against admissible changes in the shapes of the potentials $V_a(r)$, $V_p(r)$, and $V_{\rm ex}(r)$.

It is possible to describe the elastic scattering of a slow electron by an atom either relativistically (the Dirac equation) or nonrelativistically (the Schroedinger equation). In our previous papers,[6,7] we used the example of low-energy elastic scattering of electrons by atoms of the inert gases Ar, Kr, and Xe to show that it is important to use a relativistic approach to treat the scattering of slow electrons. We found that for electron energies less than 1 eV the discrepancy between computed values of the total cross section for the relativistic and nonrelativistic cases increases with increasing atomic number Z, and that the energy behavior of the cross section was quite sensitive to the choice of models for the atom–electron interaction potential.

In this paper we continue our analysis of this situation. Using a single-particle treatment of the scattering, we calculate values of the elastic scattering cross section of electrons by a uranium atom with various types of interactions between the electron and the atom obtained from the literature. Using the method of phase functions,[2,3] we solve the radial relativistic Dirac and nonrelativistic Schroedinger equations for the problem of elastic scattering by a spherically symmetric potential $V(r)$ which decreases at infinity faster than r^{-1}. In all cases, we checked the results obtained by the phase method by direct numerical integration of the radial Dirac and Schroedinger equations and subsequent matching with their asymptotic solutions.

2. SOLUTION OF THE RADIAL DIRAC EQUATION BY THE METHOD OF PHASE FUNCTIONS

The solution to the Dirac equation in a centrosymmetric field is given in the form of the following bispinor:[8]

$$\Psi_{jlm}(\mathbf{r})=\begin{cases} g(r)\Omega_{jlm}(\mathbf{r}/r), \\ if(r)\Omega_{jl'm}(\mathbf{r}/r), \end{cases} \tag{2}$$

where $\Omega_{jlm}(\mathbf{r}/r)$ is a spherical spinor, $l'=2j-1$, j is the quantum number corresponding to a definite value of the total angular momentum operator $\hat{\mathbf{J}}=\hat{\mathbf{L}}+\hat{\mathbf{S}}$ (where $\hat{\mathbf{S}}$ is the electron-spin operator), and $g(r)$ and $f(r)$ are radial functions that satisfy the system of differential equations

$$\frac{d}{dr}[rg(r)]+\frac{\kappa}{r}[rg(r)]-\frac{1}{\hbar c}[E+mc^2-V(r)]rf(r)=0, \tag{3}$$

$$\frac{d}{dr}[rf(r)]-\frac{\kappa}{r}[rf(r)]-\frac{1}{\hbar c}[E-mc^2-V(r)]rg(r)=0, \tag{4}$$

Here $\kappa=-j(j+1)+l(l+1)-1/4$, E is the total energy of the electron, m is the electron mass, and c is the velocity of light in vacuum.

In a spherically symmetric potential $V(r)$ that decreases at infinity faster than r^{-1}, the large $(g(r))$ and small $(f(r))$ components of the electron wave function in the continuous spectrum $(E>mc^2)$ have the following asymptotic behavior as $r\to\infty$:

$$g(r)=\sqrt{\frac{E+mc^2}{2E}}\ \frac{\sin(pr-\pi l/2+\delta_{jl})}{pr}, \tag{5}$$

$$f(r)=i^{l-l'+1}\sqrt{\frac{E-mc^2}{2E}}\ \frac{\sin(pr-\pi l'/2+\delta_{jl})}{pr}, \tag{6}$$

where $p=\sqrt{E^2/c^2-m^2c^2}/\hbar$ is the wave vector and δ_{jl} is the elastic scattering phase.

In accordance with the method of phase functions, we replace the radial functions $g(r)$ and $f(r)$ by a phase function $\delta_{jl}(r)$ and an amplitude function $A_{jl}(r)$ (see Ref. 3). Then

$$g(r)=\sqrt{\frac{E+mc^2}{2E}}\,A_{jl}(r)[\cos\,\delta_{jl}(r)j_l(pr) - \sin\,\delta_{jl}(r)\,\eta_l(pr)], \tag{7}$$

$$f(r)=i^{l-l'+1}\sqrt{\frac{E-mc^2}{2E}}\,A_{jl}(r)[\cos\,\delta_{jl}(r)j_{l'}(pr) - \sin\,\delta_{jl}(r)\,\eta_{l'}(pr)], \tag{8}$$

where $j_l(pr)$ and $n_l(pr)$ were defined in Ref. 9.

Substituting (7) and (8) into (3) and (4), we obtain directly the equations for the phase and amplitude functions with the boundary conditions $\delta_{jl}(0)=0$, $A_{jl}(\infty)=1$ (see Ref. 2):

$$\frac{d}{dr}\,\delta_{jl}(r)=\frac{V(r)(pr)^2}{\hbar c}\left\{\sqrt{\frac{E+mc^2}{E-mc^2}}\,[\cos\,\delta_{jl}(r)j_l(pr) - \sin\,\delta_{jl}(r)\,\eta_l(pr)]^2 + \sqrt{\frac{E-mc^2}{E+mc^2}}\,[\cos\,\delta_{jl}(r)j_{l'}(pr) - \sin\,\delta_{jl}(r)\,\eta_{l'}(pr)]^2\right\}, \tag{9}$$

$$\frac{d}{dr}\,A_{jl}(r)=-\frac{A_{jl}(r)V(r)(pr)^2}{\hbar c} \times\left\{\sqrt{\frac{E+mc^2}{E-mc^2}}\,[\cos\,\delta_{jl}(r)j_l(pr) - \sin\,\delta_{jl}(r)\,\eta_l(pr)][\cos\,\delta_{jl}(r)\,\eta_l(pr) - \sin\,\delta_{jl}(r)j_l(pr)] + \sqrt{\frac{E-mc^2}{E+mc^2}}\,[\cos\,\delta_{jl}(r)j_{l'}(pr) - \sin\,\delta_{jl}(r)\,\eta_{l'}(pr)][\cos\,\delta_{jl}(r)\,\eta_{l'}(pr) + \sin\,\delta_{jl}(r)j_{l'}(pr)]\right\}. \tag{10}$$

Let us go to the nonrelativistic limit in Eqs. (9) and (10). In this case $\sqrt{(E+mc^2)/(E-mc^2)}\to 2mc/\hbar k$, $\sqrt{(E-mc^2)/(E+mc^2)}\to 0$, $p\to k$ (the nonrelativistic expression for the kinetic energy is $E_k=(\hbar k)^2/2m$), and we obtain equations for the phase and amplitude functions[1)] (see Ref. 3) with boundary conditions $\delta_l(0)=0$, $A_l(\infty)=1$:

$$\frac{d}{dr}\,\delta_l(r)=\frac{2m}{h^2}\,V(r)kr^2[\cos\,\delta_{jl}(r)j_l(kr) - \sin\,\delta_{jl}(r)\,\eta_l(kr)]^2, \tag{11}$$

$$\frac{d}{dr}\,A_l(r)=-\frac{2m}{h^2}\,A_l(r)V(r)kr^2[\cos\,\delta_l(r)j_l(kr) - \sin\,\delta_l(r)\,\eta_l(kr)][\cos\,\delta_l(r)\,\eta_l(kr) - \sin\,\delta_l(r)j_l(kr)]. \tag{12}$$

When a nonrelativistic particle is scattered by a potential whose characteristic depth is of order mc^2, it is necessary to include relativistic corrections. It is clear from Eq. (9) that for such potentials we cannot neglect the second term and simply make the transition to Eq. (11).

In conclusion, we give the expressions that relate the elastic scattering cross section σ to the elastic scattering phases δ_{jl} in the relativistic case.[8] The scattering cross section is expressed in terms of a sum of partial cross sections:

$$\sigma=\sum_{lj}\,\sigma_{lj}=\frac{4\pi}{p^2}\sum_{lj}\,|\kappa|\sin^2\,\delta_{lj}\,. \tag{13}$$

For the case of low energy scattering $\delta_{l,l+1/2}\approx\delta_{l,l-1/2}\approx\delta_l$, and Eq. (13) reduces to the nonrelativistic expression

$$\sigma=\sum_{l=0}^{\infty}\,\sigma_l=\frac{4\pi}{k^2}\sum_{l=0}^{\infty}\,(2l+1)\sin^2\,\delta_l\,. \tag{14}$$

3. MODEL POTENTIAL FOR SCATTERING OF AN ELECTRON BY A URANIUM ATOM

The state of an electronic shell of a uranium atom is described by using the relativistic Hartree–Fock–Slater method. In order to determine the stability of our results against a change in the average atomic potential $V_a(r)$, let us change the centrosymmetric Slater potential $V_{Slt}(r)$, which determines the electronic atomic wave functions required to obtain the electron density $\rho(r)$. These changes will be made in two ways:

1. By changing the initial electronic configuration of the uranium atom. For this, in addition to the ground electronic configuration of uranium $5f^3_{5/2}6d^1_{3/2}7s^2_{1/2}$, we used the following possible configurations: $5f^4_{5/2}6d^0_{3/2}7s^2_{1/2}$, $5f^2_{5/2}6d^2_{3/2}7s^2_{1/2}$, $5f^1_{5/2}6d^3_{3/2}7s^2_{1/2}$, and $5f^0_{5/2}6d^4_{3/2}7s^2_{1/2}$. The choice of this sequence of configurations for the uranium atom $5f^{\,n}_{5/2}6d^{4-n}_{3/2}7s^2_{1/2}$, where $n=1, 2, 3, 4$, is motivated by the fact that a change in the population of the spatially compact $5f_{5/2}$ orbit due to transition of electrons from the spatially extended orbit $6d_{3/2}$ can cause a considerable change to the potential $V_{Slt}(r)$. We illustrate this assertion by calculating the average squared radius of the orbit for the sequence of configurations $5f^{\,n}_{5/2}6d^{4-n}_{3/2}7s^2_{1/2}$:

$$\langle 5f_{5/2}|\left(\frac{r}{a_0}\right)^2|5f_{5/2}\rangle=\begin{cases}1.73, & n=0,\\ 2.79, & n=4,\end{cases}$$

$$\langle 6d_{3/2}|\left(\frac{r}{a_0}\right)^2|6d_{3/2}\rangle=\begin{cases}8.24, & n=0,\\ 19.48, & n=4,\end{cases}$$

where a_0 is the Bohr radius, whose value is $a_0=\hbar^2/me^2=0.5291772\times10^{-10}$ m;

2. By changing the form of the correlation-exchange potential used in the iterative procedure for obtaining $V_{Slt}(r)$. We will use the following forms of the correlation-exchange interaction:

a) the standard Slater form[10] with two choices of the Slater exchange parameter α_{Slt}: $\alpha_{Slt}=2/3$ and $\alpha_{Slt}=1$;

b) the form based on Ref. 11;

c) the form from Ref. 12;

d) the form from Ref. 13.

In addition, we calculated the atomic potential using the nonrelativistic Thomas–Fermi–Dirac model[4] with a subsequent integration of the relativistic Dirac equation.

The polarization potential we used had the form[14]

$$V_p(r)=\begin{cases}V_p^{SR}(r), & r\leqslant r_c,\\ V_p^{LR}(r), & r>r_c,\end{cases} \tag{15}$$

$$V_p^{SR}(r)=\begin{cases}0.0622\ln r_s-0.096+0.018r_s\ln r_s-0.02r_s, & r_s\leqslant 0.7\\ -0.1231+0.03796\ln r_s, & 0.7<r_s<10,\\ -0.876r_s^{-1}+2.65r_s^{-3/2}-2.8r_s^{-2}-0.8r_s^{-5/2}, & r_s\geqslant 10,\end{cases} \tag{16}$$

$$V_p^{LR}(r)=-e^2\alpha_{\rm U}/2r^4, \tag{17}$$

where $r_s=\{3/4\pi\rho(r)\}^{1/3}$, e is the charge of an electron, $\alpha_{\rm U}$ is the dipole polarizability of uranium, and r_c is the point where $V_p^{SR}(r)$ and $V_p^{LR}(r)$ intersect.

This potential was used both in our previous papers[6,7] to calculate elastic scattering by inert gases, and in Ref. 14 to describe elastic scattering by alkali metal atoms. At this time, we know of no experimental values of the dipole polarizability of uranium $\alpha_{\rm U}$. In Ref. 1, the following value of the polarizability of uranium was obtained using the Thomas-Fermi-Dirac model:[4] in atomic units, $\alpha_{\rm U}/a_0^3=29.36$, which is somewhat larger than the experimentally known polarizability of the more compact xenon atom $\alpha_{\rm Xe}/a_0^3=27$. Actually, the dipole polarizability is a variational parameter, whose value we will take to be $\alpha_{\rm U}/a_0^3=30$; for this value we find $r_c=3.7575268$.

In order to study the question of stability of our results with respect to changes in the shape of the polarization potential, we also used a polarization potential of the form[5]

$$V_p(r)=-\frac{e^2\alpha_{\rm U}}{2(r^2+r_v^2)^2}, \tag{18}$$

where r_v is a fitting parameter of the model determined from agreement between numerical calculations and experimental data, and is usually rather well described by the expression $r_v=(\alpha/2)^{1/3}$. For our value of the polarizability of uranium $\alpha_{\rm U}$ we pick r_v to satisfy $r_v/a_0=2.45$.

As a model exchange potential we used the expression proposed in Ref. 15:

$$V_{\rm ex}(r)=-\frac{2}{\pi}K_F F(\eta),\quad F(\eta)=\frac{1}{2}+\frac{1-\eta^2}{4\eta}\ln\left|\frac{1+\eta}{1-\eta}\right|. \tag{19}$$

where

$$K_F=[3\pi^2\rho(r)]^{1/3},\quad \eta=\frac{K}{K_F}.$$

If we include relativistic corrections we have

$$K=\sqrt{\frac{(I+E+\sqrt{c^2K_F^2+m^2c^4})^2}{c^2}-m^2c^2}, \tag{20}$$

while without them we have

$$K=\sqrt{K_F^2+k^2+2I}, \tag{21}$$

where I is the ionization potential of the atom.

In order to study the stability of our results with respect to variations of the shape of the model exchange interaction we used[16]

$$V_{\rm ex}(r)=\frac{1}{2}\{E_k-V_a(r)-V_p(r)\}-\frac{1}{2}\sqrt{(E_k-V_a(r)-V_p(r))^2-4\pi\rho(r)\frac{e^2\hbar^2}{m}}. \tag{22}$$

TABLE I. Numerical calculations of the total elastic scattering cross section σ (in atomic units) as a function of the electron kinetic energy E_k for a U atom with an unperturbed potential $V_a(r)$ obtained from the correlation-exchange term of Ref. 11 with the third set of parameters, with the polarization potential $V_p(r)$ of Ref. 14, and the exchange potential $V_{ex}(r)$ of Ref. 15.

E_k,	σ/a_0^2				E_k,	σ/a_0^2			
eV	1	2	3	4	eV	1	2	3	4
10^{-8}	104.0	461.9	106.9	462.0	4.410	114.2	67.4	112.7	65.4
0.010	144.9	511.0	145.1	511.2	4.840	100.8	61.8	99.8	60.8
0.040	217.4	545.4	219.2	545.6	5.290	91.4	60.2	89.9	58.7
0.090	359.0	579.1	359.3	579.2	5.760	84.4	60.4	82.8	58.8
0.165	665.6	659.4	665.4	659.2	6.250	80.2	62.7	77.9	60.4
0.250	1022.4	784.9	1022.0	784.6	7.840	74.8	72.7	72.8	70.8
0.365	1122.4	911.4	1122.1	911.0	9.000	76.0	81.1	74.4	79.7
0.490	993.0	921.3	992.8	920.8	10.240	78.9	89.1	77.8	88.1
0.640	838.2	830.7	838.0	830.5	11.560	82.9	96.5	82.2	95.7
0.810	710.7	702.8	710.5	702.7	12.960	87.4	102.7	86.9	102.2
1.000	607.5	577.4	607.3	577.4	14.440	91.8	107.8	91.5	107.5
1.440	448.6	376.6	448.4	376.8	16.000	96.1	111.8	95.8	111.6
2.250	282.3	193.2	282.1	193.1	20.250	104.2	116.9	104.1	116.8
2.560	241.0	155.2	240.7	154.8	25.000	108.1	116.3	108.0	116.3
2.890	205.7	125.6	205.2	125.0	30.250	107.9	111.8	107.9	111.8
3.240	175.7	102.9	175.1	102.2	36.000	104.7	105.1	104.7	105.1
3.610	150.9	86.3	149.9	85.3	42.250	99.5	97.3	99.5	97.3
4.000	130.9	75.1	129.3	73.4					

Note. Column 1 is the total cross section including relativistic corrections based on Eqs. (13), (9), and (10); column 2, the total cross section without including relativistic corrections, is based on Eqs. (14), (11), and (12); column 3 is the total cross section (relativistic case) obtained by direct solution of the Dirac equation; and column 4 is the total cross section (nonrelativistic case) obtained by direct solution of the Schröedinger equation.

4. NUMERICAL CALCULATIONS

We solved the systems of differential equations (9), (10) (11), and (12) numerically using the Adams predictor–corrector method.[17] In addition, we found the value of the phase by direct integration of the radial Dirac Eqs. (3) and (4) with subsequent matching to the asymptotic solutions (7) and (8), and also by integration of the Schroedinger equation without including relativistic corrections.

The results of these numerical calculations, which we present here, lead us to the following conclusions.

1. In accordance with assertions of our previous papers,[6,7] at low electron kinetic energies the values of the elastic scattering cross section differ significantly for the relativistic and nonrelativistic cases.

2. The values of the cross section obtained for uranium when $E_k < 5$ eV holds are sensitive to changes in the shapes of all three terms of the total interaction potential between an electron and an atom, i.e., the unperturbed atomic potential $V_a(r)$, the polarization potential $V_p(r)$, and the exchange potential $V_{ex}(r)$.

3. The presence or absence of a Ramsauer effect, and also the particular electron kinetic energy at which the elastic scattering cross section is a minimum when there is a Ramsauer effect, are contingent on the detailed potential shape.

In what follows we will illustrate these assertions with specific examples.

Table I shows the results of calculations of the total cross section for elastic scattering of an electron by a uranium atom with the unperturbed potential $V_a(r)$ obtained by using the correlation-exchange term from Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ taken from Ref. 14, and the exchange potential $V_{ex}(r)$ taken from Ref. 15. In Table I we list the results as a function of the electron kinetic energy E_k: the first column is the total cross section, including relativistic corrections based on Eqs. (13), (9), and (10); the second is the total cross section without including relativistic corrections based on Eqs. (14), (11), and (12); the third is the total cross section (relativistic case) obtained by direct integration of the Dirac equation; and the fourth is the total cross section (nonrelativistic case) obtained by direct integration of the Schrödinger equation. By comparing columns 1 and 2 (3 and 4), we can estimate the magnitude of the relativistic effects associated with electron scattering.

The solution to the Dirac and Schroedinger equations by the method of phase functions allows us to consider the radial dependence of both the phase functions $\delta_{jl}(r)$ and $\delta_l(r)$ and the amplitude functions $A_{jl}(r)$ and $A_l(r)$. If the depth of the effective potential $V(r)$ Eq. (1) is of order mc^2, then the second terms in Eqs. (9) and (10) which we neglected in going to the nonrelativistic limit in Eqs. (11) and (12), contribute to the phase and amplitude functions. Once we know the radial behavior of these functions, we can determine at what value of radius vector r relativistic corrections play a significant role and how the set of phases $\delta_{jl}(r)$ develops. From this we can estimate the influence of the effective potential shape. Figure 1 shows plots of the phase and amplitude functions versus distance from the cen-

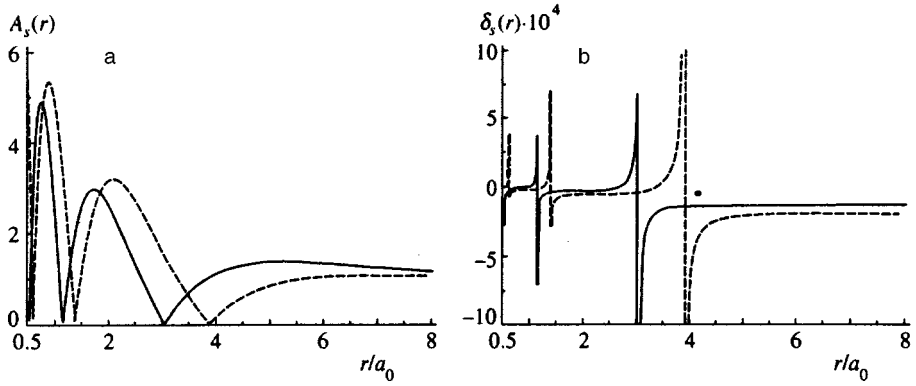

FIG. 1. Dependence of the amplitude (a) and phase (b) functions on distance from the center of the nucleus (in atomic units) for an s-wave at electron kinetic energy $E_k = 10^{-8}$ eV. The unperturbed potential $V_a(r)$ is taken from from the correlation-exchange term of Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ is from Ref. 14, and the exchange potential $V_{\text{ex}}(r)$ is from Ref. 15. The solid curves are relativistic calculations, the dashed curves nonrelativistic.

ter of the nucleus for s-waves and an electron kinetic energy $E_k = 10^{-8}$ eV, and the unperturbed potential $V_a(r)$, polarization potential $V_p(r)$, and exchange potential $V_{\text{ex}}(r)$ chosen above. The solid curve is the result of a calculation for the relativistic case, while the dashed curve is for the nonrelativistic case. The jumps in the phase function ($-\pi \leqslant \delta_s \leqslant \pi$) and zeros in the amplitude function corresponding to them match the zeros of the radial wave functions $g(r)$ (Ref. 7) and $f(r)$ (Ref. 8). From Fig. 1 it is easy to see that these points shift on the radial scale as we go from the nonrelativistic to the relativistic case. Thus, the relativistic corrections for the uranium atom must be included both in the phase and the amplitude functions.

In order to check our calculations of the phase functions $\delta_{jl}(r)$ and amplitude functions $A_{jl}(r)$, we calculated the wave radial functions $g(r)$ and $f(r)$ from Eqs. (7) and (8) and compared them with the same functions obtained directly from an integration of the radial Dirac equations (3) and (4) with subsequent matching to the asymptotic solutions (5) and (6). The results of the two methods agree completely. The phase functions $\delta_l(r)$ and amplitude functions $A_l(r)$ from the nonrelativistic Schrödinger equation were checked in the same way.

The relativistic and nonrelativistic calculations differ considerably for kinetic energies less than 1 eV, since the characteristic momentum of an electron within the atom is so large that relativistic corrections play an important role. Therefore, in what follows we give results only for the relativistic case.

The slight discrepancy between the results of the two computational methods we have used at low electron energies is explained by complications that arise when systems of differential equations are solved numerically by the phase method; nevertheless, when the phases are calculated by these two methods, they agree to an accuracy of $\sim 0.01\%$; if this were not so, there would be a rather large difference in the total scattering cross section. For definiteness we also give results obtained by using the solution to the relativistic Dirac equation by the phase method (9) and (10).

Let us consider the effect of changing the unperturbed atomic potential $V_a(r)$. Figure 2 shows the results of numerical calculations of the total cross section σ for elastic scattering of an electron by a uranium atom as a function of the electron kinetic energy E_k, using the polarization potential $V_p(r)$ from Ref. 14 and exchange potential $V_{\text{ex}}(r)$ from Ref. 15; the unperturbed atomic potential of uranium was calculated using the relativistic Hartree–Fock–Slater method with various exchange-correlation potentials. Curve *1* is the Slater form (with parameter $\alpha_{Slt} = 2/3$); *2* is the potential from Ref. 11 with the third set of parameters; *3* is from Ref. 12; *4* is from Ref. 13; *5* is the Slater form (with parameter $\alpha_{Slt} = 1$); and *6* is a calculation of the unperturbed atomic potential using the nonrelativistic Dirac model.[4] From this figure it is easy to see that model perturbations of the average atomic field lead to different behavior of the function σ for electron kinetic energies less than 1 eV, whereas for $E_k > 20$ eV all of these models give similar results.

Table II lists the results of calculations in which the unperturbed atomic potential $V_a(r)$ is modified by changing the initial electron configuration of the uranium atom; $V_a(r)$ is obtained from the correlation-exchange term given in Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ is taken from Ref. 14 and the exchange potential $V_{\text{ex}}(r)$ from Ref. 15. The first column is the total cross section for electron configuration $5f^4_{5/2}6d^0_{3/2}7s^2_{1/2}$, the second for $5f^3_{5/2}6d^1_{3/2}7s^2_{1/2}$, the third for $5f^2_{5/2}6d^2_{3/2}7s^2_{1/2}$, the

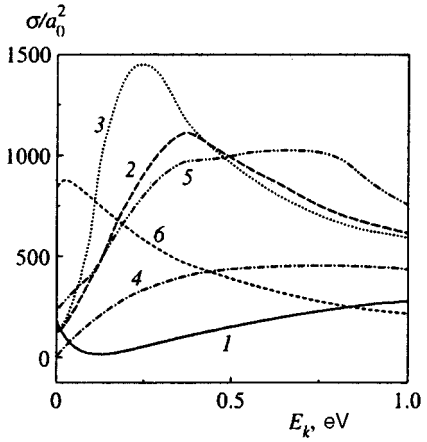

FIG. 2. Dependence of the total elastic scattering cross section for uranium σ (in atomic units) on the electron kinetic energy E_k with the polarization potential $V_p(r)$ from Ref. 14, the exchange potential $V_{\text{ex}}(r)$ from Ref. 15, and an unperturbed atomic potential obtained within the framework of the relativistic Hartree–Fock–Slater method. The exchange-correlation potentials were taken from: Ref. 10, with $\alpha_{Slt} = 2/3$ (curve *1*), Ref. 11 with the third set of parameters (curve *2*); Ref. 12 (curve *3*); Ref. 13 (curve *4*); and Ref. 10, with $\alpha_{Slt} = 1$ (curve *5*). Curve *6* is calculated using the nonrelativistic Thomas–Fermi–Dirac model (see Ref. 4).

TABLE II. Numerical calculations of the total elastic scattering cross section σ (in atomic units) as a function of the electron kinetic energy E_k for uranium with an unperturbed potential $V_a(r)$ obtained from the correlation-exchange term of Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ of Ref. 14, and the exchange potential $V_{ex}(r)$ of Ref. 15.

E_k,	σ/a_0^2					E_k,	σ/a_0^2				
eV	1	2	3	4	5	eV	1	2	3	4	5
10^{-8}	88.6	104.0	97.5	90.3	84.6	4.410	112.9	114.2	116.1	118.2	120.0
0.010	125.3	144.9	134.2	126.2	121.1	4.840	100.1	100.8	102.1	103.5	104.8
0.040	228.9	217.4	199.9	190.4	191.6	5.290	90.5	91.4	91.8	92.4	92.8
0.090	664.3	359.0	319.1	295.2	288.3	5.760	83.5	84.4	84.2	83.9	83.5
0.165	1691.6	665.6	570.5	493.4	445.2	6.250	78.7	80.2	79.2	78.2	76.8
0.250	1616.5	1022.4	892.7	759.6	658.5	7.840	73.8	74.8	73.1	71.0	68.0
0.365	1213.7	1122.4	1051.7	957.0	861.3	9.000	75.4	76.0	74.1	71.8	68.5
0.490	978.9	993.0	962.6	922.7	875.2	10.240	78.9	78.9	76.9	74.4	71.3
0.640	818.6	838.2	821.0	803.7	784.1	11.560	83.4	82.9	80.8	78.2	75.0
0.810	700.7	710.7	698.7	688.4	678.3	12.960	88.1	87.4	85.0	82.2	79.0
1.000	605.1	607.5	599.1	592.1	585.7	14.440	92.7	91.8	89.4	86.4	83.1
1.440	451.7	448.6	446.5	444.0	441.2	16.000	96.8	96.1	93.6	90.5	87.1
2.250	283.8	282.3	285.8	288.0	289.4	20.250	104.2	104.2	102.4	99.7	96.5
2.560	241.6	241.0	245.2	248.2	250.4	25.000	106.9	108.1	107.6	105.8	103.5
2.890	205.5	205.7	210.0	213.5	216.1	30.250	105.6	107.9	108.8	108.4	107.2
3.240	175.0	175.7	179.8	183.4	186.2	36.000	101.4	104.7	106.7	107.4	107.4
3.610	149.8	150.9	154.4	157.7	160.4	42.250	95.6	99.5	102.2	103.6	104.4
4.000	129.3	130.9	133.6	136.2	138.4						

Note. The total cross sections are listed here for the following configurations: column 1—$5f_{5/2}^4 6d_{3/2}^0 7s_{1/2}^2$, column 2—$5f_{5/2}^3 6d_{3/2}^1 7s_{1/2}^2$, column 3—$5f_{5/2}^2 6d_{3/2}^2 7s_{1/2}^2$, column 4—$5f_{5/2}^1 6d_{3/2}^3 7s_{1/2}^2$, and column 5—$5f_{5/2}^0 6d_{3/2}^4 7s_{1/2}^2$.

fourth for $5f_{5/2}^1 6d_{3/2}^3 7s_{1/2}^2$, and the fifth for $5f_{5/2}^0 6d_{3/2}^4 7s_{1/2}^2$. The model variations of the average atomic field also lead to different values of σ for electron kinetic energies less than 1 eV. Note the absence of a Ramsauer effect for all the calculations presented in Table II.

In order to determine the sensitivity of our results to a change in the shape of the polarization interaction, we carried out additional calculations in which we used the unperturbed atomic potential $V_a(r)$ obtained from the exchange-correlation potential of Ref. 11 with the third set of

TABLE III. Overlap integrals I_b of radial wave functions for the scattered and atomic electrons in the n-shell as a function of the electron kinetic energy E_k for uranium with an unperturbed potential $V_a(r)$ obtained from the correlation-exchange term from Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ from Ref. 14 and the exchange potential $V_{ex}(r)$ from Ref. 15.

E_k,	I_b, %							E_k,	I_b, %						
eV	$7s_{1/2}$	$6p_{1/2}$	$6p_{3/2}$	$6d_{3/2}$	$5d_{5/2}$	$5f_{5/2}$	$4f_{7/2}$	eV	$7s_{1/2}$	$6p_{1/2}$	$6p_{3/2}$	$6d_{3/2}$	$5d_{5/2}$	$5f_{5/2}$	$4f_{7/2}$
10^{-8}	1.4	0.0	0.3	0.1	0.0	0.0	0.0	4.410	13.3	0.6	1.0	11.6	0.1	5.9	0.0
0.010	1.7	0.0	0.3	0.2	0.0	0.0	0.0	4.840	12.4	0.6	1.0	10.4	0.1	5.9	0.0
0.040	3.4	0.0	0.4	0.3	0.0	0.0	0.0	5.290	11.5	0.6	1.0	9.5	0.1	5.9	0.0
0.090	4.7	0.1	0.5	0.7	0.0	0.0	0.0	5.760	7.5	0.6	1.5	8.4	0.1	4.6	0.0
0.165	6.2	0.1	0.6	1.7	0.0	0.0	0.0	6.250	7.0	0.8	1.5	7.3	0.1	4.4	0.0
0.250	7.1	0.2	0.7	2.7	0.0	0.1	0.0	7.840	5.3	0.8	1.3	5.0	0.1	3.9	0.0
0.365	8.0	0.2	0.7	4.5	0.0	0.1	0.0	9.000	4.4	0.8	1.2	3.9	0.2	3.4	0.0
0.490	8.6	0.2	0.7	6.0	0.0	0.2	0.0	10.240	3.6	0.8	1.1	3.4	0.2	3.0	0.0
0.640	9.1	0.3	0.7	8.1	0.0	0.3	0.0	11.560	2.9	0.7	1.1	2.3	0.2	2.5	0.0
0.810	9.4	0.3	0.7	9.6	0.0	0.4	0.0	12.960	2.4	0.7	1.0	1.7	0.2	2.2	0.0
1.000	9.5	0.3	0.7	11.4	0.0	0.6	0.0	14.440	1.9	0.7	0.9	1.3	0.2	1.8	0.0
1.440	23.8	0.5	1.1	17.5	0.0	2.3	0.0	16.000	1.5	0.6	0.9	0.9	0.2	1.6	0.0
2.250	20.9	0.6	1.1	18.0	0.1	3.4	0.0	20.250	0.8	0.6	0.7	0.4	0.2	1.2	0.0
2.560	19.2	0.6	1.1	17.7	0.1	4.0	0.0	25.000	0.4	0.6	0.7	0.2	0.2	0.8	0.0
2.890	17.6	0.6	1.1	16.8	0.1	4.6	0.0	30.250	0.1	0.6	0.6	0.1	0.2	0.7	0.0
3.240	16.2	0.6	1.1	15.3	0.1	5.2	0.0	36.000	0.0	0.5	0.5	0.1	0.2	0.6	0.0
3.610	15.1	0.6	1.1	14.2	0.1	5.7	0.0	42.250	0.0	0.4	0.4	0.1	0.2	0.5	0.0
4.000	14.1	0.6	1.1	13.1	0.1	5.8	0.0								

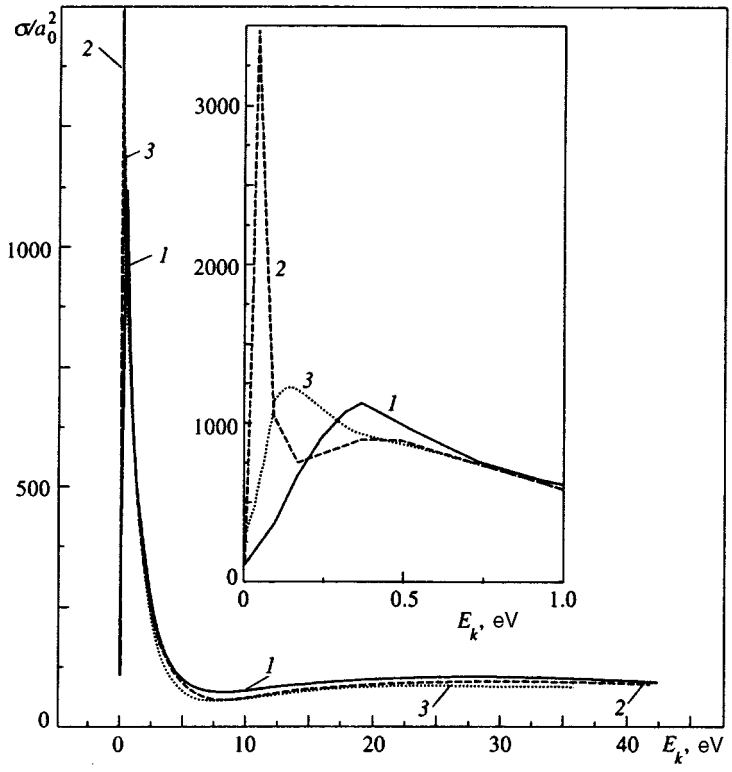

FIG. 3. Dependence of the total elastic scattering cross section for uranium σ (in atomic units) on the electron kinetic energy E_k with the unperturbed atomic potential $V_a(r)$ obtained from the exchange-correlation potential of Ref. 11 with the third set of parameters, the exchange potential $V_{\text{ex}}(r)$ of Ref. 15, and polarization potentials $V_p(r)$ taken from: Ref. 14 (curve *1*), Ref. 5 (curve *2*), and $V_p(r)\equiv 0$ (curve *3*).

parameters and the exchange potential $V_{\text{ex}}(r)$ of Ref. 15, and took the polarization potential from Ref. 14 and Ref. 5; we also calculated the scattering with no polarization potential. Figure 3 shows the results obtained when the two model polarization interactions are used. It is clear that they are in good agreement with one another for energies greater than 1 eV, but differ markedly at smaller energies. The cross section functions shown do not have a deep minimum at small kinetic energies, i.e., there is no Ramsauer effect.

In order to determine the effect of variations in the exchange potential we carried out a fourth series of calculations in which we used the unperturbed atomic potential $V_a(r)$ obtained from the exchange-correlation potential of Ref. 11 with the third set of parameters and the polarization potential $V_p(r)$ from Ref. 14, while varying the exchange potential $V_{\text{ex}}(r)$. In the first series of calculations, which we mentioned already, the exchange potential was taken from Ref. 15; the second set of calculations included no exchange potential at all; in the third we used the potential from Ref. 16; and in the fourth, the Slater potential (with $\alpha_{Slt}=2/3$). The results are shown in Fig. 4. For kinetic energies less than 5 eV, the behavior of the cross section is totally different from one calculation to the next. Clearly, the cross section is very sensitive to the form of the exchange interaction at these low kinetic energies. There is no Ramsauer effect for all four types of calculations.

In Ref. 18 we presented data from a numerical calculation of the dependence of the partial scattering phases on electron kinetic energy E_k in eV for uranium, using the relativistic phase function method for all the types of electron–atom interaction potentials chosen by us.

In this paper we calculate the degree of orthogonality of the wave functions of the scattered and atomic (bound) electrons. Consider first the orthogonality of the radial wave functions with the same quantum numbers j and l. The degree of orthogonality of the radial wave functions is determined by the overlap integral

$$I_b=\left|\frac{\int_0^{r_{njl}}[g_{Ejl}(r)g_{nlj}(r)+f_{Ejl}(r)f_{nlj}(r)]dr}{\sqrt{\int_0^{r_{njl}}[g_{Ejl}^2(r)+f_{Ejl}^2(r)]dr\int_0^{r_{njl}}[g_{njl}^2(r)+f_{njl}^2(r)]dr}}\right|, \tag{23}$$

where r_{njl} is the radius of normalization of the wave function of a bound electron, $g_{Ejl}(r)$ and $f_{Ejl}(r)$ are the large and small components of the continuous-spectrum wave function of the scattered electron, and $g_{njl}(r)$ and $f_{njl}(r)$ are the large and small components of the wave function of a bound atomic electron from Eq. (6). The radius of normalization r_{njl} is chosen from the condition

$$\int_0^{r_{njl}}[g_{njl}^2(r)+f_{njl}^2(r)]dr=1-\varepsilon, \tag{24}$$

where $\varepsilon=10^{-7}$. The calculations show that the value of I_b exceed 1% for electron from the outer shells of the atom with minimum binding energy. Table III shows how the overlap integral I_b depends on the kinetic energy E_k of an electron for uranium with the unperturbed potential $V_a(r)$ derived from the correlation-exchange term taken from Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ taken from Ref. 14, and the exchange potential $V_{\text{ex}}(r)$ taken from Ref. 15. The first column is for $s_{1/2}$ waves and $7s_{1/2}$ atomic electrons with binding energy $E_b=4.743$ eV; the second is for $p_{1/2}$, $6p_{1/2}$, $E_b=29.526$ eV respectively; the third

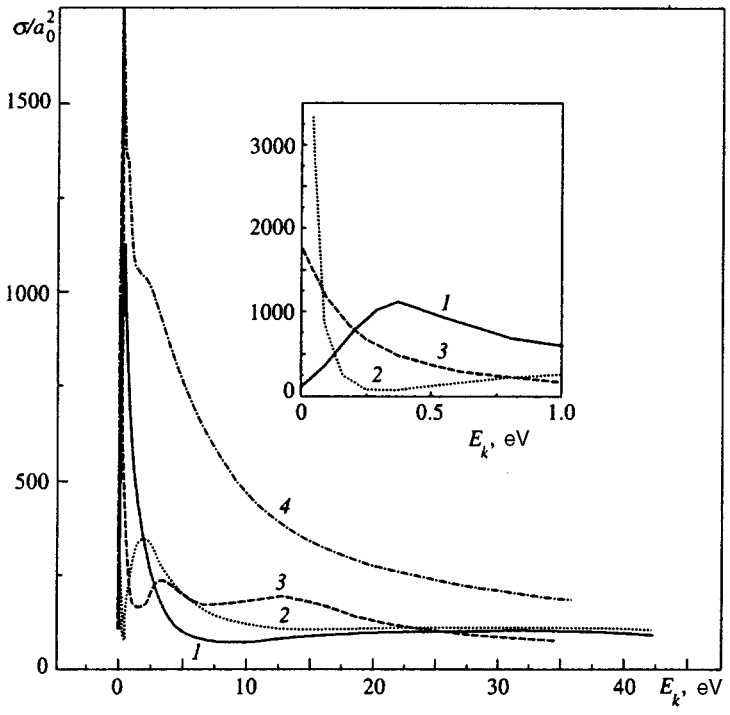

FIG. 4. Dependence of the total elastic scattering cross section for uranium σ (in atomic units) on the electron kinetic energy E_k with the unperturbed atomic potential $V_a(r)$ obtained from the exchange-correlation potential of Ref. 11 with the third set of parameters, the polarization potential $V_p(r)$ of Ref. 14 and exchange potentials $V_{\rm ex}(r)$ taken from: Ref. 15 (curve *1*); Ref. 16 (curve *2*); $V_{\rm ex}(r)\equiv 0$ (curve *3*); and the Slater potential of Ref. 10 and $\alpha_{Slt}=2/3$ (curve *4*).

for $p_{3/2}$, $6p_{3/2}$, $E_b=20.627$ eV; the fourth for $d_{3/2}$, $6d_{3/2}$, $E_b=2.776$ eV; the fifth for $d_{5/2}$, $5d_{5/2}$, $E_b=94.462$ eV; the sixth for $f_{5/2}$, $5f_{1/2}$, $E_b=3.225$ eV; and the seventh for $f_{7/2}$, $4f_{7/2}$, $E_b=365.490$ eV. Other types of the mean atomic, polarization, and exchange potentials give roughly the same maximum value of I_b, but at a different value of E_k.

It is possible that the models of elastic and exchange potentials we have used, which describe the elastic scattering of electrons by inert gas atoms and certain alkali metal atoms rather well, may be unsuitable for describing scattering by uranium. However, we can assert that at low energies the computed behavior of the total cross section for elastic scattering of an electron by uranium depends markedly on the model interaction potential. Note that it is necessary to use a relativistic model of the scattering of an electron by a heavy atom. The relativistic effects are especially evident in the *s*- and *p*-waves for $E_k<5$ eV. This fact prevents us from making any reliable predictions regarding the presence or absence of a Ramsauer effect in the scattering from average-potential models.

5. CONCLUSION

Reliable assertions about the shape and magnitude of the energy dependence of the cross section for elastic scattering of a slow electron by a uranium atom, and the presence or absence of a Ramsauer effect, can be made only with the help of experiments that refine the theoretical concepts used to describe the elastic scattering process. There are a number of simple model potentials that are in good agreement with one another, but only for energies of several tens of eV. More first-principles theoretical calculations would have to be made using the continuous-spectrum multiconfigurational relativistic Hartree–Fock method for the complete atom-plus-electron system; however, such calculations would require a great deal more computational power. We know of similar calculations carried out by Saha[19] for atoms of certain inert gases in the nonrelativistic approximation, which require tens of hours of machine time on a Cray-level supercomputer. In our paper Ref. 7, we compared the calculations of Ref. 19 for the elastic-scattering cross sections of argon atoms with relativistic and nonrelativistic single-particle calculations based on an effective spherically symmetric potential, and concluded that it is necessary to include relativistic corrections in describing the scattering at electron kinetic energies less than 5 eV.

The authors are grateful to V. Yu. Dobretsov for providing us with programs without which the numerical calculations would have been very difficult. Partial financial support for this work was obtained from a grant by the Russian ''Kurchatov Institute'' Science Center.

*) E-mail: x1064@kiae.su

1) The phase function $\delta_l(r)$ and amplitude function $A_l(r)$ derived from the Schroedinger equation depend only on the radial quantum number l, and at nonrelativistic energies we have $\delta_l\approx\delta_{l+1/2,l}\approx\delta_{l-1/2,l}$, $p\approx k$.

[1] L. P. Kudrin and A. A. Drozdov, Atomnaya Energiya **27**, 39 (1969) [in Russian].
[2] F. Calogero, *Variable Phase Approach to Potential Scattering* (Academic Press, New York, 1967; Russ transl. Mir, Moscow, 1972).
[3] V. V. Babikov, *The Method of Phase Functions in Quantum Mechanics (Nauka, Moscow, 1988).*
[4] P. Gombàs, *Theorie und Lösungsmethoden des Mehrteilchenproblems der Wellenmekanik (Many-Body Problems in Wave Mechanics: Theory and Solution Methods)* (Birkhäuser, Basel, 1950; Russ transl. IL, Moscow, 1952).
[5] D. R. Hartree, *The Calculation of Atomic Structures* (Wiley, New York, 1957; Russ. transl. IL, Moscow, 1960).
[6] D. P. Grechukhin and A. V. Lomonosov, Preprint NAE 5789/12 (1994).
[7] D. P. Grechukhin and A. V. Lomonosov, JETP Lett. **60**, 779 (1994).
[8] A. I. Akheizer and V. B. Berestetskiĭ, *Quantum Electrodynamics* (Nauka,

Moscow, 1969 [3rd edition]; Eng. transl. Wiley, New York, 1965).
[9] *Handbook of Special Functions*, M. I. Abramowitz and I. A. Stegun eds. (Dover, New York, 1965; Nauka, Moscow, 1969).
[10] J. C. Slater, *The Self-Consistent Field for Molecules and Solids* (McGraw-Hill, New York, 1974; Mir, Moscow, 1978).
[11] U. von Barth and L. Hedin, J. Phys. C **4**, 2064 (1971).
[12] J. P. Perdew and A. Zunger, Phys. Rev. B **23**, 5048 (1981).
[13] S. H. Volsko, L. Wijk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
[14] J. K. O'Connel and N. F. Lane, Phys. Rev. A **27**, 1893 (1983).
[15] J. Hara, J. Phys. Soc. Jpn. **22**, 710 (1967).
[16] M. E. Riley and D. G. Truhler, J. Chem. Phys. **63**, 2182 (1975).
[17] G. A. Korn and T. M. Korn, *Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1961; Nauka, Moscow, 1972).
[18] D. P. Grechukhin and A. V. Lomonosov, Preprint NAE 5992/12 (1996).
[19] H. P. Saha, Phys. Rev. A **43**, 4712 (1991).

Translated by Frank J. Crowne

Mechanisms for long-range forces in the ''three atoms+electron'' system

F. M. Pen'kov*)

Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia
(Submitted 10 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1229–1235 (April 1997)

The Born–Oppenheimer approximation is used to obtain an equation for the effective interaction of three atoms bound by a single electron. For low binding energies long-range forces arise between the atoms in an ''electron+atom pair'' that lead to bound states when the size of the three-atom cluster is several tens of angstroms. A system made up of alkali metal atoms is considered as an example. © *1997 American Institute of Physics.* [S1063-7761(97)00804-4]

1. INTRODUCTION

Nowadays, methods for creating matter in the cluster phase and determining its properties are widely investigated. Thus, a model was proposed by Manykin *et al.*[1,2] for matter consisting of highly excited atoms in which the stability is determined by the large overlap of wave functions from the valence electrons. Later this model was used to explain the cluster properties of a cold plasma of cesium atoms.[3] In this paper we propose a different mechanism for maintaining a very diffuse cluster based on the small affinity energy of an electron for a neutral atom.

Previously,[4] it was shown by direct analysis of the Faddeev equations[5] that the effective interaction potential of two atoms in a ''two atoms+electron'' system for real electron–atom affinity energy scales has long-range Efimov[6] and quasi-Coulombic terms proportional to $1/r^2$ and $1/r$ respectively, which determine the asymptotic part of the spectrum of these systems. In this case the quasiclassical limit for the effective potential coincides with the effective potential of Ref. 7 obtained from the Schrödinger equation in the Born–Oppenheimer approximation. The self-similarity of the effective potential at distances larger than the characteristic range of the pair forces makes it possible to use convenient pairwise separable Yamaguchi potentials for the analysis.

It turns out that in a ''three atoms+electron'' system the effective potential for interatomic interactions also has long-range parts with coupling constants that are larger than those of a ''two atoms+electron'' system. In what follows, the terms ''triatomic system'' and ''diatomic system'' are used for simplicity without mentioning the electron, whose presence here also determines the diatomic and triatomic effective potentials.

The effective potential of such a triatomic system is local, but the three-particle potential cannot be broken up into a sum of pairwise interactions. Three-particle forces are a serious problem for analyzing the spectrum of such a system, while the quasiclassical nature of the motion makes numerical calculations of the states of real atomic systems practically impossible. In addition, the range of energies of most interest is above the diatomic threshold, where two atoms of the three can form a bound state while the third particle lies in the continuum spectrum. Therefore, a complete analysis requires investigation not only of the spectrum but also the series of resonances in which the triatomic system may be found. In this paper a scheme is given for constructing the effective potential of the triatomic subsystem for a four-particle problem within the framework of the Born–Oppenheimer approximation. The properties of this system will be discussed only for those configurations of three atoms that are necessary to illustrate assertions about the long-range character of the effective potential and the quasiclassical nature of the spectrum.

2. EFFECTIVE POTENTIAL

In order to treat a four-particle system consisting of three atoms and an electron, we introduce Jacobi coordinates, which are convenient since the electron coordinate is invariant with respect to permutation of the atoms when they are used. Let $\mathbf{r}_\alpha$ and $\boldsymbol{\rho}_\alpha$ be the standard three-particle Jacobi coordinates of the atoms, i.e., $\mathbf{r}_\alpha$ joins a pair of atoms and $\boldsymbol{\rho}_\alpha$ is a vector directed from the center of mass of this pair to the third atom. The label $\alpha = 1, 2, 3$ identifies one of the three possible systems. Then the electron coordinate x will be a vector from the position of the electron to the center of mass of the three atoms.

In view of the smallness of the electron mass compared with the atomic masses, our scheme for constructing the effective potential in the Born–Oppenheimer approximation consists of finding the quantity $U(\mathbf{r},\boldsymbol{\rho})$ from the equation.[1)]

$$-\frac{1}{2\mu}\Delta_x\Psi+\sum_{i=1}^{3} v(x_i)\Psi=U(\mathbf{r},\rho)\Psi, \tag{1}$$

where μ is the electron mass and $v(x_i)$ is the pairwise interaction potential between an electron and the atom with label i. In what follows we will use rank-1 separable potentials that act only in an S-wave. The effect of such a potential on the wave function can be written in the form

$$v(t)\Psi=\nu(t)\langle\nu(t')\Psi(\mathbf{t}')\rangle,$$

where the angle brackets imply integration over all the space of the variable $\mathbf{t}'$.

The separable form of the pairwise potential defines a simple scheme for constructing equations for the potential U. In this case we must take a Fourier transform with respect to the variable $\mathbf{x}$. Then the coordinates $\boldsymbol{\rho}$ and $\mathbf{r}$, denoted below by a six-dimensional vector $\mathbf{R}$, drop out of the discussion. The Fourier transforms of the quantities $v(x_i)$ $\Psi(\mathbf{R},\mathbf{x})$ have the following simple form:

$$\exp(i\mathbf{k}\cdot\mathbf{a}_i)\,\nu(k)\langle\,\nu(k)\exp(-i\mathbf{k}\cdot\mathbf{a}_i)\Psi(\mathbf{R},\mathbf{k})\rangle,$$

where the vector $\mathbf{a}_i$ connects the position of the atom with label i and the center of mass of the triatomic system. In what follows we will express $\Psi(\mathbf{R},\mathbf{k})$ in terms of the unknown quantities in angle brackets. Then by projecting it onto the function $\nu(k)\exp(-i\mathbf{k}\mathbf{a}_i)$ for each i we obtain a homogeneous system of equations which is solved by setting its determinant equal to zero. In this way we obtain the equation

$$[G(0)-1]^3+2G_1G_2G_3-(G_1^2+G_2^2+G_3^2)[G(0)-1]=0 \tag{2}$$

for the effective potential U, which enters into the Green's functions $G_i \equiv G(r_i)$ averaged over momentum:

$$G(r)=\left\langle \nu(k)\,\frac{\exp(i\mathbf{k}\cdot\mathbf{r})}{U-k^2/2\mu}\,\nu(k)\right\rangle. \tag{3}$$

Here the vector $\mathbf{r}_i$ joins two atoms, i.e., it coincides with the coordinates $\mathbf{r}_\alpha$ for a corresponding choice of the set of Jacobi coordinates. For definiteness, a pair of particles will be labeled according to the label of the third particle.

Equation (2) with the definition (3) is valid for any separable potentials. In what follows we will consider the Yamaguchi potential, which has the form

$$\nu(k)\,\nu(k')=-\frac{4\pi}{\mu}\,\frac{(\kappa+\beta)^2\beta}{(\beta^2+k^2)(\beta^2+k'^2)},$$

the quantity κ is a wave number for a real ($\kappa>0$) or virtual ($\kappa<0$) state with energy $\varepsilon=-\kappa^2/2\mu$, while β determines the inverse range of the pair potential, and consequently its depth $\beta^2/2\mu$. Simple integration gives an expression for $G(r)$:

$$G(r)=\left(\frac{\beta+\kappa}{\beta+u}\right)^2\left(\frac{2\beta}{(\beta-u)^2}\,\frac{e^{-ru}-e^{-r\beta}}{r}-\frac{\beta+u}{\beta-u}\,e^{-r\beta}\right). \tag{4}$$

Here we define the wave number $u=\sqrt{-2\mu\,U}$ of the potential for $\mathrm{Re}\,u>0$. In what follows we require the value of the function (4) at zero:

$$G(0)=\left(\frac{\beta+\kappa}{\beta+u}\right)^2, \tag{5}$$

and in the limiting case $r\gg\beta^{-1}$:

$$G(r)\to\left(\frac{\beta+\kappa}{\beta+u}\right)^2\,\frac{2\beta}{(\beta-u)^2}\,\frac{e^{-ru}}{r}. \tag{6}$$

3. PROPERTIES OF THE EFFECTIVE POTENTIAL

Let us consider the effective potential at zero. Setting r_i equal to zero, we obtain from Eq. (2) the simple expression

$$3G(0)=1,$$

which shows that the wave number of the effective potential corresponds to three times the depth of the pairwise electron–atom interaction, since there is no factor of 3 in the eigenvalues of the pairwise problem for the electron-atomic system. Taking (5) into account, the depth of the triatom potential at zero is determined by the wave number $u_{\lim}$:

$$u_{\lim}=(\sqrt{3}-1)\beta+\sqrt{3}\kappa,$$

which reduces to zero only for a virtual value of the wave number of the electron–atom interaction. However, the effective potential of a triatomic system exists for $\kappa>-0.423\beta$.

The equations given above for the effective potential are valid for any separable interaction used to define the details of the pairwise force at small distances. Below we will treat the case of a resonant interaction, where the wave number of the bound state is much smaller than the inverse range of the pairwise forces. For the Yamaguchi potential the resonant nature is determined by the condition $\kappa/\beta\ll 1$. In this case the amplitude of the pairwise electron–atom interaction is determined by the pole nearest to the bound state (real or virtual), and does not depend on the details of the pairwise forces, i.e., on the choice of potential. For the effective potential of a triatomic system this independence will be found for $r_i\gg\beta^{-1}$.

Let us consider Eq. (2) in the resonant case. Taking (6) into account, it simplifies to an expression that does not depend on β:

$$1-2\,f_1f_2f_3-f_1^2-f_2^2-f_3^2=0, \tag{7}$$

where

$$f_i=\frac{e^{-\kappa r_i}e^{-c_i}}{c_i},\quad c_i=(u-\kappa)r_i\,.$$

The quantities c_i determine the behavior of the effective potential with coordinate r_i:

$$U=-\frac{c_i^2}{2\mu r_i^2}-\frac{c_i\kappa}{\mu r_i}+\varepsilon,$$

and have the sense of coupling constants which depend on the mutual atomic positions.

The form of Eq. (7) allows us to draw some preliminary conclusions about the behavior of the effective potential. First of all, the larger the quantity f_i is, the smaller the coupling constant corresponding to it. Secondly, the condition for solvability of Eq. (7) imposes restrictions on the values of f_i:

$$f_i^2+f_j^2\leqslant 1,\quad i\neq j,$$

for any pair of coordinates. Thus, the limiting value $f_i=1$, and consequently the minimum value for c_i is achieved only in the case where the remaining two components equal zero, corresponding to the configuration $\rho_\alpha\gg r_\alpha$ for any pair of atoms. This case corresponds to two isolated atoms whose effective interaction potential has clearly expressed long-range terms with coupling constant $c_{\lim}$. In particular, in the region $\kappa r_i\ll 1$ the limiting value of the coupling constant $c_{\lim}=0.567$ is a solution to the equation $c=\exp(-c)$. For any other configuration of three atoms we have $f_i<1$ and $c_i>c_{\lim}$. Thus, the pairwise diatomic potential is an upper limit on the triatomic interaction.

As in the diatomic case,[4] for $\kappa>0$ the asymptotic value of u reduces to κ, corresponding to real electron–atom cou-

pling. For $\kappa r_i \ll 1$ the functions f_i depend only on c_i, and therefore the behavior of the effective potential is determined by these variables. In the diatomic case this dependence immediately leads to long-range potentials of the Efimov and quasi-Coulombic type. In the triatomic case the situation is more complicated, since there are three independent coordinates that determine the triangle of particles.

Of course, for spatial configurations of three bodies specified by a single coordinate r we return to the case of long-range forces, as in the diatomic case,[4] but with different values of the constant c. Some examples of this are configurations with $\rho_i \to 0$, for which the atoms are located on the same line symmetrically with respect to the atom with label i, or the configuration of an equilateral triangle. In the first case, assuming for definiteness $\boldsymbol{\rho}_1 = 0$ we obtain $r_2 = r_3 = r_1/2$. Then Eq. (7) gives an equation for c:

$$c^2 = (8+c)e^{-c},$$

with the unique solution $c = 1.473\ldots$

For the equilateral triangle configuration ($r_1 = r_2 = r_3 = r$) Eq. (7) reduces to the form

$$c = 2e^{-c}, \tag{8}$$

which differs from the diatomic case[4] by a factor of 2 and therefore gives a larger coupling constant $c = 0.8526\ldots$.

It is not difficult to show that this configuration of atoms is unstable, i.e., any deviation from the configuration of an equilateral triangle gives a deeper effective potential. In order to verify this, let us consider a variation δu which distorts the configuration for a fixed moment of inertia, i.e., $R^2 = (r_1^2 + r_2^2 + r_3^2)/3$. We note that constancy of the moment of inertia mR^2 implies independence of the kinetic energy with respect to such variations.

Varying Eq. (2) with respect to u and r_i^2, taking into account the constancy of R and the sign of the derivatives $\partial G/\partial u$, we obtain

$$\delta u = M\left[\frac{3}{2}\,(\delta_1^2+\delta_2^2)+(\delta_1+\delta_2)^2\right], \quad \delta_i = \delta r_i^2, \quad M>0.$$

Thus, the wave number reaches its minimum and the effective potential its maximum value in the ''equilateral triangle'' atomic configuration. Hence using this maximum value we can obtain a lower bound on the number of levels and draw conclusions about the long-range character for arbitrary atomic positions.

The equation of motion for three atoms contains six independent coordinates. It is convenient to choose hyperspherical coordinates (see, e.g., the review Ref. 8). Since it is sufficient to investigate only S-wave motion in each pair in order to confirm the long-range character of the effective potential, the only variables in the Schrödinger equation will be the hyperradius R and the hyperangle χ:

$$R^2 = \frac{2}{3}\,\rho^2 + \frac{1}{2}\,r^2, \quad \tan\chi = \sqrt{\frac{3}{4}}\,\frac{r}{\rho}.$$

In these variables the motion of the atoms is described by a Schroedinger equation, which for the wave function $\Phi(R,\chi)$, after separating out the standard factor $1/r\rho$, takes the form

$$-\frac{1}{2m}\left(\frac{\partial^2}{\partial R^2}\,\Phi + \frac{1}{R}\,\frac{\partial}{\partial R}\,\Phi + \frac{1}{R^2}\,\frac{\partial^2}{\partial \chi^2}\,\Phi\right) + U(R,\chi)\Phi = E\Phi, \tag{9}$$

where m is the mass of an atom. In this case the function Φ can take on zero values for $R = 0$, $\chi = 0$, $\chi = \pi/2$; since we are discussing the problem as an eigenvalue problem, Φ should be quadratically integrable.

The effective potentials under discussion have the scale of the potential energy for an electron interacting with an atom, i.e., they consist of fractions of an electron volt, whereas the de Broglie kinetic energy of the atoms on the atomic scale of distances amounts to only a few thousandths of an electron volt. Thus, the motion of the atoms is essentially quasiclassical. In order to verify this, let us consider the solutions to Eq. (9) with maximum value of the effective potential determined by the equilateral-triangle configuration. Since this value is independent of χ, the variables separate and the solution can be written as a sum of functions $A_K(R)B_K(\chi)$; the functions $B_K(\chi)$, which are found by solving the eigenvalue problem, are required to vanish for $\chi = 0$ and $\chi = \pi/2$. The solutions are obvious and give the eigenvalues $-4K^2$ ($K = 1,2,3$), leading to an additional repulsive potential of the form $4K^2/2mR^2$. The equation for $A(R)$ is quasiclassical. Actually, introducing the constant $\gamma = \sqrt{m/\mu}$ and taking into account that $\gamma \gg 1$, we can look for an expansion of the action S with respect to γ:

$$S = \gamma s_0 + s_1 + \ldots\ .$$

Substituting $A(R)$ in the form $\exp(iS)$ into Eq. (9), we obtain the usual quasiclassical solution with Bohr–Sommerfeld quantization rules in the form

$$\gamma\int_{R_{\min}}^{R_{\max}} \sqrt{u^2 + 2\mu E - \frac{4K^2-0.25}{\gamma^2 R^2}}\; dR = \pi\left(n+\frac{3}{4}\right). \tag{10}$$

The latter condition can be used to estimate the number of levels if we set the energy equal to the electron–atom binding energy or zero for the virtual states. In what follows we will sum over all K, which gives a number $n > 0$. Since the effective potential at small distances depends on the model chosen both for the electron–atom and the atom–atom pairwise interaction, it is reasonable to choose the bound $R_{\min}$ in the range greater than the characteristic scales of these interactions. This choice of the left-hand boundary corresponds to a model with an additional interatomic repulsion at small distances. By increasing $R_{\min}$ we can impose the vanishing of the bound state. This boundary also will define the real scale of the long-distance effective potential of the equilateral-triangle configuration.

In order to demonstrate these assertions, it is good to consider a number of alkali metals with rather small electron-affinity binding energies for an atom in a S-state. For ^{6}Li and ^{133}Cs we estimate affinity energies of 0.62 and 0.47 eV respectively.[9] We calculate our estimates of the number of levels over a region larger than 5 Bohr radii, i.e., 2.65 Å. In this case we use the asymptotic equation (7). Then the system $^6Li_3^-$ has more than 1.6×10^3 levels and disappears when the repulsion radius is greater than 20 Å. The

system $^{133}Cs_3^-$ has more than 4.0×10^4 levels and disappears when the repulsion radius is greater than 34 Å. Since these calculations are for configurations with very shallow effective potentials, these estimates are lower bounds.

4. CONCLUSION

The mechanisms for quantum-mechanical long-range behavior in systems of heavy particles with almost classical dynamics do not change when we go from a diatomic problem[4] to a triatomic one, and are explained by the large characteristic size of the electronic shell for the added electron. This physics of the long-range behavior should also be present in systems with larger numbers of atoms. Such clusters should exhibit certain special features: negative charge, a large average distance between the atoms, and for atoms with unfilled shells of valence electrons magnetic properties as well, connected with a tendency to align their spins antiparallel to the spin of the external electron that binds the system. In addition to atoms with rather small real electron affinities, atoms with small virtual affinities can also be candidates for the formation of such clusters. In this paper we have considered the effective interaction of three ground-state atoms mediated by an external electron, but the mechanism for such an interaction is an intrinsic property of the interatomic degrees of freedom and does not exclude excited atoms. In this case additional long-range attractive potentials may arise between the atoms in such a system.

This work was carried out under Project K-40-96.

*)E-mail: penkov@thsun1.jinr.dubna.su

1)In this paper we will set Planck's constant $\hbar = 1$.

[1] E. A. Manykin, M. I. Ozhovan, and P. P. Poluektov, Zh. Éksp. Teor. Fiz. **84**, 442 (1983) [Sov. Phys. JETP **57**, 256 (1983)].
[2] E. A. Manykin, M. I. Ozhovan, and P. P. Poluektov, Ukr. Phys. J **34**, 146 (1989).
[3] R. Svenson, L. Holmid, and L. Lundgren, J. Appl. Phys. **70**, 1489 (1991).
[4] F. M. Pen'kov, Zh. Éksp. Teor. Fiz. **106**, 1046 (1994) [JETP **79**, 568 (1994)].
[5] L. D. Faddeev, Zh. Éksp. Teor. Fiz. **39**, 1459 (1960) [Sov. Phys. JETP **12**, 1014 (1960)].
[6] V. Efimov, Yad. Fiz. **12**, 1080 (1970) [Sov. J. Nuclear Phys. **12**, 589 (1971)].
[7] A. C. Fonseca, E. F. Redish, and P. E. Shanley, Nucl. Phys. A **320**, 273 (1979).
[8] R. T. Pack and A. Parker, J. Chem. Phys. **87**, 3888 (1897).
[9] A. A. Radtsig and B. M. Smirnov, *Parameters of Atoms and Atomic Nuclei* (Energoatomizdat, Moscow, 1986), p. 116 [in Russian].

Translated by Frank J. Crowne

Coherent repopulation of hyperfine structure levels in the field of a bichromatic resonant radio-frequency wave

D. F. Zaretskiĭ and S. B. Sazonov

Kurchatov Institute, 123182 Moscow, Russia
(Submitted 11 October 1996)
Zh. Éksp. Teor. Fiz. **111**, 1236–1244 (April 1997)

We consider the resonant interaction between atoms with hyperfine energy levels and a bichromatic radio-frequency field. Nuclear Zeeman levels of an impurity center in a magnetic host form a structure of this kind. Using the spin-density-matrix formalism, we solve the problem of coherent repopulation of a system of three of these levels under the action of a bichromatic resonant radio-frequency wave, taking into account transverse relaxation, and note the connection between this effect and the well-known phenomenon of coherent population capture when a laser bichromatic field interacts resonantly with a three-level system. We discuss various possibilities for observation of this effect experimentally. © *1997 American Institute of Physics.* [S1063-7761(97)00904-9]

1. INTRODUCTION

The resonant interaction of laser fields with atoms can lead to significant repopulation of levels of these atoms. In particular, this effect is observed when a bichromatic laser wave interacts with atoms with hyperfine structure.[1] This repopulation can be accompanied by the phenomenon of coherent population capture.[1–3] In optics this effect is used to obtain amplification without inversion, the creation of anomalously transparent medium, etc.[4]

It is reasonable to assume that an analogous effect should occur when systems of atomic hyperfine levels resonant interact resonantly with bichromatic radio-frequency waves. An example of such a system could be nuclear Zeeman levels of impurity centers in the magnetic field of a host. In some cases, the Zeeman nuclear levels are not equidistant due to the quadrupole electric interaction of the nuclei with the crystal field of the host.[5]

The primary difference between RF resonance and optical resonance is the fact that, as a rule, all the Zeeman levels are populated from the very beginning in the RF case. Populations of these levels less than unity are observed only at ultra-low temperatures. Furthermore, both longitudinal and transverse spin relaxations play a significant role in RF resonance. It is noteworthy that in the NMR region, as a rule the transverse relaxation time T_2 is considerably shorter than the longitudinal relaxation time T_1 (for spin depolarization).

One of the commonest methods for recording repopulation of the Zeeman nuclear levels of radioactive impurity centers is to measure the angular anisotropy of their decay products.[6]

2. COHERENT REPOPULATION OF LEVELS FOR A THREE-LEVEL SYSTEM WITH TRANSVERSE RELAXATION

Let us consider the coherent interaction of a three-level nuclear system with a resonant bichromatic radio-frequency field, in which the first component of the bichromatic wave induces resonance transitions between levels 1 and 3, while the second component induces resonance transitions between levels 2 and 3. Level 3 is common to both. Let us assume that all initial populations are nonzero. The values of the populations as functions of time are found by solving the system of equations for the spin-density matrix. In these equations we will include only transverse relaxation (relaxation time T_2). For impurity centers this approximation is valid, because, as a rule, $T_2 \ll T_1$ (where T_1 is the spin–relaxation time). However, this limitation is not fundamental. In the equations for the density matrix, we can introduce terms associated with T_1, but then the solution will be more complicated.

In addition, we will use the following approximation.

1. Each frequency of the bichromatic wave coincides with the frequency for a transition between the corresponding levels (the resonant approximation).

2. Throughout the entire interaction, the relative phase constant of the two components of the wave $\Delta\varphi$ remains fixed. For simplicity we consider only the cases $\Delta\varphi=0$ and $\Delta\varphi=\pi$.

3. The matrix elements for the magnetic dipole interaction of the nuclear spins with the magnetic field of the ith component of the bichromatic wave V_i $(i=1,2)$, which stimulate the corresponding transition, are assumed to be equal in absolute value ($V_2=V_1$ for $\Delta\varphi=0$ and $V_2=-V_1$ for $\Delta\varphi=\pi$).

4. The three-level system consists of nonequidistant levels.

There are several different ways that a three-level system can interact with a resonant field, depending on where the common third level is located compared to the other two. We will consider a system of three nonequidistant Zeeman levels for which the common third level, which participates in the interaction with both components of the resonant bichromatic field, is located between levels 1 and 2. (The analogous scheme in optics is referred to as the Ξ scheme.)

The interaction of the three-level system with the resonance field is described by an equation for the spin density matrix σ_{ij}, which is the best way to include relaxation processes. In our case, this equation has the form

$$\dot{\sigma}+\hat{\Gamma}\sigma=-i[(\hat{H}_0+\hat{V}_{\rm int}),\sigma]. \tag{1}$$

 1063-7761/97/040682-05$10.00

Here $\hat{H}_0$ is the Hamiltonian of the three-level system, $\hat{V}_{\rm int}$ is the interaction operator between the system and the field, and $\hat{\Gamma}$ is an operator that describes the transverse relaxation process; here $\hbar=1$. Following Ref. 7, we convert from Eq. (1) for σ to an equation for the matrix ρ in the interaction picture, using the relations

$$\rho_{ij}=\sigma_{ij}\exp(i(E_j-E_i)t),$$
$$V_{ij}=(\hat{V}_{\rm int})_{ij}\exp(i(E_j-E_i)t). \tag{2}$$

Eventually we obtain the following ninth-order system of equations for ρ_{ij} in the resonance approximation for $\Delta\varphi=0,\ \pi$, for which the $(\hat{V}_{\rm int})_{ij}$ are real:

$$\begin{aligned}
&\dot{\rho}_{11}=-iV_1(\rho_{31}-\rho_{13}),\\
&\dot{\rho}_{22}=-iV_2(\rho_{32}-\rho_{23}),\\
&\dot{\rho}_{33}=-iV_1(\rho_{13}-\rho_{31})-iV_2(\rho_{23}-\rho_{32}),\\
&\dot{\rho}_{13}+\Gamma\rho_{13}=-iV_1(\rho_{33}-\rho_{11})+iV_2\rho_{12},\\
&\dot{\rho}_{31}+\Gamma\rho_{31}=-iV_1(\rho_{11}-\rho_{33})-iV_2\rho_{21},\\
&\dot{\rho}_{23}+\Gamma\rho_{23}=-iV_2(\rho_{33}-\rho_{22})+iV_1\rho_{21},\\
&\dot{\rho}_{32}+\Gamma\rho_{32}=-iV_2(\rho_{22}-\rho_{33})-iV_1\rho_{12},\\
&\dot{\rho}_{12}+\Gamma\rho_{12}=-iV_1\rho_{32}+iV_2\rho_{13},\\
&\dot{\rho}_{21}+\Gamma\rho_{21}=-iV_2\rho_{31}+iV_1\rho_{23}.
\end{aligned} \tag{3}$$

Here Γ is the transverse relaxation width: $\Gamma=T_2^{-1}$. In order to simplify this system of equations (3) we introduce the quantities

$$\chi_{13}=\rho_{31}-\rho_{13},\quad \chi_{23}=\rho_{32}-\rho_{23},\quad \chi_{12}=\rho_{12}+\rho_{21},$$

and also use the normalization condition

$$\rho_{11}+\rho_{22}+\rho_{33}=3A.$$

As a result, we go from a system of equations of ninth order to a system of equations of fifth order:

$$\begin{aligned}
&\dot{\rho}_{11}=-iV_1\chi_{13},\\
&\dot{\rho}_{22}=-iV_2\chi_{23},\\
&\dot{\chi}_{13}+\Gamma\chi_{13}=-iV_1(4\rho_{11}+2\rho_{22}-6A)-iV_2\chi_{12},\\
&\dot{\chi}_{23}+\Gamma\chi_{23}=-iV_2(4\rho_{22}+2\rho_{11}-6A)-iV_1\chi_{12},\\
&\dot{\chi}_{12}+\Gamma\chi_{12}=-iV_2\chi_{13}-iV_1\chi_{23}.
\end{aligned} \tag{4}$$

The characteristic equation of this system (4) is an irreducible algebraic equation of fifth order, whose solution is difficult to find. We can solve this equation approximately, along with the system of equations (4), by assuming that the magnitude of the interaction energy for the three-level system with the bichromatic field V is much larger than the transverse relaxation width Γ. In this case, the period of the Rabi oscillations will be much smaller than the time T_2. As our zeroth approximation we take Eq. (4) with no relaxation: $\Gamma=0$. In this case the characteristic equation for the energy eigenvalues has the form

$$k^5+10V^2k^3+16V^4k=0. \tag{5}$$

Its solutions are the values

$$\begin{aligned}
&k_1=0,\quad k_2=i\sqrt{2}V,\quad k_3=-i\sqrt{2}V,\quad k_4=2i\sqrt{2}V,\\
&k_5=-2i\sqrt{2}V.
\end{aligned} \tag{6}$$

Using (6), we obtain for the solutions to the characteristic equation of system (4) to accuracy up to first order of smallness in the small parameter $\xi=\Gamma/V$

$$\begin{aligned}
&k_1=-3\Gamma/4,\quad k_2=i\sqrt{2}V-\Gamma/2,\quad k_3=-i\sqrt{2}V-\Gamma/2,\\
&k_4=2i\sqrt{2}V-5\Gamma/8,\quad k_5=-2i\sqrt{2}V-5\Gamma/8.
\end{aligned} \tag{7}$$

Let us assume that all the levels of the impurity center are populated at time $t=0$, with populations unequal to unity in general: the diagonal matrix elements of the density are $\rho_{ii}(0)=A_i$. The initial conditions for the off-diagonal density matrix elements require some discussion. We can assume that the rf field is turned on instantaneously. In this case the phases of the off-diagonal matrix elements at time $t=0$ will be fixed:

$$\rho_{ij}(0)=\sqrt{A_iA_j}\exp(i(\alpha_i-\alpha_j)),$$

where α_i are phase constants. Turning the field on instantaneously implies that the turn-on time is much shorter than the time between interactions that change the phases of the amplitudes of the level populations of the impurity center. Among the processes that change the phase are interactions of an impurity center with phonons, magnons, etc. Because this time is the same order of magnitude as T_2, the turn-on time should be much shorter than T_2. Then, taking into account that the wave functions in quantum mechanics are determined only up to a phase factor, the dependence on the time-independent phase constants α_i can be eliminated from the system (4).

When the field is turned on adiabatically over a time $\gg T_2$, the phases of the off-diagonal matrix elements can cycle many times during the turn-on time, and the initial conditions in this case should have the form $\rho_{ij}(0)=0\,(i\neq j)$. On the other hand, in order to bring about a Rabi-oscillation regime, the quantum system should interact with the field coherently over times that do not exceed T_2, i.e., the period of the Rabi oscillations should be much smaller than T_2. Then, according to (7), it is necessary that the field be sufficiently strong: $V\gg\Gamma$.

The solution to the system of Eqs. (4) to first-order accuracy in the parameter ξ for instantaneous turn-on of the strong field for $\Delta\varphi=\pi$ has the form

$$\begin{aligned}
\rho_{11}(t)=A+&[(A_1+A_2+2\sqrt{A_1A_2}-2A)/8]\exp(-3\Gamma t/4)\\
&+\{\{128(A_1-A_2)\cos(\sqrt{2}Vt)+4\sqrt{2}\xi(11A_1\\
&-5A_2+6\sqrt{A_1A_2})\sin(\sqrt{2}Vt)+[96(A_1+A_2\\
&-2A)-64\sqrt{A_1A_2}]\cos(2\sqrt{2}Vt)+\xi[15\sqrt{2}(A_1\\
&+A_2)-42\sqrt{2}A-10\sqrt{2A_1A_2}]\sin(2\sqrt{2}Vt)\}/\\
&256\}\exp(-\Gamma t/2).
\end{aligned} \tag{8}$$

Analogously, for $\Delta\varphi=0$ we obtain

$$\rho_{11}(t)=A+[(A_1+A_2-2\sqrt{A_1A_2}-2A)/8]\exp(-3\Gamma t/4)$$
$$+\{\{128(A_1-A_2)\cos(\sqrt{2}Vt)+4\sqrt{2}\xi(11A_1$$
$$-5A_2-6\sqrt{A_1A_2})\sin(\sqrt{2}Vt)+[96(A_1+A_2$$
$$-2A)+64\sqrt{A_1A_2}]\cos(2\sqrt{2}Vt)+\xi[15\sqrt{2}(A_1$$
$$+A_2)-42\sqrt{2}A+10\sqrt{2A_1A_2}]\sin(2\sqrt{2}Vt)\}/$$
$$256\}\exp(-\Gamma t/2). \quad (9)$$

For $\rho_{22}(t)$ we obtain expressions analogous to (8) and (9) with the difference that terms proportional to $\cos(\sqrt{2}Vt)$ and $\sin(\sqrt{2}Vt)$ have opposite signs. The expressions for $\rho_{33}(t)$ are easy to obtain from the normalization condition

$$\rho_{11}(t)+\rho_{22}(t)+\rho_{33}(t)=3A,$$

where A is the average initial population of the levels:

$$A=(A_1+A_2+A_3)/3.$$

From these expressions it follows that the magnitude of the population depends significantly on the phase difference $\Delta\varphi$ of the components of the bichromatic wave. In particular, if $A_1=A_2=A_3=A$ and the duration of the pulse τ satisfies the condition $V^{-1}\ll\tau\ll\Gamma^{-1}$, then for the quantity $\overline{\rho_{11}}$, which in this case equals $\overline{\rho_{22}}$ (the bar above the matrix element implies time-averaging away the Rabi oscillations),

$$\overline{\rho_{11}}=3A/4+(3A/16)(\tau/T_2) \quad \text{for } \Delta\varphi=0,$$
$$\overline{\rho_{11}}=5A/4-(3A/16)(\tau/T_2) \quad \text{for } \Delta\varphi=\pi. \quad (10)$$

Hence it is clear that for a short field pulse ($\tau/T_2\ll 1$) the populations of levels 1 and 2 will differ significantly from the population of level 3 ($\approx 25\%$) even when the populations of all levels are the same before the field is turned on. Furthermore, this difference will change for different values of $\Delta\varphi$. The reason for this effect is the coherent addition of population amplitudes at the common level 3, i.e., it is analogous to the phenomenon of coherent population capture in a three-level atomic system interacting resonantly with a bichromatic laser field. The repopulation of NMR levels at impurity centers should be observable if these centers are radioactive by examining the angular anisotropy of the decay products.

Since the populations of the levels depend significantly on the transverse relaxation time T_2, this effect affords an additional way to measure this quantity.

Let us now briefly discuss the case of adiabatic turn-on of the strong field. For this we obtain the solution to system (4) for zero initial conditions of the nondiagonal density matrix elements. In this case, ρ_{11} and ρ_{22} are given by the same expressions, which do not depend on the phase difference of the components of the bichromatic field. To second order in the small parameter ξ we have

$$\rho_{11}(t)=A-(3A\xi^2/512)\exp(-3\Gamma t/4)+[3\sqrt{2}A\xi$$
$$\times\sin(\sqrt{2}Vt)/32-3\sqrt{2}A\xi\sin(2\sqrt{2}Vt)/64$$
$$+3A\xi^2\cos(2\sqrt{2}Vt)/512]\exp(-\Gamma t/2). \quad (11)$$

As is clear from (11), no significant repopulation of the levels takes place in this case; however, small-amplitude Rabi oscillations appear in the system. The amplitude of these oscillations is proportional to T_2^{-1}.

3. RAPID REPOPULATION OF A THREE-LEVEL NUCLEAR SYSTEM USING PULSED RESONANCE FIELDS

As we showed above, a resonance bichromatic radio-frequency strong field can bring about significant repopulation of the nuclear levels. Furthermore, the degree of repopulation can be increased substantially using a simple procedure: to produce a major alteration of the level populations, we combine the resonant interaction with the bichromatic field with a turn-off of one of the components of this field. In this case, the single component of the resonant field that is not turned off will equalize the level populations which it couples. When both components of the bichromatic field are turned on once more, the coherent repopulation of all the levels of the three-level system begins again, but now with different initial conditions.

Let us investigate the behavior of the three-level system in these pulsed fields in detail, for simplicity neglecting relaxation and describing the quantum mechanical system using the Schrödinger equation for the state amplitudes. We emphasize that the formalism of wave functions can be used when relaxation effects are negligible. As follows from the results of the previous section, this approximation is valid for pulsed bichromatic fields if the pulse is shorter than all the relaxation times.

The wave function of the three-level system has the form

$$\Psi(t)=a_1(t)\Psi_1+a_2(t)\Psi_2+a_3(t)\Psi_3,$$

where Ψ_i are eigenfunctions of the Hamiltonian $\hat{H}_0$. The time-dependent perturbation theory equations for the amplitudes $a_i(t)$ in the resonance approximation will take the form

$$\dot{a}_1=-iV_1a_3, \quad \dot{a}_2=-iV_2a_3, \quad \dot{a}_3=-iV_1a_1-iV_2a_2, \quad (12)$$

here, as in (3), V_i is the dipole matrix element for the transition between the corresponding ith level and level 3 under the action of the field component with frequency ω_i and phase φ_i that is resonant with this transition. The solution to (12) for the initial conditions $a_i(0)=\sqrt{A_i}$ has the form

$$a_1(t)=(V_2A_-+V_1A_+\cos(\Omega t))/\Omega^2$$
$$-i\sqrt{A_3}V_1\sin(\Omega t)/\Omega,$$
$$a_2(t)=(V_2A'_+\cos(\Omega t)-V_1A'_-)/\Omega^2$$
$$-i\sqrt{A_3}V_2\exp(i\Delta\varphi)\sin(\Omega t)/\Omega,$$
$$a_3(t)=-\sqrt{A_3}\cos(\Omega t)-iA_+\exp(-i\varphi_1)\sin(\Omega t)/\Omega,$$
$$A_+=\sqrt{A_2}V_2\exp(-i\Delta\varphi)+\sqrt{A_1}V_1,$$
$$A_-=\sqrt{A_1}V_2-\sqrt{A_2}V_1\exp(-i\Delta\varphi), \quad (13)$$
$$A'_+=\sqrt{A_2}V_2+\sqrt{A_1}V_1\exp(i\Delta\varphi),$$

$$A'_- = \sqrt{A_1} V_2 \exp(i\Delta\varphi) - \sqrt{A_2} V_1,$$
$$\Omega^2 = V_1^2 + V_2^2.$$

From (13) it is clear that in a system of nuclear levels that are all populated at time $t=0$, total coherent population capture, i.e., vanishing population of level 3 at all times, is impossible. For this it is necessary that $A_3=0$ at the very least. However, it is clear from (13) that for certain relationships between the system parameters the interaction with the resonant bichromatic field can bring about a drastic readjustment of the populations in our case as well, even when A_3 is not equal to zero. The magnitude of the repopulation depends both on the values of the initial level populations and on the ratio of phases and intensities of the components of the bichromatic rf field.

Let us consider two characteristic situations. Suppose that $A_+=0$ and $A_3\neq 0$. Then A_+ can become zero for $\Delta\varphi = \pi$ and

$$V_1 = V_2\sqrt{A_2/A_1}. \tag{14}$$

In order to estimate the values of the level repopulations, let us convert from the amplitudes $a_i(t)$ to the populations $\rho_{ii} = \overline{|a_i|^2}$ averaged over the Rabi oscillations. From (13) we obtain

$$\begin{aligned}
\overline{\rho_{11}} &= V_2^2|A_-|^2/\Omega^4 + V_1^2|A_+|^2/2\Omega^4 + V_1^2|A_3|^2/2\Omega^2,\\
\overline{\rho_{22}} &= V_1^2|A_-|^2/\Omega^4 + V_2^2|A_+|^2/2\Omega^4 + V_2^2|A_3|^2/2\Omega^2,\\
\overline{\rho_{33}} &= |A_+|^2/2\Omega^2 + |A_3|^2/2.
\end{aligned} \tag{15}$$

From this we obtain the following relations for the values of the average populations in the case under discussion when $\Delta\varphi=\pi$ and when (14) holds:

$$\begin{aligned}
\overline{\rho_{11}} &= A_1 + A_2A_3/2(A_1+A_2),\\
\overline{\rho_{22}} &= A_2 + A_1A_3/2(A_1+A_2),\\
\overline{\rho_{33}} &= A_3/2.
\end{aligned} \tag{16}$$

Assume that initially all the level populations are the same: $A_1=A_2=A_3=1$. Then after the bichromatic field pulse acts the population of level 3 decreases by a factor of 2, while the populations of the other levels are increased to 5/4. Let us assume that the field that couples levels 1 and 3 is turned off. Due to the action of the remaining components of the field on the system, the populations of levels 2 and 3 are comparable and become equal to 7/8. Let us once again turn on both components of the field, changing their intensities according to (14). After the second bichromatic field pulse, according to (16) the following level populations are established:

$$\overline{\rho_{11}} = 389/272, \quad \overline{\rho_{22}} = 308/272, \quad \overline{\rho_{33}} = 119/272.$$

It is clear that the emptying of the third level will be larger in magnitude than it is after the first pulse. Thus, by repeatedly turning on and turning off the field many times, we can rapidly induce a significant decrease in the magnitude of the population of level 3 compared to the other levels.

Consider a different situation: $A_-=0$ and $A_3\neq 0$. These relations can hold for $\Delta\varphi=0$ and

$$V_1 = V_2\sqrt{A_1/A_2}. \tag{17}$$

The expressions for the average populations in this case will have the form

$$\begin{aligned}
\overline{\rho_{11}} &= A_1(A_1+A_2+A_3)/2(A_1+A_2),\\
\overline{\rho_{22}} &= A_2(A_1+A_2+A_3)/2(A_1+A_2),\\
\overline{\rho_{33}} &= (A_1+A_2+A_3)/2.
\end{aligned} \tag{18}$$

According to (18), as a result of the action of the bichromatic field the population of level 3 increases to 3/2, while the populations of levels 1 and 2 decrease to 3/4. Let us turn on the component that couples levels 1 and 3. Then, under the action of the other field components the populations of 2 and 3 both become 9/8. Repeatedly turning on both components of the bichromatic field with changes in their intensities according to Eq. (17) leads according to (18) to the following level populations:

$$\overline{\rho_{11}} = 6/10, \quad \overline{\rho_{22}} = 9/10, \quad \overline{\rho_{33}} = 15/10.$$

In this case, conversely, the populations of both levels 1 and 2 can be decreased compared to the population of level 3. Thus, it is clear that by repeatedly turning the bichromatic resonance field and its components on and off, we can rapidly make large changes in the pattern of populations of the three-level system.

4. CONCLUSIONS

The results we have obtained show that coherent repopulation of the levels of hyperfine structure in the field of a bichromatic resonance radio-frequency wave is entirely realistic for observing the effect. However, for this we require a number of conditions to be fulfilled.

1. The times for turning the bichromatic field on and off should be considerably shorter than all the relaxation times.

2. During the field pulse τ it is necessary to maintain a regime of Rabi oscillations, i.e., it is necessary for the condition $V_\tau \gg 1$ to hold. This implies that the field should be strong.

3. The field pulse τ should be shorter than both relaxation times (T_1, T_2). If $\tau > T_1, T_2$, then it is clear from Eqs. (8) and (9) that the populations of all the levels will become equal with time.

4. For a system of Zeeman nuclear levels, the energy of interaction with the field should not exceed the fixed amount by which they are separated, i.e., $|\Delta E_{13} - \Delta E_{23}| > V$, where $\Delta E_{ij} = E_i - E_j$ and E_i are the energy levels.

5. It is particularly important to address the question of how these processes are affected by inhomogeneous broadening. Due to the bichromatic character of the resonance field, inhomogeneous broadening will not significantly alter the way it affects the three-level system if the broadening is the same for all the levels. For the case of coherent population capture, as was shown in Ref. 3, a frequency detuning of the field from resonance that is the same for all the transitions will not interfere with the effect. However, the inhomogeneous broadening cannot be too large, because a criterion

for the system to be treated as three-level is that the transition frequency difference $|\Delta E_{13}-\Delta E_{23}|$ must be larger than the value of the inhomogeneous broadening.

6. In our calculations we set the widths of the transverse relaxation for all transitions equal. This may not be so in practice, although the widths can hardly differ in order of magnitude. For the pulse processes under discussion, whose durations are smaller than all the relaxation times, the assumption of identical relaxation widths is not fundamental. On the other hand, introducing widths that are different in magnitude would give rise to additional complications in the final result.

7. We have assumed that heating of the medium by the rf field does not lead to a significant increase in the level broadening. It is, of course, desirable that the size of the sample be less than the skin-depth thickness. In this case, as was shown in the papers by Chaplin,[5] heating of the medium is insignificant even at ultra-low temperatures.

8. Repopulation of fine-structure nuclear levels in the fields of a pulsed radio-frequency wave is analogous to repopulation of atomic levels in strong laser fields, for which the field width is larger than all the relaxation widths. From this point of view, these effects are nonlinear in the magnitude of the field. However, their formal description can be carried out by using the equations for the density matrix, in which the field enters in linearly.

It should be emphasized that the repopulation of these levels also takes place when all the levels are populated equally before turning on the bichromatic field, i.e., at high temperatures. Furthermore, as Sec. 3 showed, by turning on the bichromatic and monochromatic rf fields in a specific order, we can augment this repopulation when the system is in the lowest Zeeman level, i.e., ''cooling'' of the nuclear subsystem takes place.

Repopulation of the hyperfine levels is also possible in the gas phase. However, in this case, we find as a rule that $T_1 \approx T_2$, and the theoretical calculations given in this paper require refinement since we have not included longitudinal relaxation.

In conclusion, one of the authors (D. F. Zaretskiĭ) is deeply grateful to Prof. D. Chaplin (Canberra, Australia) who initiated this work, and without whose advice and help it would not have been completed. The authors are also grateful to Prof. R. Couseman (Louvain, Belgium) for discussing the possibility of observing these effects experimentally.

This work was carried out with the partial financial support of the Russian Fund for Fundamental Research (Grant No. 96-02-17612a).

[1] E. Arimondo and G. Orriols, Nuovo Cimento Lett. **17**, 333 (1976).
[2] G. Orriols, Cimento **53**, 1 (1979).
[3] E. A. Korsunskiĭ, B. C. Matisov, and Yu. V. Rozhdestvenskiĭ, Zh. Éksp. Teor. Fiz. **100**, 1438 (1991) [Sov. Phys. JETP **73**, 797 (1991)].
[4] M. O. Scully, Phys. Rep. **219**, 191 (1992).
[5] P. T. Callagan, P. J. Back, D. H. Chaplin *et al.*, Hyperfine Interactions **22**, 39 (1983).
[6] N. J. Stona, and H. Postma, *Low Temperature Nuclear Orientation* Elsevier Scientific Publishers (1986).
[7] V. S. Butylkin, Yu. O. Khronopulo *et al.*, *The Resonant Interaction of Light with Matter*, [in Russian] Nauka, Moscow, (1977).

Translated by Frank J. Crowne

Weibel instability associated with inverse bremsstrahlung absorption of intense electromagnetic radiation

A. Yu. Romanov, V. P. Silin, and S. A. Uryupin

P. N. Lebedev Institute of Physics, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 7 August 1996)
Zh. Éksp. Teor. Fiz. **111**, 1245–1257 (April 1997)

We investigate the Weibel instability in a plasma with a nonstationary three-temperature electron distribution generated by inverse bremsstrahlung of an intense elliptically polarized electromagnetic wave. We show that electron–ion collisions in this strong high frequency field are simultaneously the cause of the development of the instability and the reason it is suppressed. We find plasma and radiation parameters for which spontaneous quasistationary solenoidal electromagnetic fields can increase by many times. © *1997 American Institute of Physics.* [S1063-7761(97)01004-4]

1. INTRODUCTION

It has long been known (see, e.g., Refs. 1–4) that a plasma with an anisotropic electron velocity distribution is unstable against the generation of quasistationary electromagnetic fields. One of the most well-studied plasma instabilities of this kind is the Weibel instability.[1] According to the fundamental postulates of the theory, the threshold for development of the Weibel instability in a fully ionized plasma is determined by electron–ion collisions, which according to the prevailing wisdom prevent the instability from developing (see, e.g., Refs. 5 and 6). However, the situation changes qualitatively when the plasma is in the field of an intense high-frequency laser. In this case the electron–ion collisions can generate an anisotropy in the electron distribution via backward braking absorption of the radiation, which then gives rise to the instability.[7,8] This effect is especially marked when the amplitude of the quiries velocity v_E of an electron in the laser field exceeds the thermal velocity v_T of the electrons, while the ionization state Z of the ions is large enough that $Zv_T \gg v_E \gg v_T$. Under these conditions, according to Ref. 8, heating of the electrons via inverse bremsstrahlung of the radiation leads to the formation of an anisotropic bimaxwellian electron distribution. Such a distribution can in turn give rise to the Weibel instability. However, the well-known postulates of the theory[5,6] indicate that collisions should stabilize the Weibel instability. Our goal in this work is to understand this dual manifestation of electron–ion collisions by developing a first-principles description of the way they affect the Weibel instability when the plasma interacts with an intense laser field. The corresponding theory is set forth in this paper.

The basis of our discussion is the kinetic equation for the electron distribution function, which takes into account *ab initio* the effect of a strong high-frequency electromagnetic wave on the electron–ion collisions. In the second section we show that for $Zv_T \gg v_E \gg v_T$ an elliptically polarized wave can generate a nonstationary three-temperature distribution function. In this case the maximum temperature is along the direction of the propagating wave, while the minimum is along the axis of the highest field intensity. In the third section, we assume that the system is operating far above the threshold for the instability, and analyze possible unstable configurations of quasistationary perturbations of the electromagnetic field. We establish that perturbations grow most effectively whose wave vectors are along the axis of lowest temperature and whose electric field intensity vectors are along the axis of highest temperature. In the fourth section we give a first-principles description of how electron–ion collisions in a strong optical field affect the growth rate of the most unstable quasistationary mode. We establish bounds on the time intervals, plasma parameterss, and optical field for which development of a Weibel instability is possible in a plasma with a nonstationary three-temperature electron velocity distribution.

2. KINETICS OF ELECTRONS IN A STRONG FIELD

Consider a fully ionized plasma in the field of a plane electromagnetic wave of the form

$$\mathbf{E}_0=\frac{1}{2}\,(E_x\,\exp(i\varphi_x),E_y\,\exp(i\varphi_y),0)$$
$$\times\exp(-i\omega_0 t+ik_0 z)+\text{c.c.}, \tag{2.1}$$

where φ_x and φ_y are phases, E_x and E_y are real projections of the electric field intensity vector on the coordinate axes, k_0 is the wave number, and ω_0 is the frequency of the wave. We will assume the frequency ω_0 is much larger than the plasma frequency ω_L of the electrons. Then the wave number is $k_0=\omega_0/c$, where c is the velocity of light, and we can neglect possible dependence of E_x and E_y on coordinates. We will limit ourselves to studying situations where the velocity of light is much larger than both the thermal velocity of the electrons v_T and the amplitude of their quiver velocity in the high-frequency field:

$$v_E=\frac{|e|E}{m\omega_0},\quad E=\sqrt{E_x^2+E_y^2}, \tag{2.2}$$

where e and m are the charge and mass of an electron. Under these conditions, we can use a kinetic equation to describe the motion of the electrons in the form

$$\frac{\partial}{\partial t}\,f+\frac{e}{m}\,\mathbf{E}_0\cdot\frac{\partial}{\partial\mathbf{v}}\,f=\mathrm{St}(f)+\mathrm{St}(f,f), \tag{2.3}$$

 1063-7761/97/040687-07$10.00

where $f=f(\mathbf{v},t)$, $\mathrm{St}(f)$ is the electron–ion collision integral, and $\mathrm{St}(f,f)$ is the electron–electron collision integral. Neglecting small changes in electron energy during collisions with the ions, we will use the following expression for $\mathrm{St}(f)$:

$$\mathrm{St}(f)=\frac{1}{2}\,\nu(v)\,\frac{\partial}{\partial v_i}\,(v^2\delta_{ij}-v_iv_j)\,\frac{\partial}{\partial v_j}\,f, \tag{2.4}$$

$$\nu(v)=4\pi Ze^4n\Lambda m^{-2}v^{-3}, \tag{2.5}$$

where n is the electron density, Z is the multiplicity of ionization of the ions, and Λ is the Coulomb logarithm. The weak logarithmic dependence of the parameter Λ on velocity will be neglected. Let us now discuss Eq. (2.3) for the case of a plasma whose ions are in high ionization states, i.e., $Z\gg 1$, under conditions where the amplitude of the velocity oscillations satisfies the inequalities

$$Zv_T\gg v_E\gg v_T. \tag{2.6}$$

In this case, the effective frequency of electron–ion collisions is $\sim\nu(v_E)$ and the frequency of electron–electron collisions is $\sim\nu(v_T)/Z$. In discussing the effect of high-power radiation on the plasma, we see from the left side of inequality (2.6) that we can neglect the electron–electron collision integral[8] in Eq. (2.3), and the right side of (2.6) allows us to assume the thermal motion of the electrons is slow compared to the rapid motion in the high-frequency field. Under these conditions it is useful to rewrite the kinetic equation in the variables

$$\tau=t,\quad \mathbf{u}=\mathbf{v}-\mathbf{v}_E(t), \tag{2.7}$$

$$\frac{d}{dt}\,\mathbf{v}_E(t)=\frac{e}{m}\,\mathbf{E}_0. \tag{2.8}$$

Then we have for the distribution function of electrons in the new variables $F=F(\mathbf{u},\tau)=f(\mathbf{u}+\mathbf{v}_E(\tau),\tau)$

$$\frac{\partial}{\partial\tau}\,F=\frac{\partial}{\partial u_i}\,D_{ij}(\mathbf{u},\tau)\,\frac{\partial}{\partial u_j}\,F, \tag{2.9}$$

$$D_{ij}(\mathbf{u},\tau)=\frac{1}{2}\,u^3\nu(u)|\mathbf{u}+\mathbf{v}_E(\tau)|^{-3}\{[\mathbf{u}+\mathbf{v}_E(\tau)]^2\delta_{ij}-[\mathbf{u}+\mathbf{v}_E(\tau)]_i[\mathbf{u}+\mathbf{v}_E(\tau)]_j\}. \tag{2.10}$$

According to Eq. (2.9), the only reason for the function F to depend on time is the electron–ion collisions modified by the external field, whose frequency is much smaller than the radiation frequency. In this situation, the function F may be written in the form of a sum $F=\overline{F}+\delta\widetilde{F}$, where $\overline{F}=\overline{F}(\mathbf{u},\tau)$ is that large part of the function F which is slowly varying over a period of the high-frequency field, averaged over the latter:

$$\overline{F}=\frac{\omega_0}{2\pi}\int_0^{2\pi/\omega_0}d\tau\,F(\mathbf{u},\tau), \tag{2.11}$$

and $\delta\widetilde{F}$ is the small ($|\delta\widetilde{F}|\ll\overline{F}$) highly oscillatory correction caused by the collisions. Then, averaging Eq. (2.9) over a period of the high-frequency field and omitting small terms that contain $\delta\widetilde{F}$, we find the following equations for $\overline{F}$:

$$\frac{\partial}{\partial\tau}\,\overline{F}=\frac{\partial}{\partial u_i}\,\overline{D_{ij}(\mathbf{u},\tau)}\,\frac{\partial}{\partial u_j}\,\overline{F}, \tag{2.12}$$

where

$$\overline{D_{ij}(\mathbf{u},\tau)}=\frac{\omega_0}{2\pi}\int_0^{2\pi/\omega_0}d\tau\,D_{ij}(\mathbf{u},\tau). \tag{2.13}$$

Since we are interested in patterns in the evolution of the electron distribution in a strong field of the form (2.1), we can also neglect the dependence of the components of the tensor $D_{ij}(\mathbf{u},\tau)$ on velocity $\mathbf{u}$. In this case, we can use Eqs. (2.8), (2.10) for the averaged component of the diffusion tensor to obtain

$$D_{xx}=\frac{D}{\rho^4\sqrt{1+\rho^2}}\,\frac{E_y^2}{E^2}\left\{\left(1-2\,\frac{E_x^2}{E^2}\sin^2\varphi\right)K\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)-\left(1+\rho^2-\frac{2}{1-\rho^2}\,\frac{E_x^2}{E^2}\sin^2\varphi\right)E\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)\right\}, \tag{2.14}$$

$$D_{xy}=D_{yx}=-\frac{D}{\rho^4\sqrt{1+\rho^2}}\,\frac{E_xE_y}{E^2}\cos\varphi\left\{K\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)-(1+\rho^2)E\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)\right\}, \tag{2.15}$$

$$D_{yy}=\frac{D}{\rho^4\sqrt{1+\rho^2}}\,\frac{E_x^2}{E^2}\left\{\left(1-2\,\frac{E_y^2}{E^2}\sin^2\varphi\right)K\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)\times\left(1+\rho^2-\frac{2}{1-\rho^2}\,\frac{E_y^2}{E^2}\sin^2\varphi\right)E\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right)\right\}, \tag{2.16}$$

$$D_{zz}=\frac{D}{\sqrt{1+\rho^2}}\,K\left(\sqrt{\frac{2\rho^2}{1+\rho^2}}\right). \tag{2.17}$$

In Eqs. (2.14)–(2.17), $K(\kappa)$ and $E(\kappa)$ are the complete elliptic integrals of the first kind

$$K(\kappa)=\int_0^{\pi/2}\frac{d\psi}{\sqrt{1-\kappa^2\sin^2\psi}}, \tag{2.18}$$

and second kind

$$E(\kappa)=\int_0^{\pi/2}d\psi\sqrt{1-\kappa^2\sin^2\psi}, \tag{2.19}$$

Here we have used the notation $\varphi=\varphi_x-\varphi_y$, and

$$D=\frac{\sqrt{2}}{\pi}\,v_E^2\nu(v_E), \tag{2.20}$$

$$\rho^4=1-\frac{4}{E^4}\,E_x^2E_y^2\sin^2\varphi. \tag{2.21}$$

The remaining components of the diffusion tensor equal zero: $D_{xz}=D_{zx}=D_{yz}=D_{zy}=0$. We now turn to a detailed investigation of Eqs. (2.12) in the interesting case of an el-

liptically polarized electromagnetic wave, for which $\varphi = \pi/2$. Then $D_{xy} = D_{yx} = 0$, the diffusion tensor is diagonal, and Eq. (2.12) has the form

$$\frac{\partial}{\partial \tau} \overline{F} = D_{ii} \frac{\partial^2}{\partial u_i^2} \overline{F}, \tag{2.22}$$

where the summation runs over triply repeated indices. The solution to Eq. (2.22) has an especially simple form if the initial distribution of the electrons is Maxwellian, i.e., $\overline{F}(\mathbf{u}, \tau = 0) = F_m(u) = (n/2\pi\sqrt{2\pi} v_T^3) \exp(-u^2/2v_T^2)$. In this case, we obtain from (2.22)

$$\overline{F} = \frac{n}{2\pi\sqrt{2\pi}\, v_{T_x} v_{T_y} v_{T_z}} \exp\left(- \frac{u_x^2}{2v_{T_x}^2} - \frac{u_y^2}{2v_{T_y}^2} - \frac{u_z^2}{2v_{T_z}^2} \right), \tag{2.23}$$

and the components of the thermal velocity increase with time linearly:

$$v_{T_i}^2 = v_{T_i}^2(\tau) = 2D_{ii}\tau + v_T^2. \tag{2.24}$$

Thus, when backward braking absorption of an intense high-frequency electromagnetic field takes place, anisotropic heating of the electrons occurs and a three-temperature Maxwellian velocity distribution is generated. Note that in the special cases of linear ($E_y = 0$) or circular ($E_x = E_y = E/\sqrt{2}$) polarized waves, two of the three temperatures coincide, and the distribution (2.23) reduces to the bimaxwellian distribution discussed previously in Ref. 8. The absolute values of the components of the thermal velocity depend on the values of the diffusion tensor components. In particular, for a circularly polarized wave we have from Eqs. (2.14)–(2.17) that $D_{xx} = D_{yy} = D_{zz}/2 = v_E^2 \nu(v_E)/\sqrt{8}$, indicating more efficient heating of the electrons in the direction orthogonal to the plane of polarization of the electromagnetic wave. The anisotropy in the heating is manifest most clearly for a wave with components $E_x \gg E_y$. In this case, we find from (2.14)–(2.17) that

$$D_{zz} = \sqrt{2} D \ln(2E_x/E_y) > D_{yy} = \sqrt{2} D [\ln(2E_x/E_y) - 1/2] \gg D_{xx} = D/\sqrt{2}. \tag{2.25}$$

According to (2.25), when the light is absorbed by inverse bremsstrahlung, the electrons that are most efficiently heated are those with velocities in the directions of low electric field intensity. As $E_y \to 0$, which corresponds to the case of a linearly polarized wave, according to (2.25) we can neglect the difference between the components D_{zz}, D_{yy} of the diffusion tensor. In this case Eq. (2.25) for these components has a logarithmic dependence on $v_{Ey} = eE_y/m\omega_0$. When we take into account the thermal motion of the electrons and when v_{Ey} is smaller than the electron thermal velocities, this reduces to a logarithmic dependence on v_T.[8]

3. INSTABILITY OF THE ANISOTROPIC DISTRIBUTION

Weibel instability can develop in a plasma with an anisotropic distribution of electrons (see, e.g., Refs. 1, 2, 4). Let us consider the possibility that this instability can develop when the anisotropic distribution is generated by inverse bremsstrahlung of an intense electromagnetic field. Assume that a low-frequency perturbation arises in the plasma of the form

$$\delta\mathbf{E}, \delta\mathbf{B}, \delta F \sim \exp(-i\omega t + i\mathbf{kr}) \tag{3.1}$$

at a frequency ω much smaller than the the radiation frequency ω_0, but much larger than the inverse time over which the nonstationary electron distribution F changes as given by (2.23):

$$\omega_0 \gg |\omega| \gg D_{ii}/v_{T_i}^2. \tag{3.2}$$

In order to describe the response of the plasma to such a perturbation, we use the kinetic equation for the small quasistationary correction $\delta\widetilde{F}$ to the function F, in which according to (3.2) we neglect collisions

$$i(\omega - \mathbf{ku})\delta F = \frac{e}{m} \left\{ \delta\mathbf{E} + \frac{1}{\omega} [\mathbf{u}[\mathbf{k}\delta\mathbf{E}]] \right\} \frac{\partial}{\partial \mathbf{u}} \overline{F}, \tag{3.3}$$

and the linearized Maxwell equations for the fields $\delta\mathbf{E}$ and $\delta\mathbf{B}$:

$$[\mathbf{k}\delta\mathbf{E}] = \frac{\omega}{c} \delta\mathbf{B}, \tag{3.4}$$

$$[\mathbf{k}\delta\mathbf{B}] = -\frac{4\pi}{c} ie \int d\mathbf{u}\, \mathbf{u} \delta F - \frac{\omega}{c} \delta\mathbf{E}. \tag{3.5}$$

We will assume that the perturbations of the field are solenoidal, such that $(\mathbf{k}\delta\mathbf{E}) = 0$. Then from Eqs. (3.3)–(3.5) we find

$$\Lambda_{ij}(\omega, \mathbf{k})\delta E_j = 0, \tag{3.6}$$

$$\Lambda_{ij}(\omega, \mathbf{k}) = \delta_{ij}(1 - \omega_L^2/\omega^2 - k^2c^2/\omega^2) + \frac{\omega_L^2}{n\omega^2} \int d\mathbf{u} \frac{u_i u_j}{\omega - \mathbf{ku}} \left(\mathbf{k} \frac{\partial}{\partial \mathbf{u}} \right) \overline{F}. \tag{3.7}$$

Equation (3.6) has a nontrivial solution $\delta\mathbf{E} \neq 0$ when

$$\mathrm{Det}\{\Lambda_{ij}(\omega, \mathbf{k})\} = 0. \tag{3.8}$$

The dispersion relation (3.8) allows us to investigate the possibility that an electromagnetic instability can develop as a function of the direction of the perturbation wave vector $\mathbf{k}$ and the direction of polarization of the perturbation $\delta\mathbf{E}$. Let us use Eq. (3.8) first to describe the states of the plasma that arise during absorption of a plane electromagnetic wave with wave vector $\mathbf{k}_0 = (0,0,k_0)$ and components $E_x \gg E_y$, i.e., when according to Eqs. (2.24), (2.25) all three components of the electron thermal velocity are different. This difference is especially important once the initial thermal energy of the electrons has doubled, when $v_{T_z} > v_{T_y} \gg v_{T_x}$. We note that the behavior of the electromagnetic instability under the action of a linearly polarized wave ($E_y = 0$) automatically follows from the relations given in this section if we further take $v_{T_z} = v_{T_y} > v_{T_x}$. Similar behavior occurs under the action of a circularly polarized wave ($E_x = E_y = E/\sqrt{2}$), for which case we should assume $v_{T_x} = v_{T_y} < v_{T_z}$ in discussing the stability conditions. We will analyze the most general case below, when all three components of the thermal velocity are different. With the goal of determining the configura-

tion of the most unstable perturbations, let us discuss the consequences of Eqs. (3.6)–(3.8) for all three independent orientations of the wave vector $\mathbf{k}$ and field $\delta\mathbf{E}$.

We begin our analysis with a discussion of perturbations with $\mathbf{k}=(k,0,0)$ and $\delta\mathbf{E}=(0,0,\delta E)$. Then using the distribution (2.23), from (3.6)–(3.8) we find

$$\Lambda_{zz}(\omega,k)=1-\frac{\omega_L^2}{\omega^2}-\frac{k^2c^2}{\omega^2}+\frac{\omega_L^2}{\omega^2}\frac{v_{T_z}^2}{v_{T_x}^2}\left[1-J_+\left(\frac{\omega}{kv_{T_x}}\right)\right]=0, \tag{3.9}$$

where the function $J_+(\beta)$ has the form[9]

$$J_+(\beta)=\beta\ \exp(-\beta^2/2)\int_{i\infty}^{\beta}dy\ \exp(y^2/2). \tag{3.10}$$

At low enough frequencies that $|\omega|\ll kv_{T_x}\ll kc$, using the expansion

$$J_+(\beta)=-i\beta\sqrt{\pi/2},\quad |\beta|\ll 1, \tag{3.11}$$

for $\omega=i\gamma$ we obtain from Eq. (3.9)

$$\gamma=\sqrt{\frac{2}{\pi}}\,kv_{T_x}\frac{v_{T_x}^2}{v_{T_z}^2}\left(\frac{v_{T_z}^2}{v_{T_x}^2}-1-\frac{k^2c^2}{\omega_L^2}\right). \tag{3.12}$$

Since in the cases under discussion we have $v_{T_z}>v_{T_x}$, Eq. (3.12) determines the growth rate of the Weibel instability for perturbations with small wave numbers

$$k<\frac{\omega_L}{c}\sqrt{\frac{v_{T_z}^2}{v_{T_x}^2}-1}\equiv k_m\sqrt{3}. \tag{3.13}$$

The perturbation with wave number $k=k_m$ has the maximum growth rate γ_m. In this case

$$\gamma_m=\sqrt{\frac{8}{27\pi}}\,\omega_L\frac{v_{T_z}}{c}\left(1-\frac{v_{T_x}^2}{v_{T_z}^2}\right)^{3/2}. \tag{3.14}$$

Because γ_m should be smaller than $k_mv_{T_x}$, Eq. (3.14) is valid if the degree of anisotropy of the temperatures is small:

$$v_{T_z}^2-v_{T_x}^2\ll\sqrt{9\pi/8}\,v_{T_z}^2. \tag{3.15}$$

In the range of high frequencies $kv_{T_x}\ll|\omega|\ll kc$, using the expansion

$$J_+(\beta)=1+\beta^{-2}+\ldots,\quad |\beta|\gg 1, \tag{3.16}$$

for $\omega=i\gamma$ we obtain from (3.9)

$$\gamma=kv_{T_z}(1+k^2c^2/\omega_L^2)^{-1/2}\gg kv_{T_x}. \tag{3.17}$$

This solution is realized only when the temperature anisotropy is large, i.e., $v_{T_z}\gg v_{T_x}$. The maximum value of the growth rate is reached for large wave numbers $k>\omega_L/c$, and is less than $\omega_Lv_{T_z}/c$ (compare with (3.14)).

The response of the plasma to a perturbation with $\mathbf{k}=(k,0,0)$ and $\delta\mathbf{E}=(0,\delta E,0)$ is described by the dispersion relation $\Lambda_{yy}(\omega,k)=0$, which differs from (3.9) by the replacement of v_{T_z} by the smaller velocity v_{T_y}. Consequently, the special features of the growth of these perturbations are described by Eqs. (3.12)–(3.15) and (3.17), in which we must replace v_{T_z} by v_{T_y}. Because $v_{T_y}<v_{T_z}$, the maximum growth rate for these perturbations is smaller than that given by Eqs. (3.14), (3.17).

The response to a perturbation with $\mathbf{k}=(0,k,0)$ and $\delta\mathbf{E}=(\delta E,0,0)$ is described by the function $\Lambda_{xx}(\omega,k)=0$, which differs from (3.9) through the replacement of v_{T_z} by v_{T_x} and v_{T_x} by v_{T_y}. Because we have $v_{T_y}>v_{T_x}$, in this case the perturbation is damped. If, as before, we have $\mathbf{k}=(0,k,0)$, while the field is oriented along another axis $\delta\mathbf{E}=(0,0,\delta E)$, then the dispersion relation has the form (3.9), where in place of v_{T_x} the larger quantity v_{T_y} enters in. Since the velocity v_{T_y} is close to v_{T_z}, the maximum growth rate of this perturbation is described by the expression

$$\gamma_m=\sqrt{\frac{8}{27\pi}}\,\omega_L\frac{v_{T_z}}{c}\left(1-\frac{v_{T_y}^2}{v_{T_z}^2}\right)^{3/2} \tag{3.18}$$

and turns out to be significantly smaller than that described by (3.14). Finally, the response of the plasma to perturbations with $\mathbf{k}=(0,0,k)$ and $\delta\mathbf{E}=(\delta E,0,0)$ or $\delta\mathbf{E}=(0,\delta E,0)$ is described by the functions $\Lambda_{xx}(\omega,k)$ and $\Lambda_{yy}(\omega,k)$ respectively. These functions differ from (3.9) by the replacement of v_{T_x} by v_{T_z} and v_{T_z} by v_{T_x} or v_{T_z} by v_{T_y}, respectively. Since v_{T_z} is larger than both v_{T_x} and v_{T_y}, this perturbation is damped. From this analysis it follows that perturbations that grow most effectively are solenoidal perturbations of the field with wave vectors in the direction with the smallest component of electron thermal velocity and polarizations in the direction with the largest component of the thermal velocity.

4. EFFECT OF COLLISIONS ON THE INSTABILITY

The uniqueness of the electromagnetic instability created by inverse bremsstrahlung absorption of intense electromagnetic radiation lies in the fact that the instability is caused by anisotropy of the temperature distribution generated by heating of the electrons when they collide with ions. In this sense we can say that the electron–ion collisions cause the instability. However, according to the basic postulates of the theory of the Weibel instability, collisions suppress the instability and determine the threshold for its appearance.[5,6] In order to clarify the dual role of the electron–ion collisions, let us consider a near-threshold theory of the Weibel instability. In this case, keeping in mind the results of the previous section, we will limit ourselves to discussing the most efficiently excited perturbations, i.e., those which have the configuration $\mathbf{k}=(k,0,0)$, $\delta\mathbf{E}=(0,0,\delta E)$, and $\delta\mathbf{B}=(0,\delta B,0)$ and depend only on the single coordinate as $\sim\exp(ikx)$. In writing this perturbation we relax the right side of inequality (3.2), and in place of (3.3) we use a more general equation that takes into account the effect of collisions on the quasistationary correction to the function F (2.23),

$$\frac{\partial}{\partial\tau}\,\delta F+iku_x\delta F+\frac{e}{m}\,\delta E\,\frac{\partial}{\partial u_z}\,\overline{F}+\frac{e}{mc}\,\delta B\times\left(u_x\,\frac{\partial}{\partial u_z}-u_z\,\frac{\partial}{\partial u_x}\right)\overline{F}=D_{xx}\,\frac{\partial^2}{\partial u_x^2}\,\delta F+D_{yy}\,\frac{\partial^2}{\partial u_y^2}\times\delta F+D_{zz}\,\frac{\partial^2}{\partial u_z^2}\,\delta F,\tag{4.1}$$

where the components of the electromagnetic field δE and δB are themselves determined by the correction δF in accordance with the Maxwell equations

$$\frac{\partial}{\partial\tau}\,\delta B=ikc\,\delta E,\tag{4.2}$$

$$\frac{\partial}{\partial\tau}\,\delta E=ikc\,\delta B-4\pi e\int d\mathbf{u}\;u_z\delta F.\tag{4.3}$$

In discussing Eqs. (4.1)–(4.3) it is convenient to introduce the functions

$$\delta F_x(u_x,\tau)=\int_{-\infty}^{\infty}du_y\int_{-\infty}^{\infty}du_zu_z\delta F(\mathbf{u},\tau),\tag{4.4}$$

$$F_m(u_x,\tau)=\frac{n}{\sqrt{2\pi}\;v_{T_x}}\exp\left(-\frac{u_x^2}{2v_{T_x}^2}\right),\tag{4.5}$$

which are independent of the velocity components u_y and u_z. Furthermore, taking into account the time dependence of the function $\overline{F}$ in Eqs. (2.23), (2.24), and also the first-order correction to the differential equations (4.1)–(4.3), we find approximately that

$$\delta E,\delta B,\delta F_x\sim\exp\left[\int_0^\tau d\tau'\,\gamma(\tau')\right].\tag{4.6}$$

Then from (4.1)–(4.3) we obtain

$$[\gamma^2(\tau)+k^2c^2]\,\delta E=-4\pi e\int_{-\infty}^{\infty}du_x\gamma(\tau)\,\delta F_x(u_x,\tau),\tag{4.7}$$

$$\gamma(\tau)\,\delta F_x(u_x,\tau)=\frac{e}{m}\,\delta E\left(1-\frac{v_{T_z}^2}{v_{T_x}^2}\right)F_m(u_x,\tau)+\frac{\gamma(\tau)}{\gamma(\tau)+iku_x}\,\frac{e}{m}\,\delta E\,\frac{v_{T_z}^2}{v_{T_x}^2}\,F_m(u_x,\tau)+\frac{D_{xx}}{\gamma(\tau)+iku_x}\,\frac{\partial^2}{\partial u_x^2}\,[\gamma(\tau)\,\delta F_x(u_x,\tau)].\tag{4.8}$$

At the threshold for excitation of the Weibel instability,

$$|ku_x|\gg\gamma(\tau),\quad D_{xx}/v_{T_x}^2.\tag{4.9}$$

These conditions ensure applicability of the approximation (4.6). Furthermore, according to (4.9), in solving Eq. (4.8) we first use the relation

$$[\gamma(\tau)+iku_x]^{-1}=-iP\left(\frac{1}{ku_x}\right)+\pi\delta(ku_x)\tag{4.10}$$

(where P denotes principal value), and then include collisions by perturbation theory. Taking these assumptions into account, we obtain from (4.8)

$$\gamma(\tau)\,\delta F_x(u_x,\tau)=\frac{e}{m}\,\delta E\left(1-\frac{v_{T_z}^2}{v_{T_x}^2}\right)F_m(u_x,\tau)+\frac{e}{m}\,\delta E\left[\pi\delta(ku_x)-iP\left(\frac{1}{ku_x}\right)\right]\times\left\{\frac{v_{T_z}^2}{v_{T_x}^2}\,\gamma(\tau)F_m(u_x,\tau)+\left(1-\frac{v_{T_z}^2}{v_{T_x}^2}\right)D_{xx}\,\frac{\partial^2}{\partial u_x^2}\,F_m(u_x,\tau)\right\}.\tag{4.11}$$

Substituting expression (4.11) into Eq. (4.7) and dropping a small term proportional to $\gamma^2(\tau)\ll k^2c^2$, we find for the growth rate of the instability

$$\gamma(\tau)=\gamma_m(\tau)-\sqrt{\frac{2}{27\pi}}\;\omega_L\,\frac{v_{T_z}}{c}\left(1-\frac{v_{T_x}^2}{v_{T_z}^2}\right)^{3/2}\times\left(1-\frac{k}{k_m}\right)^2\left(2+\frac{k}{k_m}\right),\tag{4.12}$$

$$\gamma_m(\tau)=\left(1-\frac{v_{T_x}^2}{v_{T_z}^2}\right)\left\{\sqrt{\frac{8}{27\pi}}\;\omega_L\,\frac{v_{T_z}}{c}\;\sqrt{1-\frac{v_{T_x}^2}{v_{T_z}^2}}-\frac{D_{xx}}{v_{T_x}^2}\right\},\tag{4.13}$$

$$k_m=k_m(\tau)=\sqrt{3}\;\frac{\omega_L}{c}\;\sqrt{\frac{v_{T_z}^2}{v_{T_x}^2}-1}.\tag{4.14}$$

In Eqs. (4.12)–(4.14), the components of the electron thermal velocity increase with time in accordance with Eq. (2.24). The maximum value of the growth rate $\gamma_m(\tau)$ is associated with perturbations with wave vector equal to $k_m(\tau)$. Equation (4.13) differs from Eq. (3.14) obtained earlier by the fact that the dependence of the growth rate on the effective electron–ion collision frequency in the strong field is explicit. As in the theory of a plasma with no radiation field, the threshold of the Weibel instability is proportional to the electron–ion collision frequency. Equation (4.13) allows us to trace how anisotropic heating of the electrons can lead to the development of a Weibel instability. At the initial time, when $v_{T_x}=v_{T_z}=v_T$ holds, the plasma is stable. As the electrons are heated, the anisotropy in the temperatures increases. As soon as the maximum growth rate $\gamma_m(\tau)$ becomes positive, solenoidal electromagnetic perturbations grow, following Eq. (4.6). According to Eq. (4.13) the time τ_0 at which the instability appears is given by the equation

$$\frac{16}{27\pi}\frac{\omega_L^2}{c^2}\,\tau_0(D_{zz}-D_{xx})=D_{xx}^2(2D_{xx}\tau_0+v_T^2)^{-2}. \qquad (4.15)$$

Note that Eq. (4.15) is valid under the condition $v_{T_z}^2(\tau_0)=2D_{zz}\tau_0+v_T^2<v_E^2$, i.e., when the electron thermal velocity is small compared to the amplitude of the quiver velocity in the high-frequency field of the radiation. If the intensity of the radiation and frequency of electron–ion collisions satisfy $Zv_T\gg v_E\gg v_T$ and

$$\sqrt{\frac{8\pi}{27}}\,\sqrt{2\ln\frac{2E_x}{E_y}-1}\,\frac{v_E}{c}\,\frac{\omega_L}{\nu(v_T)}\gg 1, \qquad (4.16)$$

then the instability appears at time

$$\tau_0=\frac{27}{16}\,\frac{c^2}{v_T v_E}\,\frac{\nu(v_T)}{\omega_L^2}\left[2\ln\frac{2E_x}{E_y}-1\right]^{-1}, \qquad (4.17)$$

when the velocity $v_{T_x}(\tau_0)$ is close to the initial velocity v_T, while $v_{T_z}(\tau_0)$ is much less than v_E. In this case we have from (4.17) the estimate $\tau_0\sim(v_E/v_T)\nu^{-1}(v_T)\sim(v_T^2/v_E^2)\nu^{-1}(v_E)$ as a boundary condition on the applicability of (4.16). In other words, this time is found to be much smaller than the characteristic time between electron–ion collisions, which is $\sim 1/\nu(v_E)$. The appearance of anisotropy in the electron distribution is caused by the high intensity of the heating radiation. If $Zv_T\gg v_E\gg v_T$ holds as before, but the frequency of electron–ion collisions is somewhat larger,

$$1\gg\sqrt{\frac{8\pi}{27}}\,\sqrt{2\ln\frac{2E_x}{E_y}-1}\,\frac{v_E}{c}\,\frac{\omega_L}{\nu(v_T)}\gg\frac{v_T^3}{v_E^3}\left(2\ln\frac{2E_x}{E_y}\right)^{3/2}, \qquad (4.18)$$

then the instability appears later, at time

$$\tau_0=\frac{3}{4}\,\pi^{2/3}\left\{\frac{c^2}{v_E^2}\,\frac{1}{\omega_L^2\nu(v_E)}\,\frac{1}{2\ln(2E_x/E_y)-1}\right\}^{1/3}. \qquad (4.19)$$

At this time the velocity $v_{T_x}(\tau_0)$ is considerably larger than the initial velocity v_T. However, due to the right-hand inequality (4.18), the largest component of the thermal velocity $v_{T_z}(\tau_0)$ is still smaller than v_E. Here it is appropriate to point out that during the period prior to v_{T_z} becoming comparable to v_E, both in conditions (4.16) and (4.18), the level of solenoidal excitations described by Eqs. (4.6), (4.13) of the electromagnetic field increases by many times compared with its initial value.

The latter assertion follows from use of the linear theory of the instability to derive a simple estimate of the ratio $B^2/B_{\rm sp}^2$ of the energy density of the quasistationary field at time $\tau_m\sim v_E^2/2D_{zz}$, when $v_{T_z}\sim v_E$ holds, to the initial energy density of the spontaneous electromagnetic fields. According to Eq. (4.13), for perturbations of the field with wave vector $k_m\sim\sqrt{3}(\omega_L/c)(v_{T_z}/v_{T_x})$ (4.14), we have

$$\frac{B^2}{B_{\rm sp}^2}\sim\exp(2\gamma_m\tau_m)\sim\exp\left[\sqrt{\frac{2\pi}{27}}\,\frac{1}{\ln(2E_x/E_y)}\,\frac{\omega_L v_E}{c\,\nu(v_E)}\right].$$

It is clear from inequalities (4.16), (4.18) that this ratio is much larger than unity. The linear theory gives the maximum possible increase in the initial noise, which is not realized when nonlinear effects become important before the time τ_m is reached, leading to saturation of the Weibel instability. Numerical investigations of the nonlinear stage of the Weibel instability show (see, e.g., Ref. 2) that the energy density of the magnetic field in the saturated state is no more than 10% of the kinetic energy of the initial anisotropic electron distribution. In light of this observation, if we estimate the kinetic energy density of electrons in a strong high-frequency field by nmv_E^2, we find for the energy of the quasistationary magnetic field a value $B_m^2/4\pi\sim 0.1nmv_E^2$, or

$$B_m[\mathrm{G}]\sim 2.5\times 10^5\left[\frac{n[\mathrm{cm}^{-3}]}{10^{20}}\right]^{1/2}\left[\frac{I[\mathrm{W/cm}^2]}{10^{15}}\,\lambda^2[\mu\mathrm{m}]\right]^{1/2},$$

where λ is the wavelength of the radiation. According to this last estimate, for interaction with radiation from a neodymium laser with $\lambda\sim 1\ \mu$m and $I\sim 4\times 10^{15}$ W/cm^2 in a plasma with electron density $n\sim 10^{20}$ cm^{-3}, ionization state $Z\sim 10$, and electron temperature ~ 150 eV, a magnetic field of ~ 0.5 mG is generated. Here we have chosen the electron temperature and ionization state of the ions so that the conditions for applicability of the theory given here are satisfied, assuming that the inequality $Zv_T\gg v_E\gg v_T$ holds.

We note that the limiting value of the magnetic field B_m may not be reached if the time τ_m is smaller than the time for saturation of the instability. In this case, from the time τ_m on the evolution of the electron distribution is described by perturbation-theory equations for a small high-frequency field in a hot plasma, where the amplitude of the electron velocity oscillations is smaller than the effective thermal velocity. In this case, it is well known that the temperature anisotropy is relaxed not only by solenoidal magnetic fields but also by the electron collisions themselves, which weakens the effectiveness of the transformation of kinetic energy of the anisotropic electron distribution into magnetic energy. Still another reason why creation of a magnetic field is hindered is the short duration of the laser radiation pulse, both compared to the saturation time of the Weibel instability

$$\tau_s[\mathrm{s}]\sim\frac{1}{\gamma_m}\ln\frac{B_m}{B_{\rm sp}}\gtrsim\sqrt{\frac{27\pi}{8}}\,\frac{c}{\omega_L v_E}\ln\frac{B_m}{B_{\rm sp}}=2\times 10^{-13}\left(\ln\frac{B_m}{B_{\rm sp}}\right)\times\left[\frac{10^{20}}{n[\mathrm{cm}^{-3}]}\right]^{1/2}\left[\frac{10^{15}}{I[\mathrm{W/cm}^2]}\,\lambda^{-2}[\mu\mathrm{m}]\right]^{1/2}$$

(where $B_{\rm sp}$ is the spontaneous magnetic field strength) and to the time

$$\tau_m[\mathrm{s}]\sim\frac{\pi}{4}\left[\frac{1}{\ln(2E_x/E_y)}\right]\nu^{-1}(v_E)=10^{-13}\times\left(\frac{4}{\Lambda}\right)\left(\frac{10}{Z}\right)\left[\frac{10^{20}}{n[\mathrm{cm}^{-3}]}\right]\times\left[\frac{I[\mathrm{W/cm}^2]}{10^{15}}\,\lambda^2[\mu\mathrm{m}]\right]^{3/2}\frac{1}{\ln(2E_x/E_y)}.$$

Taking as estimates $B_m/B_{sp}\sim 100$, $\Lambda\sim 4$, $Z\sim 10$, $n\sim 10^{20}$ cm^{-3} and $E_x/E_y\sim 3$, for light from a neodymium laser with $\lambda\sim 1\ \mu$m and intensity $I\sim 4\times 10^{15}\ W/\text{cm}^2$ we have $\tau_s\sim\tau_m\sim 0.5$ ps. For these parameters of the plasma and the light, this additional reduction in the efficiency with which the solenoidal magnetic field is generated can occur only when subpicosecond laser pulses are used.

In conclusion, we note that these investigations of the Weibel instability in a plasma heated by high-power electromagnetic radiation show that, on the one hand, it is necessary as usual to include electron–ion collisions in order to determine threshold conditions arising from the corresponding dissipation. On the other hand, these same electron–ion collisions generate the trimaxwellian electron distribution (2.23) in the process of absorbing the radiation that heats the plasma. Since this distribution is in general anisotropic with respect to velocities, it is also a cause for excitation of the Weibel instability. In this case, quasistatic magnetic fields are excited in a plasma heated by high-power radiation in a very short time under conditions (2.6).

This work was performed under the auspices of Project No. 96-02-17002 of the Russian Fund for Fundamental Research, and also with partial support of the International Science and Engineering Center (Project No. 310).

[1] E. S. Weibel, Phys. Rev. Lett. **2**, 83 (1959).
[2] R. C. Davidson, D. A. Hammer, D. Haber, and C. E. Wagner, Phys. Fluids **15**, 317 (1972).
[3] T. Okada, I. Yabe, and K. Niu, J. Phys. Soc. Jpn. **43**, 1042 (1977).
[4] R. C. Davidson, in: A. A. Galeev and R. N. Sudan (Eds.), *Basic Plasma Physics, Vol. 1* (North-Holland, Amsterdam, 1983).
[5] M. Wallace, J. U. Brackbill, C. W. Granfill, D. W. Forslund, and R. J. Mason, Phys. Fluids **30**, 1085 (1987).
[6] A. F. Alexandrov, L. S. Bogdankevich, and A. A. Rukhadze, *Principles of Plasma Electrodynamics* (Springer, Berlin, 1984).
[7] V. P. Silin and S. A. Uryupin, Zh. Eksp. Teor. Fiz. **111**, 107 (1997) [JETP **84**, 59 (1997)].
[8] B. N. Chichkov, S. A. Shumsky, and S. A. Uryupin, Phys. Rev. A **45**, 7475 (1992).
[9] V. P. Silin and A. A. Rukhadze, *Electrodynamic Properties of Plasmas and Plasmalike Media* (Gosatomizdat, Moscow, 1961), p. 91 [in Russian].

Translated by Frank J. Crowne

Nonlinear theory of a plasma microwave oscillator using cable waves

M. Birau

École Polytechnique, Palaiseau, France

M. A. Krasil'nikov, M. V. Kuzelev, and A. A. Rukhadze

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia
(Submitted 17 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1258–1273 (April 1997)

The nonstationary problem of excitation of a plasma microwave oscillator of a finite length driven by a pulsed relativistic beam is analyzed. Both analytic and numerical techniques have been used in studying the oscillator dynamics at given beam parameters (configuration, electron energy, and current pulse), with different plasma configurations, and at various lengths of the system. Oscillator characteristics such as the output power, efficiency, and output spectrum have been determined at parameters close to real experimental values. Conditions of optimal oscillator operation have been determined. © *1997 American Institute of Physics.*
[S1063-7761(97)01104-9]

1. Relativistic plasma microwave oscillators based on cable waves have been intensely studied both experimentally and theoretically over recent years. The first prototype of this oscillator was built in 1982.[1] Later detailed measurements of the oscillator were taken with the aim not only to detect the output radiation and determine its power, but also to record spectra of output radiation.[2]

Although it may seem strange, the theory has been lagging behind the experiment in recent time. The authors of most theoretical studies investigated various limiting cases, in which they could considerably simplify theoretical models and even perform analytic calculations.[3] The parameters of real experiments quoted in this paper, however, are notably different from those limiting cases (the relativistic parameter of the electron beam is moderate, the beam current is of the order of the vacuum limit, etc.), so the previous theoretical estimates do not directly apply to the reported experiments. The results of computer simulations of a beam–plasma system with parameters of a real experiment and based on the initial problem statement[4] also cannot be applied to an oscillator where space transfer of radiation, its amplification, and output coupling are important. Besides, all the studies quoted above ignored the fact that in experiments the electron beam is formed by a pulse with definite shapes of its leading and trailing edges.

Some authors, however, used the oscillator model based on the direct numerical solution of the Maxwell–Vlasov equation system.[5] Physical interpretation of such models is often difficult, like interpretation of real experiments. Moreover, definition of boundary conditions at the oscillator output coupler that are faithful to the Maxwell–Vlasov equations is a difficult problem; its more or less accurate solution can be obtained only in some limiting cases.[6] Finally, even with state-of-the-art computation techniques, numerical implementation of accurate models demands a lot of CPU time, especially as the data processing includes analysis of numerous variants, or even optimization of the model parameters.

The aim of the reported work was to create and investigate a simple model of a relativistic microwave oscillator based on a cable wave with two assumptions:

— the generation frequency ω (average frequency of the radiation spectrum) is of the order of the frequency at which the plasma wave gain (growth rate) has a maximum;

— the width of the output spectrum $\Delta\omega$ is much smaller than the average frequency ω.

These assumptions have proved to be true in modeling the oscillator by directly solving the Maxwell–Vlasov equation system[5] and are partially justified by experimental results. It is too early, however, to claim full agreement with experimental data,[2,7] since more improved measurements are necessary.

The model based on the assumptions given above was used in the study[8] of a microwave oscillator based on a volume wave. The cable plasma wave is a surface wave, and some of its features are determined by this property. Note that most real prototypes of plasma microwave oscillators were based on this type of plasma waves.

2. First let us recall some properties of cable plasma waves.[9]

Consider a metal waveguide of arbitrary cross section. Let us denote by z the longitudinal coordinate in the waveguide and by $\mathbf{r}_\perp$ the coordinate in the cross section. The waveguide contains an infinitely thin beam of plasma defined by the expression $n_p S_p \delta(\mathbf{r}_\perp - \mathbf{r}_p)$, where $n_p S_p$ is the plasma density per unit length and $\mathbf{r}_p$ is the plasma coordinate in the waveguide cross section. The plasma is magnetized by a strong external magnetic field aligned with the z-axis. We assume that the eigenfunctions of the waveguide cross section, $\varphi_n(\mathbf{r}_\perp)$, and the corresponding eigenvalues $k_{\perp n}^2$ are known ($n=1,2,3,\ldots$).

In the linear approximation, the modes of a waveguide containing plasma are determined by equations[9]

$$\frac{\partial}{\partial z}\left(\Delta_\perp + \frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{dt^2}\right)\Psi = 4\pi\rho_p S_p \delta(\mathbf{r}_\perp - \mathbf{r}_p),$$

$$\frac{\partial^2 \rho_p}{\partial t^2} = -\frac{\omega_p^2}{4\pi}\frac{\partial E_z}{\partial z}, \tag{1}$$

$$E_z = \left(\frac{\partial^2}{\partial z^2} - \frac{1}{c^2}\frac{\partial^2}{\partial t^2}\right)\Psi.$$

Here Ψ is the polarization potential, E_z is the longitudinal component of electric field, ρ_p is the plasma density perturbation, and $\Delta_\perp$ is the transverse component of the Laplace operator. On the waveguide metal surface $\Psi = 0$. The parameters Ψ and ρ_p as functions of z and t are assumed to be proportional to $\exp(-i\omega t + ik_z z)$, and Ψ is expanded in terms of the waveguide eigenfunctions:

$$\Psi = \sum_{n=1}^{\infty} A_n \varphi_n(\mathbf{r}_\perp)\exp(-i\omega t + ik_z z),$$
$$\rho_p = \tilde{\rho}_p(\mathbf{r}_\perp)\exp(-i\omega t + ik_z z). \tag{2}$$

Substitute Eq. (2) into Eq. (1), derive from the first line of Eq. (1) the coefficients A_n, and substitute them into the second line of Eq. (1). Then multiply the second line of Eq. (1) by $\delta(\mathbf{r}_\perp - \mathbf{r}_p)$ and integrate it over the waveguide cross section. Eliminating $\tilde{\rho}_p$ from the resulting equation, we derive the spectrum of plasma oscillations in the waveguide:

$$\omega^2 = \omega_p^2 \frac{\chi^2}{k_{\perp p}^2},$$
$$k_{\perp p}^2 = \left[\sum_{n=1}^{\infty} \frac{S_p}{k_{\perp n}^2 + \chi^2}\frac{\varphi_n^2(\mathbf{r}_p)}{\|\varphi_n\|^2}\right]^{-1}. \tag{3}$$

Here $\omega_p = \sqrt{4\pi e^2 n_p/m}$ is the Langmuir frequency, $\chi^2 = k_z^2 - \omega^2/c^2$, and $k_{\perp p}^2$ is the square of the transverse wave number of the thin plasma in the waveguide. The transverse components of the wave defined by Eq. (3) can be expressed in terms of Ψ according to generally known formulas.[10] In the frequency band $\omega < k_z c$ the wave with the spectrum defined by Eq. (3) is a surface wave, i.e., its field intensity decays with distance from the plasma beam.

If the waveguide is circular with radius R, Eq. (3) can be simplified. In this case, the plasma is described as a tube with an average radius $r_p < R$ and a thickness Δ_p (the plasma density per unit length is $2\pi r_p \Delta_p n_p$), and the waveguide modes are described by Bessel functions. Here we consider only the axially symmetric configuration, and after summation in the second line of Eq. (3) we obtain

$$k_{\perp p}^2 = \left\{ r_p \Delta_p I_0^2(\chi r_p)\left[\frac{K_0(\chi r_p)}{I_0(\chi r_p)} - \frac{K_0(\chi R)}{I_0(\chi R)}\right]\right\}^{-1}, \tag{4}$$

where I_0 and K_0 are modified Bessel functions of the first and second kind, respectively.

For $k_z \to 0$ we derive from Eqs. (3) and (4) an explicit expression for the spectrum:

$$\omega = k_z c\left(1 + \frac{k_{\perp p}^2 c^2}{\omega_p^2}\right)^{-1/2},$$
$$k_{\perp p}^2 = [r_p \Delta_p \ln(R/r_p)]^{-1}. \tag{5}$$

It is clear that at higher plasma density, when $\omega_p^2 \gg k_{\perp p}^2 c^2$, the phase velocity of the wave given by Eq. (5) is close to the speed of light. One can show[9] that the field configuration of this wave is similar to that of the cable (or TEM) wave in a metallic coaxial line,[10] so the wave with the spectrum defined by Eq. (5) is called a cable plasma wave. As $\omega_p \to \infty$ (the plasma turns into a metal), the wave described by Eq. (5) becomes identical to the cable wave in a coaxial line.

In the region of shorter wavelengths, when $k_z r_p \geqslant 1$, both the phase velocity and field configuration of the surface plasma wave are radically different from those of the cable wave. In this case the spectrum is similar to that of surface waves in deep water:

$$\omega = \omega_p \sqrt{k_z \Delta_p/2}. \tag{6}$$

3. An important component of the plasma oscillator is its output coupler (horn). Let us denote by $z = L$ the horn coordinate, and by $z = 0$ the plane where the electron beam is injected into the oscillator. It is obvious that the plasma occupies in the waveguide the region $0 \leqslant z \leqslant L$. Suppose that the injection plane contains a metal grid (or a cut-off section of a waveguide), which is transparent for the electron beam but does not transmit the radiation. The boundary condition for the electromagnetic field at $z = 0$ is obvious: it is total reflection. Now the boundary condition at $z = L$ should be defined.

In experiments, the horn is a metallic coaxial line with the outside radius R and inside radius r_p, i.e., the plasma coaxial line is matched to the metallic one. We must formulate boundary conditions for the field at this junction. They were previously formulated in a general but very complicated form.[5,6] For our purpose, the boundary condition at $z = L$ in a simple form is sufficient.

An electron beam generates a surface plasma wave whose phase velocity is close to the unperturbed velocity u of the electron beam. On the $z = L$ boundary, a fraction of the plasma wave is generated into the metallic coaxial line as a cable wave, and the rest is reflected back into the plasma waveguide. If the plasma wave generated by the electron beam is well matched to the cable wave ($k_z r_p \ll 1$, $\omega_p \gg k_{\perp p} c$), its reflection at the junction with the coaxial line is determined by the difference between the phase velocities, as in Fresnel's problem. Since the phase velocity of the plasma wave is $\sim u$ and in the coaxial line it equals the speed of light c, the reflectivity in the case of a relativistic velocity u is $\kappa = 1/4\gamma^2$, where $\gamma = (1 - u^2/c^2)^{-1/2}$ is the relativistic factor of the electron beam.

If $k_z r_p \sim 1$, one must take into account not only the difference between the phase velocities, but also between the transverse patterns of both the plasma and cable waves. It is obvious that the reflectivity will grow with k_z (or frequency) and tend to unity as $\omega \to \omega_p$.

By estimating the energy distribution in the plasma cable wave inside the plasma tube ($r < r_p$) and outside it ($r > r_p$)[9] and using some experimental data,[1,7] one can derive the following extrapolation formula for the reflectivity in the range of higher frequencies:

$$\kappa = \frac{1}{4\gamma^2}\left(1 + 0.25\,\frac{\omega^2 r_p^2}{u^2\gamma^2}\,\ln\frac{R}{2r_p}\right). \tag{7}$$

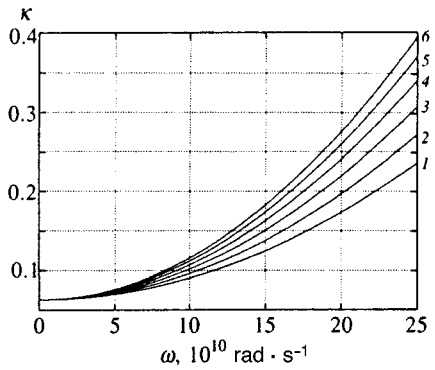

FIG. 1. Reflectivity as a function of frequency for several plasma radii: (*1*) $r_p=0.7$; (*2*) 0.8; (*3*) 0.9; (*4*) 1.0; (*5*) 1.1; (*6*) 1.2 cm.

Equation (7) is not valid at $\omega\sim\omega_p$, but in this frequency band, when plasma waves are potential, it makes no sense to consider the oscillator as a source of electromagnetic waves; furthermore, the model of an infinitely thin plasma does not apply. Therefore, we will assume in what follows that the frequencies are not very high and use the estimate given by Eq. (7). It will be shown below that this assumption is well justified. Figure 1 shows the reflectivity κ as a function of ω calculated by Eq. (7) for various plasma radii. Hereinafter the other parameters are fixed, namely, the waveguide radius is $R=1.8$ cm, the electron beam velocity $u=2.6\times10^{10}$ cm/s, and the electron relativistic factor $\gamma=2$.

4. Now let us derive the basic nonlinear equations for our analysis. If there is an infinitely thin electron beam, the first line in Eq. (1) can be simplified:

$$\frac{\partial}{\partial z}\left(\Delta_\perp+\frac{\partial^2}{\partial z^2}-\frac{1}{c}\frac{\partial^2}{\partial t^2}\right)\Psi = 4\pi\rho_p S_p\delta(\mathbf{r}_\perp-\mathbf{r}_p)+4\pi\rho_b S_b\delta(\mathbf{r}_\perp-\mathbf{r}_b), \tag{8}$$

where ρ_b is the perturbation of the charge density in the beam defined by the equation[8]

$$\rho_b=en_b\frac{\lambda}{N}\sum_j\delta(z-z_j). \tag{9}$$

Here n_bS_b is the number of electrons in the beam per unit length, z_j is the longitudinal coordinate of the jth electron, λ is a characteristic length (see below), and N is the number of electrons (large-size particles) on a section of the unperturbed beam with the length λ.

The coordinates of beam electrons are determined by the relativistic equations of motion:

$$\frac{dz_j}{dt}=v_j,\quad \frac{d}{dt}\left(\frac{v_j}{\sqrt{1-v_j^2/c^2}}\right)=\frac{e}{m}E_z, \tag{10}$$

where v_j is the velocity of the jth electron and E_z is defined by the third line in Eq. (1). As concerns plasma electrons, we apply the approach commonly used in microwave electronics and describe them in the linear approximation.[11] Therefore the perturbation of the plasma electron density is described by the second line in Eq. (1).

The equations of motion (10) are supplemented with the conditions for electron-beam injection across the cross section $z=0$. Let us denote the injection time of the jth electron (large-size particle) by t_{0j}. Given that the electrons in the injected beam move at a fixed velocity, we have the following injection conditions:

$$z_j|_{t=t_{0j}}=0,\quad v_j|_{t=t_{0j}}=u. \tag{11}$$

If the electron beam injection begins at $t=0$, then $t_{0j}>0$ for all j. It is convenient to select the time interval between sequential inputs of large-size particles into the waveguide as a constant defined by the following relation:

$$t_{0j+1}-t_{0j}=\frac{2\pi}{N\omega},\quad j=1,2,\ldots, \tag{12}$$

where ω is the average frequency of the output signal. From this equation and the second line in Eq. (11), we derive the characteristic length λ in Eq. (10):

$$\lambda=2\pi u/\omega.$$

In order to include the shapes of edges of beam current pulses, it is convenient to introduce weighting factors $q(t_{0j})$ in the sum in Eq. (9). If $q(t_{0j})=0$, the leading edge has not arrived yet, and if $q(t_{0j})=1$ the beam current has achieved its stationary value. The shape of the function $q(t_{0j})$ was determined from experimental oscilloscope traces.[1,2,7]

Let us express the polarization potential in the waveguide containing both the plasma and electron beam as

$$\Psi(z,t,r_\perp)=\frac{1}{2}\sum_{n=1}^{\infty}\varphi_n(r_\perp)[A_n(z,t)\exp(-i\omega t+ik_z z) + B_n(z,t)\exp(-i\omega t-ik_z z)+\text{c.c.}]. \tag{13}$$

Here the terms with A_n correspond to the wave resonantly excited by the electron beam and propagating from $z=0$ to $z=L$. The terms with B_n correspond to the reflected wave propagating in the opposite direction and responsible for the feedback in the oscillator. On average, the interaction between the reflected wave and the electron beam vanishes.

Since the resonant excitation of the plasma wave by the electron beam occurs through the Cherenkov effect (either single-particle or collective[12]) and in accordance with the definition of λ, the wave vector in Eq. (13) must be equated to $k_z=\omega/u$.

Equation (13) contains slow wave amplitudes $A_n(z,t)$ and $B_n(z,t)$. These are not the amplitudes of different transverse waveguide modes, since only one mode, namely the surface one, is excited. This mode is determined by infinite sums of φ_nA_n and φ_nB_n in Eq. (13). These amplitudes are slow in the sense of our basic assumptions, namely the average spectral frequency ω and the spectral width $\Delta\omega\ll\omega$. Therefore the differential operators are

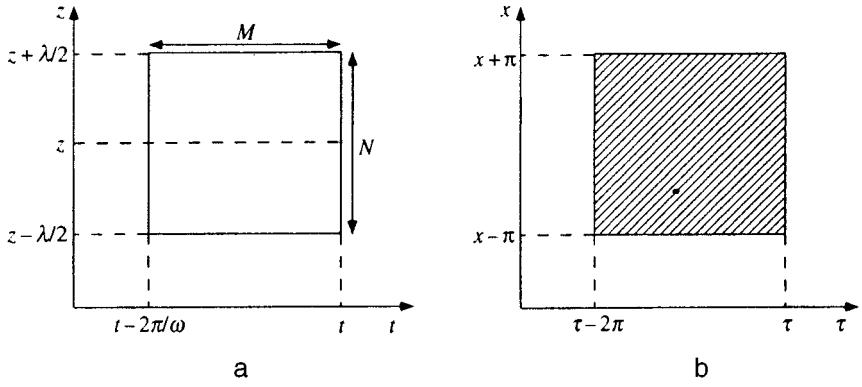

FIG. 2. Diagram illustrating the averaging procedure.

$$\frac{\partial}{\partial t}=-i\omega+\frac{\tilde{\partial}}{\partial t},\quad \frac{\partial}{\partial z}=\pm ik_z+\frac{\tilde{\partial}}{\partial z},$$

where the tilde means that only the slow amplitude is differentiated, the plus in the expression for $\partial/\partial z$ refers to the resonantly coupled wave, and the minus to the reflected wave. The slow changes in the amplitudes mean that

$$\left|\frac{1}{\omega A_n}\frac{\tilde{\partial}A_n}{\partial t}\right|\ll 1,\quad \left|\frac{1}{k_z A_n}\frac{\tilde{\partial}A_n}{\partial z}\right|\ll 1. \tag{14}$$

Similar inequalities apply to the amplitudes B_n. In what follows, the tilde in the derivatives of slow amplitudes is omitted for brevity.

By analogy with Eq. (13), the perturbation of the plasma charge density derived from the second line in Eq. (1) is expressed as

$$\rho_p=\frac{1}{2}[\rho_p^+\exp(-i\omega t+ik_z z)+\rho_p^-\exp(-i\omega t-ik_z z)+\text{c.c.}], \tag{15}$$

where the first term with the amplitude ρ_p^+ describes the resonantly coupled wave, and the second term with ρ_p^- the reflected wave. Since the plasma is infinitely thin, the amplitudes $\rho_p^\pm$ are defined only at a fixed point in the waveguide cross section, namely at $\mathbf{r}_\perp=\mathbf{r}_p$. Therefore there is no need to expand the functions $\rho_p^\pm$ in terms of the waveguide eigenfunctions $\varphi_n(\mathbf{r}_\perp)$.

Substituting Eqs. (13) and (15) into Eq. (8) and averaging over fast variables (see below), we obtain, with due account of Eq. (14), the expressions for the amplitudes A_n and B_n:

$$A_n=\hat{L}_n^{-1}4\pi(G_{pn}\rho_p^+ + en_b G_{bn}\langle\rho_b\rangle),\qquad B_n=\hat{L}_n^{-1}4\pi G_{pn}\rho_p^-. \tag{16}$$

Here $G_{\alpha n}=S_\alpha\varphi_n(\mathbf{r}_\alpha)/\|\varphi_n\|^2$, $\alpha=p$, b are geometrical factors taking into account the positions of the plasma and electron beams in the waveguide cross section, and

$$\hat{L}_n^{-1}=\pm ik_z\left(1+\frac{i}{k_z}\frac{\partial}{\partial z}\right)\frac{1}{k_{\perp n}^2+k_z^2\gamma^{-2}}\times\left[1+\frac{2ik_z^2}{k_{\perp n}^2+k_z^2\gamma^{-2}}\left(\pm\frac{1}{k_z}\frac{\partial}{\partial z}+\frac{\beta^2}{\omega}\frac{\partial}{\partial t}\right)\right] \tag{17}$$

is the inverse operator of the nth harmonic of the polarization potential ($\beta=u/c$). In deriving Eq. (17), we have taken account of Eq. (14).

The averaging procedure which has been used in deriving Eq. (16) is described by[8]

$$\langle\Phi\rangle=\frac{1}{\lambda}\int_{z-\lambda/2}^{z+\lambda/2}\Phi(z')\exp(i\omega t\mp ik_z z')dz', \tag{18}$$

where Φ is the averaged function. Therefore $\langle\rho_b\rangle$ in Eq. (16) is defined, obviously, by the formula

$$\langle\rho_b\rangle=\frac{2}{N}\sum_j q(t_{0j})\exp(i\omega t-ik_z z_j). \tag{19}$$

The summation in Eq. (19) is performed over all large-size particles on the interval $(z-\lambda/2,z+\lambda/2)$. The resulting value is assigned to the coordinate z.

In order to get rid of noise due to quantization in numerical calculations, it is worthwhile to perform additional averaging over time. Consider a time interval $(t-2\pi/\omega,t)$, where t is the current time. Divide this interval into M sections and calculate the parameter defined by Eq. (19) on each section. The averaged value $\langle\rho_b\rangle$ is defined as the arithmetic mean over all these sections. The averaging procedure is illustrated by Fig. 2a. The resulting value is ascribed to the point (t,z). At intermediate points the function is defined by cubic splines.

Substitute the coefficients A_n and B_n into the second line of Eq. (1) and the second line of Eq. (10) (with due account of Eq. (13), definition of E_z, and that the averaged interaction between the reflected wave and electron beam is zero). After simple but cumbersome calculations, we obtain the basic equations of this study:

$$\left(\frac{\partial}{\partial x}+\frac{1}{v_g}\frac{\partial}{\partial\tau}+i\frac{\eta_0}{1+f_p}\right)a^+=-\alpha_b\tilde{\alpha}\left[\frac{i}{2\gamma^2(1+f_p)}+\frac{1+f_G}{1+f_p}\left(\frac{\partial}{\partial x}+\beta^2\frac{\partial}{\partial\tau}\right)\right]\langle\rho_b\rangle,$$
$$\left(\frac{\partial}{\partial x}-\frac{1}{v_g}\frac{\partial}{\partial\tau}-i\frac{\eta_0}{1+f_p}\right)a^-=0, \tag{20}$$

$$\langle \rho_b(x,\tau)\rangle = \frac{2}{MN}\sum_j q(\tau_{0j})\exp(i\tau - ix_j)\,\theta(x,\tau),$$

$$\frac{dx_j}{d\tau} = y_j, \quad \frac{d}{d\tau}\left(\frac{y_j}{\sqrt{1-\beta^2 y_j^2}}\right)$$
$$= -\frac{i}{2}\gamma^3\left\{\left[1 - i2\gamma^2(1+f_G)\left(\frac{\partial}{\partial x}+\beta^2\frac{\partial}{\partial\tau}\right)\right]\right.$$
$$\times a^+ + \alpha_b\left[1 - i2\gamma^2(1+f_b)\left(\frac{\partial}{\partial x}+\beta^2\frac{\partial}{\partial\tau}\right)\right]\langle\rho_b\rangle\Bigg\}$$
$$\times\exp(-i\tau + ix_j) + \text{c.c.}$$

Here we have introduced the dimensionless time $\tau=\omega t$, coordinate $x=k_z z$, and velocity $y=v/u$. The explicit expressions for the dimensionless amplitudes $a^{\pm}$ of both the resonant and reflected waves is immaterial because they are expressed in relative units in what follows. The function $\theta(x,\tau)$ in the expression for $\langle\rho_b\rangle$ is unity in the hatched area in Fig. 2b and zero outside this area. Furthermore, Eq. (20) contains the following notation:

$$\alpha_p = \frac{\omega_p^2}{k_{\perp p}^2 u^2\gamma^2}, \quad \alpha_b = \frac{\omega_b^2\gamma^{-3}}{k_{\perp b}^2 u^2\gamma^2},$$

$$\tilde{\alpha} = \left[\sum_{n=1}^{\infty}\frac{1}{k_{\perp n}^2+\chi^2}\frac{\varphi_n(\mathbf{r}_b)\varphi_n(\mathbf{r}_p)}{\|\varphi_n\|^2}\right]^2 S_p k_{\perp p}^2 S_b k_{\perp b}^2, \qquad (21)$$

$$\eta_0 = \frac{1}{2\gamma^2}\left(1-\frac{1}{\alpha_p}\right), \quad v_g = \left(\beta^2 + \frac{1}{\alpha_p\gamma^2(1+f_p)}\right)^{-1}.$$

Here $k_{\perp p}^2$ is the parameter introduced in Eq. (3), $k_{\perp b}^2$ is a similar parameter with the index p replaced with b, $\tilde{\alpha}$ is the coupling constant between the beam and plasma waves (overlap integral for their fields[12]), η_0 is the detuning of the wave frequency from the exact Cherenkov resonance,[1)] v_g is the ratio between the plasma-wave group velocity and u. In Eq. (21) χ is defined by the formula $\chi=\omega/u\gamma$.

The frequency dependence of the parameters defined by Eq. (21) is due to the nonlinear dispersion of the beam and plasma waves. This dispersion also determines the parameters f_p, f_b, and f_G in Eq. (20). As $\omega\to 0$, they tend to zero, and if $\chi R>1$, they tend to $-1/2$. They are important in the frequency band where the spectrum is defined by Eq. (6). The explicit expressions for these parameters are not given in this paper because they are too cumbersome.

In a circular waveguide, when the beam is a thin-wall tube with the radius r_b and thickness Δ_b, the coupling constant $\tilde{\alpha}$ has the form (here, for definiteness, $r_p>r_b$)

$$\tilde{\alpha} = \frac{I_0(\chi r_b)[K_0(\chi r_p)I_0(\chi R) - K_0(\chi R)I_0(\chi r_p)]}{I_0(\chi r_p)[K_0(\chi r_b)I_0(\chi R) - K_0(\chi R)I_0(\chi r_b)]}. \qquad (22)$$

If $r_b=r_p$, the coupling constant $\tilde{\alpha}=1$. The factor $\tilde{\alpha}$ drops with the difference r_p-r_b, and this decrease is larger at higher frequencies. In the high-frequency band

$$\tilde{\alpha} = \exp\left(-2\frac{\omega}{u\gamma}|r_p - r_b|\right).$$

Equations (20) are supplemented with additional boundary conditions (feedback conditions):

$$a^-(x=\omega L/u,\tau) = \kappa a^+(x=\omega L/u,\tau),$$
$$a^+(x=0,\tau) = -a^-(x=0,\tau), \qquad (23)$$

where the reflectivity κ is determined by Eq. (7), and the minus in the second line corresponds to the phase reversal due to reflection from a metallic surface. In addition, the initial values $a^{\pm}(x,\tau=0)=0$ and electron-beam injection conditions (11) are required. To model the excitation mechanism of electromagnetic waves, we assumed that the input electron beam had a small ($\sim 1\%$) random modulation of relativistic momentum (in another version, a small random density modulation).

The generation efficiency is estimated using relative kinetic losses in the beam described by the formula

$$\eta = \frac{W_{\text{in}} - W_0 - W_{\text{out}}}{W_{\text{in}}}. \qquad (24)$$

Here W_{in} is the energy of all electrons injected into the waveguide by time τ, W_{out} is the energy of all electrons that have passed across the cross section $z=L$ by time τ, W_0 is the energy of all electrons in the waveguide ($0<z<L$) at the moment τ.

5. Let us analyze the equation system (20) in the linear approximation. It will be sufficient for our analysis to consider only the resonant wave. One can prove that in the linear approximation the equation for $\langle\rho_b\rangle$ has the form

$$\left(\frac{\partial}{\partial\tau}+\frac{\partial}{\partial x}\right)^2\langle\rho_b\rangle$$
$$= -\left[1 - i2\gamma^2(1+f_G)\left(\frac{\partial}{\partial x}+\beta^2\frac{\partial}{\partial\tau}\right)\right]a^+$$
$$-\alpha_b\left[1 - i2\gamma^2(1+f_b)\left(\frac{\partial}{\partial x}+\beta^2\frac{\partial}{\partial\tau}\right)\right]\langle\rho_b\rangle, \qquad (25)$$

and the equation for a^+ is the same as in the system (20).

Let us suppose that $\partial/\partial\tau=0$, which corresponds to amplification of a steady signal at fixed frequency, and let us seek a solution in the form $\langle\rho_b\rangle, a^+ \propto \exp(i\delta x)$. From Eq. (25) and the first equation in (20) we derive the dispersion relation

$$\left(\delta + \frac{\eta_0}{1+f_p}\right)[\delta^2 - \alpha_b(1 + 2\gamma^2(1+f_b)\delta)]$$
$$= -\frac{\tilde{\alpha}\alpha_b}{2\gamma^2(1+f_b)}[1 + 2\gamma^2(1+f_G)\delta]^2. \qquad (26)$$

Given the definition of the dimensionless coordinate x and Eq. (15) for the perturbation of the plasma charge density, one can see that the total longitudinal wave number in the plasma–beam system is determined by the relationship

$$k_{\parallel} = k_z(1+\delta) = \frac{\omega}{u}(1+\delta). \qquad (27)$$

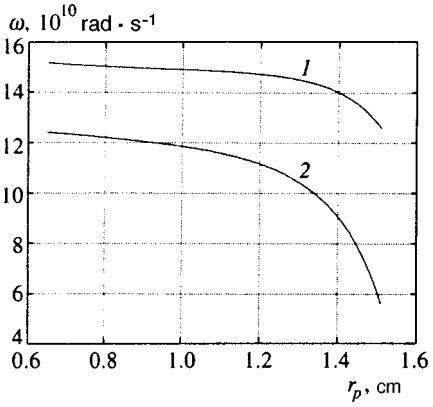

FIG. 3. Resonant frequencies versus plasma radius: (*1*) Raman resonance; (*2*) Cherenkov resonance. The electron beam radius is 0.65 cm.

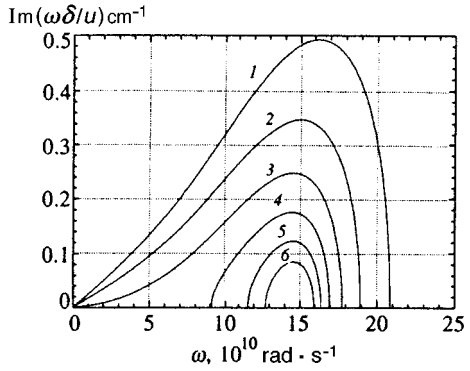

FIG. 4. Spatial instability growth rate versus frequency for several plasma radii: (*1*) $r_p=0.7$; (*2*) 0.8; (*3*) 0.9; (*4*) 1.0; (*5*) 1.1; (*6*) 1.2 cm.

The imaginary part of Eq. (27) determines the gain of oscillations.

In the low-frequency limit, Eq. (26) transforms to the equation described in detail in Ref. 13. We are not going to investigate Eq. (26) by analytic techniques because one can hardly derive simple and instructive analytic solutions for the interesting parameters of the plasma–beam system. The parameters of the system are real experimental values[2]: the beam current $J_B^{\max}=2$ kA, the beam radius $r_b=0.65$ cm, the beam thickness $\Delta_b=0.1$ cm, the plasma frequency $\omega_p=35\times10^{10}$ rad/s, the beam velocity $u=2.6\times10^{10}$ cm/s ($\gamma=2$). In the reported study, we varied the waveguide length, $L=10$, 15, 20 cm, and the plasma radius r_p.

The dispersion relation (26) is necessary for deriving formulas important for further analysis. By equating the first factor in parentheses on the left-hand side of Eq. (26) to zero,

$$\delta=-\frac{\eta_0}{1+f_p}, \tag{28}$$

we can determine the wave number of the resonant surface plasma wave. This can be easily proved by substituting Eq. (27) into Eq. (3) and expanding Eq. (3) in powers of δ. It follows from Eq. (27) that at $\delta=0$ we have $k_\parallel=k_z=\omega/u$, which denotes exact equality between the beam velocity u and phase velocity of the plasma wave, i.e., the Cherenkov resonance. The Cherenkov resonance frequency is derived from the equation $\eta_0=0$. This frequency is plotted against r_p in curve *2* of Fig. 3 (the system parameters were quoted above). The amplification at the Cherenkov resonance frequency is called amplification due to the single-particle Cherenkov effect.[12]

Zeros of the second factor on the left-hand side of Eq. (26) determine the longitudinal wave numbers of the beam charge-density waves, the fast and slow ones.[14] For the slow wave we have

$$\delta=\alpha_b\gamma^2(1+f_b)+\frac{1}{2}\sqrt{[\alpha_b 2\gamma^2(1+f_b)]^2+4\alpha_b}. \tag{29}$$

The frequency at which the phase velocities of the plasma wave and slow beam wave are equal is called the Raman resonance frequency.[2)] The Raman resonance frequency is determined by solving the equation system (28), (29). This frequency is plotted against r_p in curve *1* of Fig. 3. The amplification at the Raman resonance frequency is called amplification due to the Raman effect or, more often, collective Cherenkov effect.[14,12] It is reasonable to select one of the frequencies mentioned above for the frequency in Eqs. (13), (15) (hence in Eq. (20)), and (21) and in the formula for the reflectivity κ.

The question of the frequency selection can be answered by investigating numerical solutions of Eq. (26). Figure 4 shows the parameter $\mathrm{Im}(\omega\delta/u)$, i.e., the growth rate calculated at different plasma radii r_p, as a function of frequency. Figures 3 and 4 indicate that the growth rate maximum occurs at the Raman resonance frequency, whereas the Cherenkov resonance frequency shows no features. Moreover, at $r_p>1.1$ cm, the growth rate is zero at the Cherenkov resonance frequency. Therefore, in what follows we substitute into Eqs. (20), (21), and (26) the Raman resonance frequency, since excitation of the oscillator at this frequency is most probable. Otherwise the spectra derived numerically from Eq. (20) will show a corresponding deviation.

Now let us clearly define the terms. The gain band is the frequency band where the growth rate is nonzero (Fig. 4). The regimes in which the growth rate occurs at zero frequency are termed Compton. The regimes in which the growth rate does not extend to zero are called Raman. When the plasma is brought away from the electron beam, the Compton regime of amplification and generation transforms to the Raman regime. It is noteworthy, however, that for parameters of real experiments, the Compton and Raman regimes as the limiting cases investigated in detail previously[9,12,13] cannot exist. We operate actually in the intermediate region of parameters, which is most difficult from the theoretical viewpoint.

6. Let us consider numerical solutions of Eq. (20). The parameters of the beam current pulse are derived from experimental data[1,2]: the full width is 37 ns, the leading edge width is 10 ns, and the trailing edge width is 14 ns.

The basic characteristics of the plasma oscillator given in this paper are the following: the output power as a function of time at different plasma radii; the oscillator response time as a function of the plasma radius; the total relative kinetic losses in the beam during the entire injection time at various waveguide lengths; the distribution of mode ampli-

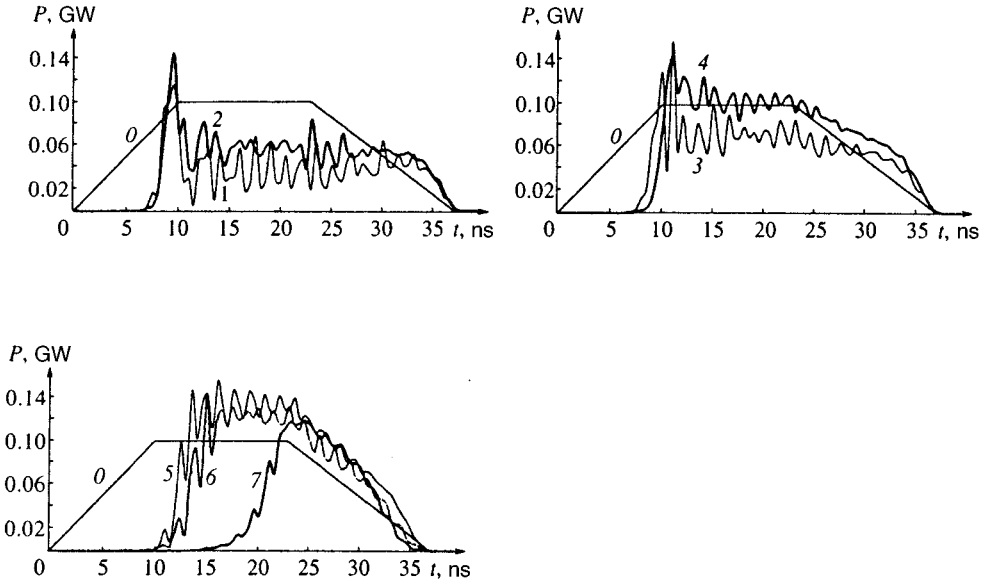

FIG. 5. Output power versus time for $L=15$ cm: (O) shape of the electron current pulse $P_{\text{beam}}/10$; (*1*) $r_p=0.79$; (*2*) 0.81; (*3*) 0.83; (*4*) 0.85; (*5*) 0.89; (*6*) 0.91; (*7*) 0.95 cm.

tudes over the waveguide length and phase planes of the beam electrons at the moment before the current pulse trailing edge; the output spectra during the oscillator operation.

The response time is defined as the time when the efficiency reaches about 0.001 (the injection onset occurs at $t=0$). The output spectrum is given by

$$S^2(\Omega)=\tilde{a}\,\tilde{a}^*, \quad \tilde{a}=\int_0^\infty a^+(\omega L/u,\tau)e^{i\Omega\tau}d\tau. \tag{30}$$

Here $\Omega=\tilde{\omega}/\omega$, where ω is the Raman resonance frequency, and $\tilde{\omega}$ is the output frequency.

Note that at some arbitrary time, the parameter η defined by Eq. (24) is not, strictly speaking, the generation efficiency. But after the current pulse leading edge, when η is constant, it determines the lost fraction of the beam energy. This loss is due to the following drains: plasma heating, reflection of electrons in the region $z<0$, and radiation. In our model, the plasma heating is neglected, and in real conditions it is not essential. Numerical calculations indicate that there are no reflected electrons. Thus, the energy is lost only to radiation.

Figure 5 shows the output power and beam current pulse shape for $L=15$ cm and several plasma radii. One can see that the oscillator response time increases with the separation between the electron and plasma beams. At $r_p=1$ cm the oscillator does not operate because the current pulse is too short. If the current pulse were longer, the oscillator could operate even at this plasma radius.[3)]

Figure 6a shows the oscillator response time as a function of the plasma radius for several waveguide lengths L. In the Compton regime, when the separation between the beam and plasma is small, the response time increases with the plasma radius slowly. In the Raman regime this increase is notably faster. The calculated response times in the range of 5–30 ns are in agreement with experimental data.[7]

Figure 6b shows the resulting radiation efficiency η for different oscillator lengths. Each system length has a corresponding optimal configuration of the electron and plasma beams in the cross section at which the energy loss to radiation of the electron beam is maximum. The efficiency maximum decreases with the oscillator length and shifts to the range of large separations between the plasma and electron beam.

At each fixed length L, the oscillator has certain features. There is an optimal plasma radius, at which the radiation efficiency peaks. At larger and smaller r_p the efficiency is lower. The drop in the efficiency with r_p has a simple explanation: the oscillator response time increases because of the lower coupling constant, and only a fraction of the beam

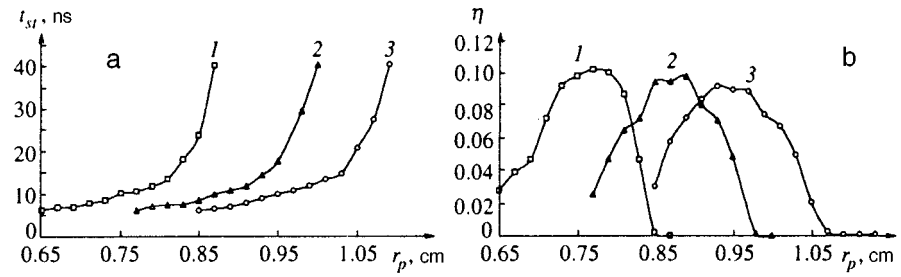

FIG. 6. (a) Response time of the oscillator as a function of the plasma radius for several tube lengths: (*1*) $L=10$; (*2*) 15; (*3*) 20 cm. (b) Oscillator efficiency as a function of the plasma radius at several tube lengths: (*1*) $L=10$; (*2*) 15; (*3*) 20 cm.

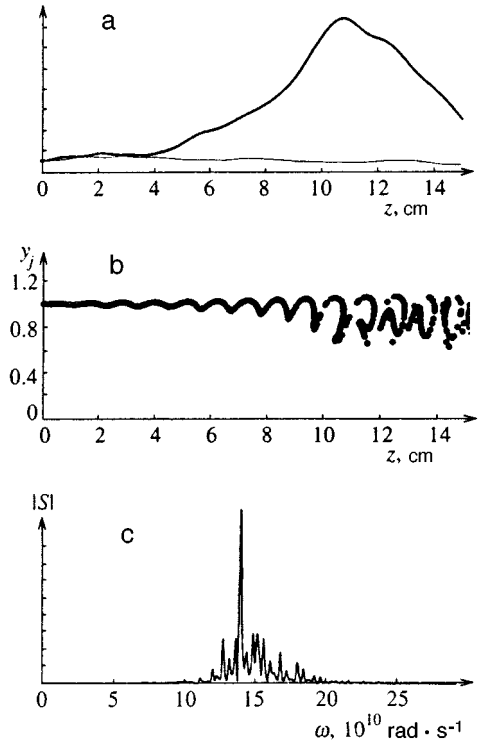

FIG. 7. (a) Amplitude distribution of the resonant (thick line) and reflected (thin line) waves at $t=22$ ns for $L=15$ cm and $r_p=0.79$ cm. (b) Phase plane of the electron beam at $t=22$ ns for $L=15$ cm and $r_p=0.79$ cm. (c) Spectrum of the resonant plasma wave output radiation for $L=15$ cm and $r_p=0.79$ cm.

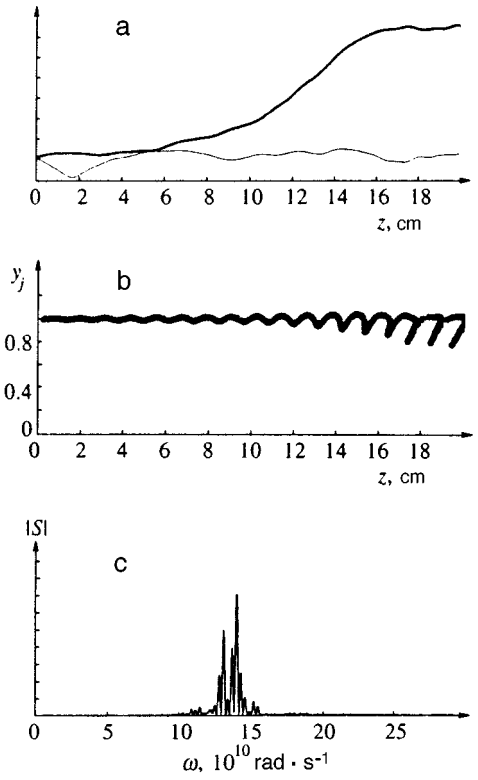

FIG. 8. (a) Amplitude distribution of resonant (thick line) and reflected (thin line) waves at the moment $t=22$ ns for $L=20$ cm and $r_p=1.03$ cm. (b) Phase plane of the electron beam at the moment $t=22$ ns for $L=20$ cm and $r_p=1.03$ cm. (c) Spectrum of resonant plasma wave output radiation for $L=20$ cm and $r_p=1.03$ cm.

current is utilized. At very large r_p the pulse is even shorter than the response time.

The efficiency decrease with decreasing r_p also has an explanation. The generation parameters approach those of the Compton regime as the plasma radius drops. It is known that in the Compton regime the coupling between the electron beam and plasma is stabilized due to capture of beam electrons by the plasma wave field.[15] The captured beam is divided into bunches, which oscillate in potential wells created by the plasma wave, and they sometimes absorb energy from the wave, sometimes emit their energy into it. The efficiency is maximum at such a plasma radius that the electrons are captured, with due account of the feedback effects, near the output coupler, where they leave the oscillator. At a smaller plasma radius, the coupling between the electron beam and plasma is stronger, the capture point shifts inside the system, and electrons leave the system in the phases of post-capture oscillations, when they absorb energy of the plasma wave. For this reason, the oscillator efficiency drops.

Figure 7a shows the distribution of the amplitudes of the resonant (thick line) and reflected (thin line) waves at a fixed time $t=22$ ns, $L=15$ cm, and $r_p=0.79$ cm. The phase plane for the beam electron at the same moment and same oscillator parameters is shown in Fig. 7b. The ordinate is the dimensionless velocity $y_j=v_j/u$. Finally, Fig. 7c shows the output radiation spectrum for this case. Figure 4 indicates that at $r_p=0.79$ cm the oscillator operates in the Compton regime. It is characterized by post-capture oscillations of the beam electrons in the field of the resonant plasma wave, as shown by the phase trajectory in Fig. 7b. Figure 7a shows that electrons are captured inside the oscillator cavity at $z\sim 11$ cm. These parameters correspond to the nonoptimal Compton regime. The fast capture of the electron beam, drag by the plasma wave, oscillations, and destruction of electron bunches lead to the wide output spectrum (Fig. 7c). Most of the radiation is emitted in the band between the Cherenkov (12×10^{10} rad/s) and Raman (15×10^{10} rad/s) resonance frequencies.

Figure 8a–c show the curves similar to those in Fig. 7a–c, but at $L=20$ cm and $r_p=1.0$ cm, for which the oscillator parameters are close to those of the Raman regime. The lack of significant spatial oscillations of the resonant wave amplitude due to electron capture is evident. The nonlinear stabilization of the plasma–beam instability in this case is due to the reversal of the beam wave,[9] which can be seen in the phase plane of the beam electrons (Fig. 8b). The absence of oscillations of the resonant plasma wave results in a narrower output spectrum (Fig. 8c).

In our opinion, no reliable data about the spectra of plasma microwave oscillators have been published to this time, so we do not compare the details of calculated and experimental spectra. This issue can be reconsidered when new results of the experiment which is currently under way are available.

[1)]We assume that $k_z=\omega/u$ in Eq. (3). We then obtain $\alpha_p=1$ or $\eta_0=0$.

[2)]In the literature this mechanism is also termed the anomalous Doppler effect, which entails interaction between waves with energies of opposite sign.[14]

[3)]The start condition (current) of the oscillator onset is known from the theory of plasma microwave oscillators.[9] This condition approximately expressed by the formula $\mathrm{Im}(\omega\delta L/u)>\ln(3/\kappa)$ is satisfied by the parameters of our model. We do not apply this condition for two reasons: first, the current pulse is short, second, the onset condition was derived for limiting cases and does not apply to the specific parameters of our model, but it should yield the desired parameters within an order of magnitude.

[1] M. V. Kuzelev, F. Kh. Mukhametzyanov, M. S. Rabinovich *et al.*, Zh. Éksp. Teor. Fiz. **83**, 1358 (1982) [Sov. Phys. JETP **56**, 780 (1982)]; Dokl. Akad. Nauk SSSR **267**, 829 (1978).

[2] M. V. Kuzelev, O. T. Loza, A. V. Ponomarev *et al.*, Zh. Éksp. Teor. Fiz. **109**, 2048 (1996) [JETP **82**, 1102 (1996)].

[3] M. V. Kuzelev, A. A. Rukhadze, Yu. V. Bobylev *et al.*, Zh. Éksp. Teor. Fiz. **91**, 1620 (1986) [Sov. Phys. JETP **64**, 956 (1986)].

[4] T. D. Pointon and J. S. De Groot, Phys. Fluids **31**, 908 (1988).

[5] M. V. Kuzelev, A. R. Maikov, A. D. Poezd *et al.*, Dokl. Akad. Nauk SSSR **300**, 1112 (1988).

[6] A. R. Maikov, A. G. Sveshnikov, and S. A. Yakunin, Zh. Vych. Mat. Mat. Fiz. **25**, 883 (1984).

[7] A. G. Shkvarunets, A. A. Rukhadze, and P. S. Strelkov, Fiz. Plasmy **20**, 682 (1994) [Plasma Phys. Rep. **20**, 613 (1994)].

[8] M. A. Krasil'nikov, M. V. Kuzelev, and A. A. Rukhadze, Zh. Éksp. Teor. Fiz. **108**, 521 (1995) [JETP **81**, 280 (1995)].

[9] M. V. Kuzelev and A. A. Rukhadze, *Electrodynamics of Dense Electron Beams in Plasma* [in Russian], Nauka, Moscow (1990).

[10] B. Z. Katsenelembaum, *High Frequency Electrodynamics* [in Russian], Nauka, Moscow (1966).

[11] A. A. Ivanov, *Physics of Highly Nonequilibrium Plasma* [in Russian], Atomizdat, Moscow (1977).

[12] M. V. Kuzelev and A. A. Rukhadze, Usp. Fiz. Nauk **152**, 285 (1987) [Sov. Phys. Usp. **30**, 507 (1987)].

[13] A. F. Aleksandrov, M. V. Kuzelev, and A. N. Khalilov, Fiz. Plasmy **14**, 455 (1988) [Sov. J. Plasma Phys. **14**, 267 (1988)].

[14] M. V. Nezlin, Usp. Fiz. Nauk **120**, 181 (1976).

[15] V. D. Shapiro and V. I. Shevchenko, Izv. Vyssh. Uchebn. Zaved.: Radiofiz. **19**, 767 (1976).

Translation provided by the Russian Editorial office.

Mechanisms of resonant laser ionization

A. G. Leonov and D. I. Chekhov

Moscow Physicotechnical Institute, 141700 Dolgoprudnyi, Moscow Region, Russia

A. N. Starostin

Troitsk Institute of Innovative and Thermonuclear Research, 142092 Troitsk, Moscow Region, Russia
(Submitted 12 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1274–1296 (April 1997)

An experimental investigation and numerical simulation of resonant laser breakdown are performed. As a result, quantitative agreement between the experimental data on the parameters of a dense resonant plasma (the electron density and the electron temperature) and the results of calculations in the range of detunings of the laser radiation from resonance $\Delta\lambda > 2-2.5$ nm, in which the spatial instability of the intense resonant laser beam and the absorption of radiation are minimal, is obtained for the first time. It is shown that the previously proposed mechanism of resonant breakdown associated with laser-induced associative ionization introduces only a small correction to the final extent of ionization of the resonant plasma and scarcely alters its temperature. The influence of quantum stimulated inverse bremsstrahlung processes, which are usually described as collisions of the second kind in the resonance case, on the energy gain by electrons is analyzed for the first time in reference to specific experimental findings. The numerical calculations show that at detunings of the order of the Rabi frequency, the mechanism by which electrons gain energy through the resonant system does not reduce to collisions of the second kind and can significantly increase the density of the resonant plasma. However, in this range of detunings the laser beam is still strongly perturbed by instability processes, precluding a proper comparison of the theory with experiment. At large $\Delta\lambda$ the classical and quantum cases differ from one another only slightly, and the values of N_e calculated for both mechanisms lie within the measurement error. © *1997 American Institute of Physics.* [S1063-7761(97)01204-3]

1. INTRODUCTION

The resonant laser plasma produced when laser radiation with a quantum energy close to the energy of some resonant atomic or molecular transition acts on a gaseous medium has attracted attention for a long time and has been thoroughly investigated in many studies (see, for example, the reviews in Refs. 1–3 and the references cited therein). This is largely because the resonant plasma, which has a high density $N_e \simeq 10^{15}-10^{17}$ cm^{-3} at a relatively low electron temperature $T_e \simeq 0.2-0.5$ eV, is a unique physical system and, as was noted in Ref. 4, the achievement of such parameters by other methods is infeasible in practice.

The broad range of the parameters of experiments in which the formation of a resonant laser plasma has been studied, viz., the density of the medium (which generally consists of an alkali-metal or alkaline-earth-metal vapor), the intensity and duration of the laser pulses, the energy spectra of the atoms irradiated, etc., has greatly hindered the formulation of a general theory for the phenomenon under consideration, and there are presently several models of resonant breakdown.

The model proposed in Refs. 5 and 6 is most generally accepted for the case of high vapor density and pulsed laser action. According to Refs. 5 and 6, in the first stage of breakdown the resonant transition becomes saturated by laser radiation, and initial ionization appears in processes involving excited atoms, viz., associative ionization, multiphoton (usually two-quantum) ionization of a resonant level, and laser-induced collisions. The seed electrons then quickly gain energy in superelastic collisions with resonantly excited atoms of the medium, quickly populating higher-lying levels and ionizing them. A certain auxiliary role is again played here by the laser radiation, which accelerates the rise in electron density by causing the one-quantum photoionization of the highly excited levels. Ultimately the development of electron avalanche leads to essentially complete ionization of the medium in a narrow channel along the laser beam.

We note that as was shown theoretically in Refs. 2 and 7, in a strong resonant field the rate at which electrons gain energy in quenching collisions can differ significantly from the rate postulated in Refs. 5 and 6.

Besides the mechanism just considered, other mechanisms for the formation of a resonant plasma have been proposed in several papers. In particular, it was concluded in Refs. 8 and 9 on the basis of an analysis of numerous experimental data that laser-induced associative ionization involving resonantly excited atoms plays a major role in the development of breakdown.

It follows from the foregoing that for resonant breakdown of a medium in any case the intensity of the laser radiation need be only of the order of magnitude of the saturation intensity, which is low in metal vapors ($\mathcal{I}_s \simeq 10-1000$ W/cm^2) as a consequence of the large value of the dipole moment and the small width of the resonant transition (of course, in real experiments because of the great

amount of absorption, as well as the finite time needed for the development of electron avalanche, significantly greater intensities, $\mathscr{I} \simeq 10^5 - 10^8$ W/cm^2 with a pulse width of $10^{-8} - 10^{-6}$ s, are needed to create a plasma channel with a length of $\sim 1-10$ cm in a vapor with a density of $10^{15} - 10^{17}$ cm^{-3}). In the case of nonresonant breakdown, in which the mechanism for heating the electrons is determined by inverse bremsstrahlung, the threshold intensities are known to be much greater: $10^9 - 10^{11}$ W/cm^2.[10] In the case of resonant breakdown, the role of traditional inverse bremsstrahlung is negligible.[11]

Although various models of resonant laser ionization qualitatively account for many experimental findings, the calculations performed on their basis[4–6,11–14] are consistent with the experimental data only in order of magnitude. It should, however, be noted that there is no detailed comparison of the results of theory and specific experiments in the literature. On the other hand, the setting up of experiments, especially in dense vapors, greatly impedes their unequivocal interpretation. To a considerable extent, this is because the nonlinear interaction of intense laser radiation with a dense resonant medium causes not only ionization, but also several other well known phenomena, particularly spectral and spatial instability of the laser wave (see, for example, Refs. 15 and 16 and the references cited therein). As a result of such instability, the divergence of the laser beam increases considerably and its spatial structure becomes highly distorted, making the intensity of the radiation, which largely determines the ionization kinetics of the resonant plasma, uncertain. In addition, the presence of instability precludes properly taking into account the strong resonant absorption which the laser beam experiences in a dense medium. All this makes it impossible to properly compare the experimental data with theory and to establish the real contribution of any particular mechanism to the overall pattern of resonant breakdown.

Nevertheless, it is possible to eliminate or, at least, strongly suppress the negative effect of instability, if the frequency of the laser radiation ν_L is shifted relative to the resonant frequency ν_0, since the growth rate of the instability decreases rapidly as the detuning $\Delta\nu = \nu_L - \nu_0$ increases. At the same time, because of the large value of the ratio $\mathscr{I}/\mathscr{I}_s$, it can be expected that the extent of excitation of the medium will still be considerable even when $\Delta\nu$ is fairly large, although the absorption of laser radiation decreases as the detuning increases. In addition, studying the dependence of resonant ionization on the detuning permits variation of the degree of saturation of the medium at a constant intensity of the laser radiation, making it possible to separate the influences of different processes on the formation of the resonant plasma and, in particular, to evaluate the influence of a strong field on the heating rate of the electrons.

Nevertheless, to date there have been essentially no such investigations. To the best of our knowledge, the dependence of the density of a resonant plasma on the detuning in a stationary vapor was measured in only one study.[17] However, in Ref. 17 the width of the resonance curve was of the order of the spectral width of the laser radiation $\Delta\nu_L$, and its intensity was so low that the Rabi frequency was also $\sim \Delta\nu_L$, and thus strong-field effects could not be detected in Ref. 17. In addition, there was no comparison of experiment with theoretical calculations in Ref. 17. The dependence of the plasma density on the detuning from resonance was measured in Refs. 18 and 19 during an investigation of the erosion plasma formed when laser radiation interacts with the surface of a metal target. However, here too the strong nonuniformity of the erosion torch essentially precludes a quantitative interpretation of the experiment.

In the present work, which is devoted to an experimental investigation and theoretical simulation of resonant breakdown at various detunings, an attempt is made to fill this gap. Detailed data on the characteristics of a resonant laser plasma as a function of $\Delta\nu$ have been obtained for the first time in the case of sodium vapor, and we have determined the range of detuning over which the influence of the instability of the laser beam is minimal, but considerable ionization is still observed. In addition, we compare experimental data and the results of numerical calculations, and we use the results to evaluate the roles of the various ionization mechanisms.

Section 2 describes the experimental setup and the diagnostic methods. The results of the experiments are presented in Sec. 3. Section 4 is devoted to a description of a theoretical model of ionization. Finally, Sec. 5 contains the results of numerical simulations, a discussion of these results, and a comparison with the experimental data.

2. EXPERIMENTAL SETUP AND DIAGNOSTIC METHODS

The experiments were carried out using a setup consisting of a tunable dye laser, a heated cell with sodium vapor, and a set of diagnostic instruments (see Fig. 1). The design of the heated cell was similar to that described in Ref. 20 and permitted the creation of a cylindrical column of sodium vapor with a height of 1.5 cm and a density greater than 10^{16} cm^{-3}. Before performing the experiments, the cell was evacuated to 10^{-5} mm Hg and then filled with an inert gas (argon) at a pressure of 190 mm Hg and heated to the required temperature ($T \sim 700$ K). The relatively high argon pressure was necessary to prevent the condensation of sodium vapor on the windows and other cold parts of the cell. The density of the sodium atoms and its distribution over the radius of the cell (Fig. 2) were calculated by taking the Abel transform of data obtained by the Rozhdestvenskiĭ hook method (for further details, see Ref. 21), taking advantage of the cylindrical symmetry of the vapor column. We note that all the experiments were performed at the vapor density N_0 at the maximum of the distribution, which equals 1.8×10^{16} cm^{-3}. The argon density in the hot zone was two orders of magnitude greater, $N_{\mathrm{Ar}} \simeq 2.6\times 10^{18}$ cm^{-3}.

The tunable dye laser was excited by the second-harmonic emission of a YAG:Nd^{3+} laser and generated linearly polarized radiation with a spectral width of 0.05 nm and pulse width $\tau_L = 16$ ns. The pulse-to-pulse variability of the output energy of the dye laser did not exceed a few percent. The laser was tuned to a specific wavelength to within 0.02 nm relative to the standard spectrum of a neon lamp using a DFS-451 grating spectrograph, in whose focal plane there was a photodiode array, which simultaneously

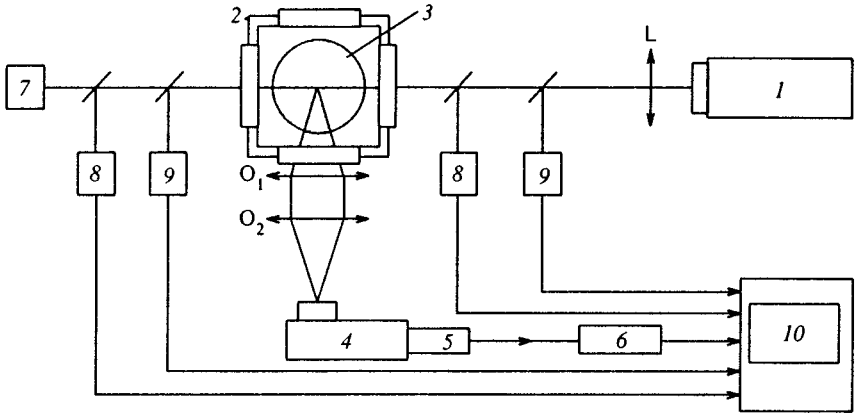

FIG. 1. Experimental setup: *1* — dye laser, *2* — cell, *3* sodium vapor column, *4* — grating monochromator, *5* — FEU-84 photomultiplier, *6* — V9-5 gated voltmeter, *7* — IMO-2N calorimeter, *8* — photodiode array, *9* — FÉK-31KP coaxial photocells, *10* — personal computer.

recorded the spectrum of the laser radiation and the reference spectrum. The laser beam was directed into the heated cell and focused at its center by a long-focus objective lens ($F=150$ cm). The radius of the beam r_L at the half-height of the energy distribution, which was measured by another photodiode array in the focusing region in the cold cell (i.e., in the absence of vapor), was 0.3 mm (see Fig. 3). In the central portion of the cell of radius 3 cm the value of r_L remained unchanged to within the experimental error (± 0.015 mm). We note that in discussing the experimental data below, we shall construe the intensity of the laser radiation $\mathscr{I}_L$ as the quantity calculated from the formula $\mathscr{I}_L=E_L/(\tau_L S_L)$, where E_L is the energy of the laser pulse and $S_L=\pi r_L^2$. The energy distribution over the cross section of the laser beam after passage through the sodium vapor was measured by a third photodiode array 110 cm from the center of the cell. The total power and the energy of the laser radiation at the cell entrance and exit apertures were monitored by coaxial FÉK-31KP photocells and IMO-2N calorimeters.

The parameters of the resonant sodium vapor plasma (electron density and electron temperature) were determined using optical diagnostics from the Stark-broadened profile of the spectral lines and their relative intensity.[22,23] For this purpose, the fluorescence of the plasma at the center of the cell was collected in the direction perpendicular to the axis of the laser beam by two objective lenses, which composed an image of the central fluorescent region at the entrance slit of a grating monochromator (1200 lines/mm). The slit width was set to 20–40 μm in the experiments, and the slit itself (with a height of 1 cm) was oriented parallel to the plasma channel. Considering of the twofold demagnification of the image produced by the projection system, this means that the spatial resolution of the detection system along the channel was 2 cm. We note that the vapor density is essentially constant over this distance (± 1 cm from the center of the cell) (see Fig. 2). The resolution along the height of the channel was 0.04–0.08 mm.

An FEU-84 photomultiplier was positioned behind the exit slit, and its signal was recorded by a V9-5 gated voltmeter with a variable time delay τ_d relative to the beginning of the dye laser pulse. The intensity of the plasma fluorescence remained essentially constant within the grating time (4 ns). To obtain the fluorescence spectra of the plasma, the transmission wavelength of the monochromator was automatically scanned in the vicinity of the spectral lines selected, the photomultiplier output being averaged over 10

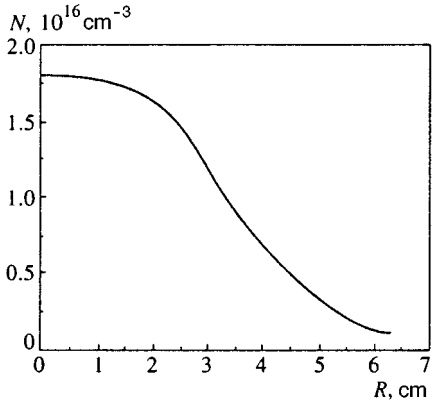

FIG. 2. Distribution of the density of the sodium vapor across the radius of the cell.

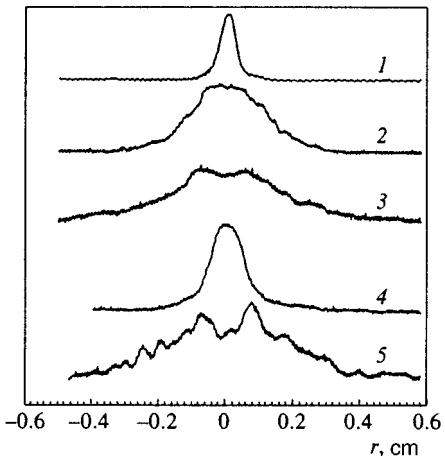

FIG. 3. Energy distribution of the radiation across the radius of the laser beam: *1* — at the center of the cold cell (in the absence of vapor); *2*—*5* — emerging from the heated cell ($N_0=1.8\times 10^{16}$ cm^{-3}) at various detunings (*2* — $\Delta\lambda=3$ nm, *3* — $\Delta\lambda=1.2$ nm, *4* — $\Delta\lambda=0.3$ nm, *5* — $\Delta\lambda=-3.6$ nm); $\mathscr{I}_L=66$ MW/cm^2.

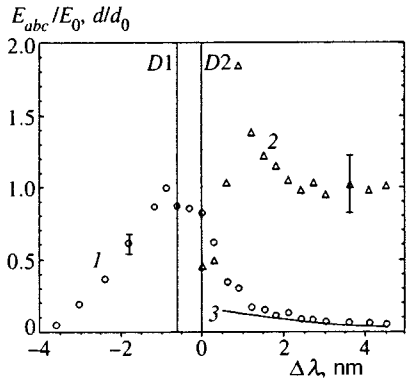

FIG. 4. Dependence of the relative energy of the laser beam E_{abs}/E_0 absorbed in the sodium vapor (*1*, *3*) and of the relative beam diameter d/d_0 emerging from the cell (*2*) on $\Delta\lambda$ (*1*, *2* — experiment, *3* — calculation); $\mathscr{I}_L = 66$ MW/cm^2. The vertical lines mark the positions of the sodium $D1$ and $D2$ lines.

pulses at each point along the spectrum. The spectral resolution of this system reached 0.02 nm.

3. RESULTS OF THE EXPERIMENTAL INVESTIGATION OF A RESONANT LASER PLASMA

3.1. Measurement of the absorption of laser radiation in sodium vapor and spatial structure of the laser beam emerging from the cell

An investigation of the spatial structure of a laser beam with an intensity of some tens of MW/cm^2, and a wavelength close to that of the $D1$ and $D2$ transitions of the sodium atom ($3^2S_{1/2}-3^2P_{1/2}$ and $3^2S_{1/2}-3^2P_{3/2}$, $\lambda_{D1}=589.592$ nm, $\lambda_{D2}=588.995$ nm), which has passed through a vapor with density $\sim 10^{16}$ cm^{-3}, i.e., under conditions characteristic of many experiments devised to investigate resonant laser ionization (compare, for example, with Refs. 9 and 24), demonstrated, as expected, the presence of developed spatial and spectral instability in the laser wave. This instability is clearly evident in the experiments in the form of broadening of the laser emission spectrum, self-focusing and self-defocusing of the laser beam as a whole, splitting of the beam into separate filaments, generation of conical emission, generation of frequency-mixed scattering components, etc. As we have already noted, the characteristics of such instability are well known, and here we only dwell briefly on the results that have direct bearing on the formation of the resonant plasma.

Figure 3 presents profiles of the laser beam recorded at the sodium vapor cell output at different values of the detuning $\Delta\lambda$ ($\Delta\lambda=\lambda_{D2}-\lambda_L$, λ_L is the wavelength of the laser radiation), and Fig. 4 shows the dependence of the diameter d of a cross section of the laser beam at the half-height of the energy distribution relative to the beam diameter d_0 measured in the absence of vapor. It follows from these data that at relatively small detunings $\Delta\lambda>0$, the laser radiation undergoes strong large-scale self-focusing,[25] which leads to considerable contraction of the beam core. The usual conical emission (see, for example, Ref. 26) is generated in this case, but not at small of $\Delta\lambda$, since the cone angle becomes so great that the diameter of the scattered beam in the plane of the recording photodiode array significantly exceeds its aperture (1.3 cm). As the detuning increases, the contraction rapidly weakens, and the cone angle decreases. Being superimposed on the beam core, the conical emission broadens the measured beam profile significantly over a narrow range of detunings, imparting a characteristic two-humped shape to it (curve *3* in Fig. 3). At still larger detunings $\Delta\lambda>2-2.5$ nm the instability effects cease to influence the spatial form of the beam, and the profile measured under such conditions (Fig. 3, curve *2*) is essentially the same as the one recorded in the absence of vapor.

In the opposite case of long-wavelength detuning ($\Delta\lambda<0$) there are no data on the ratio d/d_0 in Fig. 4, since large-scale self-defocusing processes erode the beam profile so much that recording within the aperture of the photodiode array becomes possible only when $|\Delta\lambda| \geqslant 3-4$ nm. We note that the conical structure of the laser beam scattered in the Na vapor is clearcut even at such large detunings (Fig. 3, curve *5*).

The absorption of laser radiation is also important in the vicinity of resonance. These data are also presented in Fig. 4, where the dependence of the beam energy absorbed in the vapor E_{abs} normalized by the beam energy E_0 at the cell entrance aperture plotted. The asymmetry of the curve associated with the dependence of the spatial structure of the laser radiation on the detuning considered above is appreciable. Because of instability effects, the cross-sectional area S of the beam, and therefore the absorbed power (which is proportional to S when $\mathscr{I} \gg \mathscr{I}_s$), are significantly greater at negative values of $\Delta\lambda$ than at positive values, and consequently the drop in E_{abs} with increasing $|\Delta\lambda|$ is more gradual at $\Delta\lambda<0$. We note that at small positive values of $\Delta\lambda$ a considerable fraction of the beam energy shows up in conical emission, which undergoes considerable absorption. Therefore, despite the strong contraction of the beam core, the energy of the laser pulse absorbed in the vapor remains high near resonance: $0.3E_0$ to $0.8E_0$ for $0 \leqslant \Delta\lambda \leqslant 1$ nm.

These data clearly indicate that, near resonance, as discussed in the Introduction, strong nonuniformity of the beam and considerable absorption preclude a proper determination of the intensity of the laser radiation in the zone where the fluorescence of the resonant plasma is recorded. Moreover, of long-wavelength detuning range is totally inaccessible to a comparison of the results of experiments and theoretical simulation, because, as the measurements have shown, spatial nonuniformity of the radiation remains up to values of $\Delta\lambda$ at which there is essentially no significant ionization of the vapor. However, it follows from the plots in Fig. 4 that in the wavelength range $\lambda_L<\lambda_{D2}$ (i.e., positive $\Delta\lambda$), the influence of the vapor on the spatial profile of the laser beam is already insignificant at relatively small detunings $\Delta\lambda \geqslant 2-2.5$ nm (where the degree of ionization is still high; see below), and absorption reduces the radiation intensity only slightly. This permits the use of this range of detunings to compare experimental data with the results of theoretical analysis.

3.2. Measurement of electron density and electron temperature in a resonant laser plasma

The formation of a resonant laser plasma under the conditions of our experiment was visually observed in the form of distinct white glow in the channel of the laser beam when radiation with a wavelength close to the wavelengths of the $D1$ and $D2$ transitions and an intensity exceeding several MW/cm^2 acted on the sodium vapor. As in Ref. 9, the plasma fluorescence contained numerous spectral lines, the strongest of which belonged to the $3P-nD$ $(n=3-7)$ and $3P-nS$ $(n=4-7)$ transitions in a neutral sodium atom. There were no lines of sodium ions or argon atoms in the fluorescence spectrum. This is obviously attributable to their high excitation potentials.

To determine the electron density N_e we chose the spectral lines corresponding to the $3P-4D$ ($\lambda \simeq 568$ nm) transition, since, on the one hand, they had the highest intensity and were easily detected as the intensity of the laser radiation, the detuning of the laser frequency from the frequency of the $3S-3P$ transition, and the time delay τ_d between detection and the beginning of the laser pulse were varied over broad ranges. On the other hand, there are reliable theoretical data on the broadening parameters for the $3P-4D$ transition.[22,27] In addition, at electron densities between 10^{15} and 10^{16} cm^{-3}, the Stark widths of the spectral lines for this transition are fairly large (of the order of 0.03–0.3 nm) and, as can easily be shown, greatly exceed the broadening due to other factors (Doppler broadening, collisional broadening, etc.).

At the same time, the use of spectral lines corresponding to transitions to the sodium $3P$ levels for determining the plasma density has its own drawbacks: in particular, it enables one to make measurements only in the plasma afterglow in the absence of radiation. This is because the considerable optical thickness of the plasma channel introduces a significant error into the profiles of the detected spectral lines due to the high population of the resonant levels during a laser pulse. Measurements are possible only after the population of the $3P$ levels has dropped to an acceptable level as a result of spontaneous decay and deactivation by electrons.

Under the present conditions, the lack of capture was verified by the correspondence of the intensity ratio of the principal lines in the multiplet to the theoretical value determined from the sum rule: $J(3^2P_{1/2}-4^2D_{3/2})/J(3^2P_{3/2}-4^2D_{5/2})_{\text{theor}}=0.56$. As the measurements showed, the plasma becomes optically transparent when the time delay relative to the beginning of the laser pulse is of the order of 80–100 ns; therefore, all of our experiments were performed with $\tau_d \geqslant 100$ ns. We note that absorption by Na_2 molecules (with a crosssection $\gtrsim 10^{-17}$ cm^2 (Ref. 28)), which at $T \simeq 700$ K account for $\sim 3\%$ of all atoms,[28] is negligibe at these vapor densities (see also Ref. 9). A more detailed discussion of the applicability of Stark methods to the measurement of the density of a resonant laser plasma can be found in Refs. 9 and 24.

Figure 5 presents profiles of the $3P-4D$ multiplets recorded experimentally at various detunings. They display only the two strongest lines, which correspond to the $3^2P_{1/2}-4^2D_{3/2}$ and $3^2P_{3/2}-4^2D_{5/2}$ transitions, while the third component of the multiplet (which corresponds to the $3^2P_{3/2}-4^2D_{3/2}$ transition) does not show up at all in the spectrum, since its intensity is low, and the wavelength is essentially the same as that of the $3^2P_{3/2}-4^2D_{5/2}$ transition. Even under resonance conditions, the lines in the short-wavelength wing of the $3P-4D$ multiplet that correspond to dipole-forbidden $3P-4F$ transitions, which can be excited at high electron densities, do not show up at all in the spectrum.

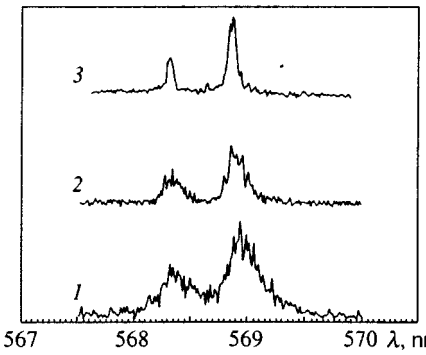

FIG. 5. Experimentally measured profiles of the lines corresponding to the $3P-4D$ transition; *1* — $\Delta\lambda=-3.6$ nm, *2* — $\Delta\lambda=2.1$ nm, *3* — $\Delta\lambda=0$ nm; $\mathscr{I}_L=66$ MW/cm^2, $\tau_d=130$ ns.

Note that the isolated-line approximation, in which the calculations of the Stark broadening parameters were performed in Refs. 22 and 27, is valid if the distance to the nearest perturbing level is much greater than the Stark width. In the sodium $4D$ levels ,this means that the influence of the $4F$ levels will be negligible, if $N_e \ll 6\times 10^{16}$ cm^{-3}. This is always the case under the present conditions.

Figure 5 clearly shows that the spectral lines are somewhat asymmetric and broadened on the red side. This means that, besides electronic broadening, ionic broadening affects their shape. Under these conditions, the line profile is not Lorentzian (collisional), and the electron density was, therefore, determined using the Holtsmark distribution, which describes the Stark profile of spectral lines with consideration of the influence of the ions in an ideal plasma. For this purpose, in the first step the Holtsmark distribution tabulated in Ref. 22 was approximated by a certain analytic function $F(\lambda,N_e)$ on the basis of the experimental data on the electron temperature (see below), and this function was then used to construct another function $\Phi(\lambda,N_e)$, which describes the superposition of the profiles of the two strongest lines of the multiplet with consideration of their relative intensities. Finally, the convolution of Φ and the instrumental function of the monochromator, which was already used to approximate the experimentally measured spectrum of the multiplet by the least-squares method (see Fig. 6), was calculated. As a result, this procedure made it possible to obtain the value of N_e with which the theoretical curve most closely describes the experimental data.

The results of the measurements of the dependence of the electron density on several parameters of the experiment are shown in Figs. 7–9. It is clear from the preliminary data that the dependence of N_e on the detuning $\Delta\lambda$ has a resonance character (Fig. 7) and that the maximum of the elec-

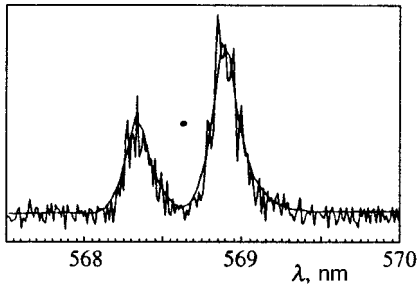

FIG. 6. Experimentally measured profile of the $3P-4D$ multiplet for $\Delta\lambda=2.1$ nm and its approximation; $\mathscr{I}_L=66$ MW/cm^2.

tron density ($N_e\simeq 8\times 10^{15}$ cm^{-3}) is achieved when the laser radiation is tuned to the frequencies of the $D1$ and $D2$ lines. One characteristic feature of this dependence is the fact that N_e decreases considerably more slowly at short-wavelength detunings than at long-wavelength detunings. This is because the laser beam is broadened to a considerably greater degree at negative detunings than when $\Delta\lambda>0$. The intensity of the radiation at a given absolute value of the detuning is significantly smaller for $\Delta\lambda<0$ than for $\Delta\lambda>0$. This is responsible for the faster drop in N_e at negative detunings.

Figure 8 presents the dependence of the electron density on the intensity of the laser beam for $\Delta\lambda=2.1$ nm, i.e., for the detuning at which, as follows from the foregoing, a more or less proper determination of the power density of the laser radiation is possible. Plots of the time dependence of N_e for two values of $\Delta\lambda$ are shown in Fig. 9. Hence it is clear, in particular, that at both detunings the plasma density decreases approximately by a factor of 2 after a time $\tau_{1/2}=100$ ns; at least-squares fit to the experimental data yields $N_e(t)\propto t^{-1}$.

The electron temperature T_e was determined in the experiments from the relative intensity of the set of spectral lines belonging to the $3P-nD$ and $3P-nS$ series of transitions. The applicability of this method is based on the fact that after completion of a laser pulse, the dense resonant plasma quickly reaches a state of local thermodynamic equilibrium.[4,29] As we know, for the existence of local thermodynamic equilibrium, the plasma density must satisfy the condition[23]

$$N_e[\mathrm{cm}^{-3}]>3.3\times 10^{13}E_{mn}^4[\mathrm{eV}]T_e^{-1/2}[\mathrm{eV}],$$

where E_{mn} is the highest energy of the transitions considered. Under the present conditions ($E_{mn}<2.5$ eV, $T_e\sim 0.5$ eV), the critical electron density is of the order of 2×10^{15} cm^{-3}, which is several times lower than the experimentally measured values, and therefore the inequality just presented is satisfied with a sufficient margin.

The data from a series of measurements of the intensities of lines are presented in Fig. 10. Here the energies of the upper levels of the corresponding transitions are plotted along the horizontal axis, and the values plotted along the vertical axis are of the logarithm of the reduced intensity J' of the emission in the line:

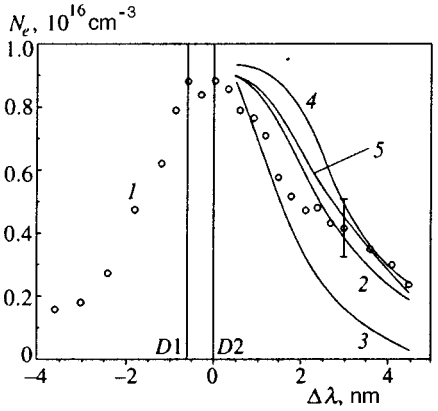

FIG. 7. Dependence of the electron density in the resonant laser plasma on the detuning $\Delta\lambda$: *1* — experiment; *2*, *3*, *5* — calculations based on the mechanism proposed in Refs. 5 and 6 without consideration of the influence of stimulated inverse bremsstrahlung on the heating of the electrons; *4* — calculation with consideration of that effect. Curve *5* corresponds to a calculation with the inclusion of laser-induced associative ionization. $\mathscr{I}_L=66$ MW/cm^2 (*1*, 2, *4*, *5*) and 33 MW/cm^2 (*3*), $\tau_d=130$ ns. The vertical lines mark the positions of the sodium $D1$ and $D2$ lines.

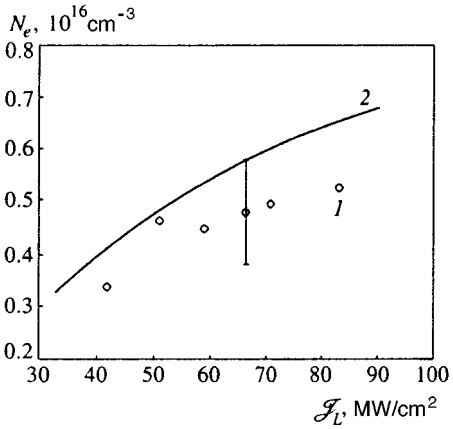

FIG. 8. Dependence of the electron density in the resonant laser plasma on the intensity of the laser radiation: *1* — experiment, *2* — calculation; $\Delta\lambda=2.1$ nm, $\tau_d=130$ ns.

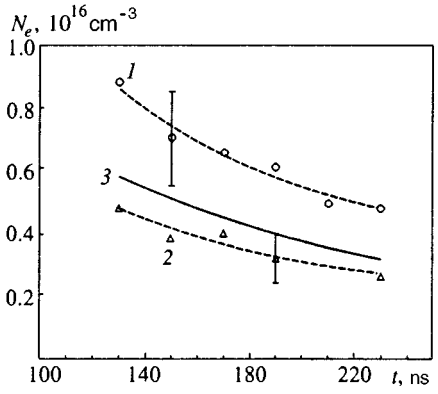

FIG. 9. Dependence of the electron density in the resonant laser plasma on time (*1*, *2* — experiment, *3* — calculation): *1*, *3* — $\Delta\lambda=2.1$ nm, *2* — $\Delta\lambda=0$, $\mathscr{I}_L=66$ MW/cm^2.

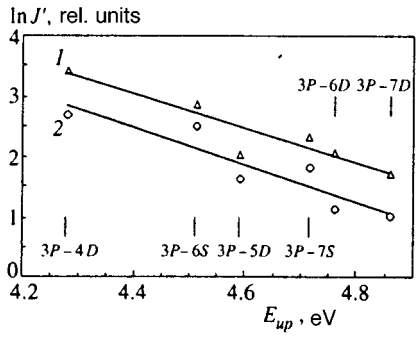

FIG. 10. Dependence of the reduced fluorescence intensity of the spectral lines on the energy of the upper level of the corresponding transition: *1* — $\Delta\lambda=0$, *2* — $\Delta\lambda=2.1$ nm; $\mathscr{I}_L=66$ MW/cm^2, $\tau_d=130$ ns.

$$\ln J' = \ln\left(\lambda^3 \sum_i J_i \Big/ \sum_i f_i g_i\right),$$

where λ is the wavelength of the transition, the indices $i=1,2$ correspond to the two components of the doublet, the J_i are the intensities of the doublet lines with a correction for the spectral sensitivity of the photomultiplier, and f_i and g_i are the oscillator strengths of the corresponding transitions in the doublet and the degrees of degeneracy of their lower level; values can be found in Ref. 30. The electron temperature is given in such a plot by the slope ($\tan\varphi=-1/T_e$) of the linear least-squares fit to the experimental data. In the present case $T_e\simeq 0.38$ eV ($\Delta\lambda=0$) and $T_e\simeq 0.34$ eV ($\Delta\lambda=2.1$ nm). Figure 11 presents the dependence of the electron temperature on the detuning. As follows from this figure, T_e decreases only slightly as $\Delta\lambda$ decreases, and this decrease lies within the experimental error. The electron temperature is also essentially constant varies as the delay τ_d varies over the range 100–250 ns.

4. THEORETICAL MODEL OF RESONANT LASER BREAKDOWN

To calculate the characteristics of the plasma formed as a result of resonant laser breakdown, we constructed a kinetic model that describes the principal processes involved in the production of the charged particles in a mixture of sodium vapor and a buffer gas (argon), allowing for the actual temporal shape of the laser pulse. It was assumed here that both the medium and the laser radiation are spatially homogeneous, and, thus, the development of breakdown with time can be described by a system of ordinary differential equations. This system included equations for the populations of a series of levels of the Na atom, equations for the number densities of the electrons, monatomic Na^+ ions, Na_2^+ and $NaAr^+$ molecular ions, and Na_2 sodium molecules, an equation for the electron temperature, and an equation for the laser pulse energy $E_{\rm dis}$ dissipated per unit volume of the resonant medium. We note that the excitation and ionization processes of Ar atoms and the excitation of Na^+ ions, which, as estimates show, are insignificant at the small values of T_e characteristic of a resonant plasma, were neglected in the model. This is also confirmed by the absence of any lines for argon and the Na^+ ion in the fluorescence spectrum of the plasma column.

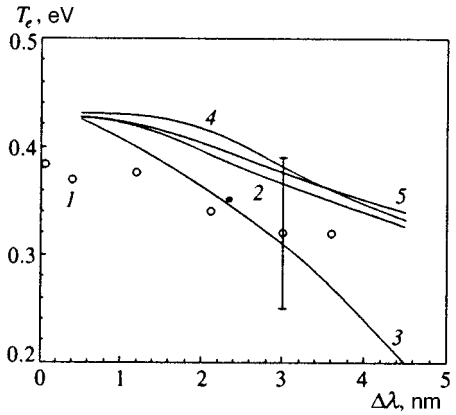

FIG. 11. Dependence of the electron temperature on the detuning $\Delta\lambda$. The notation and the parameters of the experiment correspond to Fig. 7.

In the calculations we did not assume complete saturation of the resonant transition, and the populations of the $3S$ and $3P$ levels were calculated independently, with consideration of the detuning of the frequency of the laser radiation from resonance. Because of the large value of the fine splitting $\delta\lambda\simeq 0.6$ nm (which is comparable to $\Delta\lambda$), we treated the $3P$ levels as two separate levels ($3^2P_{1/2}$ and $3^2P_{3/2}$) that interact independently with the laser field. In addition, it was necessary to take into account in the model energy exchange between the $3^2P_{1/2}$ and $3^2P_{3/2}$ states, mediated by collisions with argon atoms; the cross sections for the latter is fairly large ($\sigma(P_{1/2}\rightarrow P_{3/2})=1.1\times 10^{-14}$ cm^2, $\sigma(P_{3/2}\rightarrow P_{1/2})=0.56\times 10^{-14}$ cm^2 (Ref. 31)). We neglected energy exchange in collisions with sodium atoms, which have a similar cross section of 1.6×10^{-14} cm^2 (Ref. 32), since the density of the buffer gas is more than two orders of magnitude greater than the density of the vapor.

In calculating the cross sections of stimulated transitions, the profile of each of the two resonant absorption lines was assumed to be Lorentzian and to have a full width $\Delta\nu_{1/2}$ at half maximum equal to the sum of the widths corresponding to atomic collisional broadening and the decay of the resonant levels. Among the former we took into account only resonant collisions involving the Vlasov–Fursov mechanism, for which the collisional broadening parameters were taken from Ref. 33 ($k_{\rm res}\simeq 2.4\times 10^{-7}$ cm^3/s for the $3^2P_{3/2}$ level). However, the Weisskopf frequency, which determines the limits of applicability of the collisional approximation and can be estimated from that constant, is just only 3×10^{10} Hz (which corresponds to a detuning of 0.04 nm). Since we are only interested in fairly large $\Delta\lambda$, the contribution of the resonant collisions to the absorption cross section was determined on the basis the static wing, which is known[34] to have a symmetric Lorentzian line profile with a width $\Delta\nu_{\rm res}$ somewhat exceeding the collisional linewidth (by 16% for the $3^2P_{3/2}$ level) in the process under consideration. For a vapor density $N_0=1.8\times 10^{16}$ cm^{-3}, $\Delta\nu_{\rm res}=\pi^{-1}k_{\rm res}N_0$ is 1.4×10^9 Hz. When the plasma parameters are calculated in the wavelength range of the laser radiation $\lambda_L<\lambda_{D2}$, broadening due to collisions of excited Na ($3P$) atoms with argon atoms can be neglected, since the static wing is significant in

this case only at low-frequency detunings,[35] and the limit of the collisional approximation based on the collisional broadening parameters given in Ref. 36 also corresponds to small $\Delta\lambda \sim 0.2$ nm. Inhomogeneous Doppler broadening ($\Delta\nu_D \simeq 2\times 10^9$ Hz) likewise does not play a role in the wings of the absorption lines.

In calculating the cross sections we also took into account all the decay processes of the resonant levels included in the model under consideration. Among them, the main contributions are made by the aforementioned mixing of fine-structure states in collisions with argon atoms (which corresponds to a uniform width $\Delta\nu_{\mathrm{Ar}} = 0.46\times 10^9$ Hz for the $3^2P_{3/2}$ level), and by the electronic deexcitation of resonant levels with a transition to the $3S$ state, with a rate constant $k(3P\rightarrow 3S) \simeq 1.6\times 10^{-7}$ cm^3/s at $T_e \simeq 0.3$–0.4 eV (Ref. 37), which corresponds to a similar width $\Delta\nu_e \simeq 0.5\times 10^9$ Hz when $N_e = 10^{16}$ cm^{-3}.

Other inelastic collisional processes have a significantly smaller influence on the total broadening of the absorption lines. The contribution of spontaneous emission is small: $\Delta\nu_{\mathrm{sp}} = 10^7$ Hz.[30] We note that the short phase relaxation time of the resonant system $T_2 \ll \tau_L$ ($T_2 = (\pi\Delta\nu_{1/2})^{-1} \simeq \pi^{-1}(\Delta\nu_{\mathrm{res}} + \Delta\nu_{\mathrm{Ar}} + \Delta\nu_e)^{-1} \sim 10^{-10}$ s) and the large value of the Rabi frequency $\Omega_R\tau_L \gg 1$ ($\Omega_R = \mu E/\hbar$, where μ is the dipole moment matrix element and E is the field strength of the laser wave; when $\mathscr{I} = 70$ MW/cm^2, $\Omega_R \simeq 8\times 10^{12}$ s^{-1}) enables us to ignore coherent and time-dependent effects in the kinetic description of the interaction of radiation with the resonant transition.

Besides the resonant levels, only the two lowest-lying levels, $3D$ and $4S$, whose ionization by electron impact has been calculated in the diffusion approximation,[38] were taken into account in the kinetics of the excited states in the model under consideration. This was based on the fact that because of the low electron temperature under our conditions (0.3–0.4 eV), inelastic collisions with electrons are important to the population kinetics of the Na levels only for transitions between neighboring terms. Then the ''bottleneck'' in the ionization of laser-populated resonant states is the excitation of the $3D$ and $4S$ levels nearest to them, since the energies of the $3P-3D$, $4S$ transitions are significantly higher than the energies of the transitions between the higher-lying levels. We calculated the constants of the processes of excitation and deexcitation by electrons in the dependence on T_e in the $3S-3P$ and $3P-3D$, $4S$ transitions in the Van Regemorter approximation,[34] and the calculated deexcitation constant for the resonant transition, whose value has the strongest influence on the final degree of ionization of the vapor, coincides with the value measured in Ref. 37 to high accuracy ($\sim 5\%$).

The electronic excitation constants calculated in the same approximation for the allowed $3S-nP$ ($n \geqslant 4$), $3P-nD$ ($n \geqslant 4$), and $3P-nS$ ($n \geqslant 5$) transitions were several orders of magnitude smaller than the stepwise excitation constants, confirming what was stated above. The probability of the excitation of forbidden transitions by electron collisions is also negligible at low temperature.

The direct ionization of all the low-lying levels calculated in accordance with the recommendations in Ref. 39 also turned out to be insignificant. We note that the possibility of one-photon ionization of the $4D$ and $4S$ levels by laser radiation with a cross section $\sim 6\times 10^{-18}$ cm^2 (Ref. 40) was taken into account in their kinetics. Photoionization was neglected for the higher-lying levels, since the cross section of this process decreases rapidly as the excitation energy E_i increases ($\propto [(I-E_i)/h\nu_L]^3$, where $I = 5.14$ eV is the ionization potential of the sodium atom[30]).

The applicability of the diffusion approximation is based on the fact that the influence of radiative processes on the upper levels can be neglected at the large values of N_e typical of the present experiments ($N_e \sim 10^{16}$ cm^{-3}).[38] Radiative decay of the lower levels ($3P$, $3D$, and $4S$) was taken into account in the model, although as the calculations show, its role is negligible, and therefore the possible capture of radiation in the $3P-3S$, $3D-3P$, and $4S-3P$ transitions is insignificant. Population of the upper levels in $2\mathrm{Na}(3P)\rightarrow \mathrm{Na}(3S)+\mathrm{Na}(nL)$ reactions, whose constants are small ($10^{-11}-10^{-13}$ cm^3/s (Ref. 28)), likewise does not play a role. The constant for excitation transfer from resonant levels to a Na_2 molecule is considerably larger ($\sim 10^{-9}$ cm^3/s (Ref. 28)). However, owing to their low density (6×10^{14} cm^{-3}), this process likewise does not affect ionization kinetics in times $\sim 10^{-7}$ s. As estimates show, in this situation stepwise ionization is relatively insensitive to the details of the level-by-level kinetics in highly excited states of the sodium atom, and the incorporation of equations for two series of levels (as was done, for example, in Refs. 5, 11, 12, and 13) makes no particular sense and, in addition, is unreliable, since the constants of the pertinent elementary processes are poorly known (see also Ref. 14).

Apart from stepwise processes, processes leading to ''seed'' ionization[2,6] play an important role in the dynamical development of resonant breakdown. The main such process is associative ionization during a collision between two resonantly excited sodium atoms:

$$2\mathrm{Na}(3P)\rightarrow \mathrm{Na}_2^+ + e, \quad (k_{\mathrm{AI}} = 3.8\times 10^{-11}\ \mathrm{cm}^3/\mathrm{s}\ \ (\mathrm{Ref.\ 28})). \tag{1}$$

Although the value of this constant, which was accurately measured in special experiments at low intensity of the exciting laser field and agrees well with other data (see the discussion of this point in Ref. 28) is quite high (a constant two orders of magnitude smaller was used in the analytic model in Ref. 6), the actual contribution of associative ionization is small. When the density of the resonantly excited atoms is 10^{16} cm^{-3}, associative ionization occurring during a single laser pulse can provide an electron number density only at the level of 10^{14} cm^{-3}. In addition, this process has a significant influence on shortening the time for the development of electron avalanche by providing a considerable number of initial electrons.

We note that associative ionization involving atoms in more highly excited states has been neglected because of their relatively scarcity and the significantly smaller constant of the process.[28] Apart from associative ionization, the following less important processes were also taken into account in the model as sources of seed ionization: two-photon ion-

ization of the $3P$ levels by laser radiation with a cross section $\sigma_{2ph} \simeq 7.8 \times 10^{-49}$ cm^4/s (Ref. 41) and laser-induced Penning ionization:

$$2\mathrm{Na}(3P) + h\nu_L \rightarrow \mathrm{Na}^+ + \mathrm{Na}(3S) + e,$$

$$(\sigma_{\mathrm{PI}} = 2 \times 10^{-44} \text{ cm}^4/\text{s (Ref. 42)}). \quad (2)$$

As noted in Sec. 3, measurements of the electron density could only be performed with a considerable delay relative to the end of the laser pulse. For this reason, to compare the experimental data with the results of calculations, it is very important to have a proper description of the decay processes of the resonant plasma. One of the main such processes is three-body recombination,

$$\mathrm{Na}^+ + 2e \rightarrow \mathrm{Na} + e, \quad (3)$$

whose constant was also calculated in the diffusion approximation.[38] However, three-body recombination alone is not capable of accounting for the experimentally observed drop in electron density with time (Fig. 9). Thus, a typical relaxation time τ_r of the density of a resonant plasma as a result of this process, allowing for the dominant cooling of electrons in elastic collisions with ions, according to Ref. 29, for $T_e \simeq 0.3-0.4$ eV and $N_i \simeq 10^{16}$ cm^{-3}, is of the order of 5×10^{-7} s, which is several times the empirical value.

We note that the characteristic relaxation time of the electron density as a result of elastic collisions of electrons with argon and sodium atoms, calculated from the data on the corresponding transport cross sections presented in Refs. 23 and 43, is an order of magnitude greater than τ_r. The recorded drop in N_e cannot be attributed to diffusional spreading of the plasma column, since the rate of ambipolar diffusion at the pressures used in the experiment is low due to the low mobility μ_+ of sodium ions in argon. The mobility can be evaluated from the formula[23] $\mu_+ \simeq 97 \times 10^{19}/(N_{\mathrm{Ar}}\sqrt{\beta M}) \simeq 30$ cm^2/V$\cdot$s, where $\beta = 11.08$ amu is the polarizability of the Ar atom[30] and $M \simeq 14.6$ amu is the reduced mass. Hence the typical time of ambipolar diffusion from a cylindrical column with a radius equal to the radius of the laser beam (0.3 mm) is

$$\tau_{\mathrm{dif}} \simeq \left(\frac{r}{2.4}\right)^2 \frac{1}{\mu_+ T_e} \simeq 5 \times 10^{-6}\mathrm{s} \gg \tau_{1/2}.$$

For these reasons, to describe the decay of the plasma, the model had to take into account the possible processes leading to the formation and dissociative recombination of molecular ions, which can only be Na_2^+ and $NaAr^+$ ions under the conditions of the present experiment. The main role here is clearly played by the $NaAr^+$ ion (with a binding energy of 0.16 eV (Ref. 30)), that forms in the conversion reaction of the monatomic Na^+ ion

$$\mathrm{Na}^+ + 2\mathrm{Ar} \rightarrow \mathrm{NaAr}^+ + \mathrm{Ar}. \quad (4)$$

We were unable to find data on the constant of this reaction in the literature; therefore, its value was determined by fitting the calculated $N_e(t)$ to the experimental data (with allowance for all other plasma decay mechanisms), which yielded the value 7×10^{-31} cm^6/s (we note that this value is three times greater than the estimate obtained from the formula given in Ref. 31). In addition, Na_2^+ ions form in similar conversion reactions:

$$\mathrm{Na}^+ + 2\mathrm{Na} \rightarrow \mathrm{Na}_2^+ + \mathrm{Na}, \quad (5)$$

$$\mathrm{Na}^+ + \mathrm{Na} + \mathrm{Ar} \rightarrow \mathrm{Na}_2^+ + \mathrm{Ar}. \quad (6)$$

However, when $N_0 \sim 10^{16}$ cm^{-3}, reaction (5) is completely negligible, while (6) makes a small contribution to the total production rate of molecular ions, even despite the significant value of the conversion constant (an estimate using the formula from Ref. 31 yields a value of 6×10^{-30} cm^6/s).

A significantly greater role is played by the production of the Na_2^+ ion in the associative ionization reaction considered above, as well as in the following reaction involving charge transfer from an $NaAr^+$ ion:

$$\mathrm{NaAr}^+ + \mathrm{Na} \rightarrow \mathrm{Na}_2^+ + \mathrm{Ar}. \quad (7)$$

The value of the constant of this process, which can be determined from the polarization capture cross section (see Ref. 44) is 2×10^{-9} cm^3/s.

The dissociative recombination constant of the $NaAr^+$ ion in the model was set equal to that for the Na_2^+ ion, for which the value $3 \times 10^{-7}(0.026/T_e[\mathrm{eV}])^{1/2}$ was given in Ref. 45. The results of the calculations, however, are insensitive to its exact value, since even if the constant is diminished by an order of magnitude, dissociative recombination of the molecular ions is still restricted by their production rate. Besides dissociative recombination, for the Na_2^+ ion we also took into account another decay process, namely photodissociation under the action of laser radiation with a cross section of 1.7×10^{-17} cm^2.[46] Photodissociation by visible light is impossible for the $NaAr^+$ ion, since the ground state in the Na^+–Ar system has a single term, and transitions to the excited states lie in the ultraviolet region.

Besides the equations for the populations of the levels of atomic sodium and the densities of the monatomic and molecular ions, an equation for the density of the Na_2 molecules was also introduced into the kinetic model of resonant breakdown. The only process that takes place with the participation of these molecules and had to be taken into account in the calculations is their dissociation under electron impact, whose constant was evaluated in Ref. 44: $k \simeq 8 \times 10^{-8}$ cm^3/s. Photodissociation of Na_2 by radiation at the laser wavelength ($\lambda \sim 590$ nm) is insignificant.[28]

Note that the excitation of vibrational and rotational degrees of freedom of molecular sodium (as well as, incidentally, the excitation of the molecular ions) was neglected in the calculations. The constants of these processes are unknown, but the estimates of these constants for similar plasma parameters in Ref. 44 show that they play a minor role. The formation of the excimer NaAr* molecule, whose dissociation energy (0.07 eV (Ref. 30)) is of the order of the temperature of the heavy particles, was also neglected in the calculations.

It was assumed in the calculations that the electron energy distribution function is Maxwellian. Under our conditions this is fully justified, since at the high values $N_e > 10^{15}$ cm^{-3} characteristic of the resonant plasma, the fre-

quencies of the electron–electron collisions, which establish the Maxwellian energy distribution function, are sufficiently high. In addition, numerical simulations of the electron energy distribution function in Ref. 12 show that a deviation from the Maxwellian distribution is observed only in the earliest stage of resonant laser breakdown, in which the extent of ionization is less than 10^{-5}. All this makes it possible for us not to calculate the distribution function, but to simply use the equation for T_e, which takes into account all processes leading to energy gain or loss by electrons in the present kinetic model. In this case the contribution of inverse bremsstrahlung evaluated from the classical formulas[23] is small and does not influence the result (see also Refs. 11 and 12).

However, as shown in Refs. 2, 7, and 47, the mechanism by which electrons are heated during resonant ionization of a medium cannot be reduced to collisions of the second kind, and it must be regarded as a stimulated inverse bremsstrahlung process accompanying the inelastic scattering of electrons by an atom in the field of the resonant laser wave. We note that this phenomenon can also be interpreted in terms of inelastic collisions of electrons with a resonant atomic system split as a consequence of the dynamic Stark effect. This enables us, in particular, to take resonant stimulated inverse bremsstrahlung into account in the equation for the temperature without solving the quantum kinetic equation by modeling the terms in it that correspond to excitation and deexcitation of the resonant transition by electron impact while allowing for the influence of the laser field.

The following expression for the energy gain loss in a resonant transition as a result of resonant stimulated inverse bremsstrahlung can be obtained from the probabilities corresponding to the transitions for a two-level system presented in Ref. 47 in the approximation $\nu_L \gg \nu_R = \Omega_R/2\pi$, $\nu_L \gg \Delta\nu$, ν_R, $\Delta\nu \gg (\pi T_2)^{-1}$:

$$Q_{12} = \left[\frac{d}{dt}\left(\frac{3}{2}T_e\right)\right]_{12} = \frac{h\nu_0}{2}\left\{k_{21}\left[N_2\left(1+\frac{\Delta\nu}{\nu_R'}\right)+N_1\frac{g_2}{g_1}\left(1-\frac{\Delta\nu}{\nu_R'}\right)\right] - k_{12}\left[N_2\frac{g_1}{g_2}\left(1-\frac{\Delta\nu}{\nu_R'}\right)+N_1\left(1+\frac{\Delta\nu}{\nu_R'}\right)\right]\right\}, \quad (8)$$

where $N_{1,2}$ and $g_{1,2}$ are the populations and degrees of degeneracy of the lower and upper resonant levels, k_{12} and k_{21} are the excitation and deexcitation constants in the absence of a laser field, and $\nu_R' = (\nu_R^2 + \Delta\nu^2)^{1/2}$ is the generalized Rabi frequency. As follows from (8), when $\nu_R \gg \Delta\nu$ and $N_2 \simeq N_1$, the expression for Q_{12} transforms at saturation into the usual expression for the energy gain in collisions of the second kind, and the energy loss as a result of excitation:

$$Q_{12}' \simeq h\nu_0(k_{21}N_2 - k_{12}N_1). \quad (9)$$

The same expression is also obtained in the other limiting case of a weak field and large detunings $\Delta\nu \gg \nu_R$, but when $\Delta\nu \simeq \nu_R$, Q_{12} differs significantly from (9) and exceeds it, the difference for given values of N_1 and N_2 being greatest when $\Delta\nu = \nu_R$.

When stimulated inverse bremsstrahlung is taken into account, the expression for the energy gain in transitions from the upper resonant level to other levels m also changes:

$$Q_{2m} \equiv \left[\frac{d}{dt}\left(\frac{3}{2}T_e\right)\right]_{2m} = \frac{h\nu_m}{2}\left\{k_{2m}\left[N_2\left(1+\frac{\Delta\nu}{\nu_R'}\right) + N_1\left(1-\frac{\Delta\nu}{\nu_R'}\right)\right] - 2k_{m2}N_m\right\}, \quad (10)$$

where $h\nu_m$ is the energy of the $2 \rightarrow m$ transition, N_m is the population of level m, and k_{2m} and k_{m2} are the constants of the inelastic transitions in the absence of a laser field. As can easily be seen, Eq. (10) also transforms into the standard expression in the same two limiting cases.

It should be noted that, generally speaking, under the conditions of our experiment the three-level nature of the resonant system must also be taken into account in calculating the energy gain by electrons, since the Rabi frequency and the detuning are of the same order of magnitude as the fine splitting between the $3P$ levels. However, since we know of no generalization of the transition probabilities for stimulated inverse bremsstrahlung to the three-level case, we used the two-level approximation consisting of (8) and (10) in the equation for T_e. In this case N_2 and g_2 in (8) and (10) were construed as the total population and degree of degeneracy of the $3^2P_{1/2}$ and $3^2P_{3/2}$ levels, and the total value of the matrix element for the two transitions was used in the calculation of the Rabi frequency (see Ref. 30). We note that a detuning of 1.6 nm corresponds to the Rabi frequency for $\mathcal{I}_L = 66$ MW/cm^2. The two-level approximation is also justified by the fact that, as follows from the calculations, the deviations of the ratios between the populations of the $3^3P_{1/2}$ and $3^2P_{3/2}$ levels from those determined by the statistical weights of those levels do not exceed 10%.

5. NUMERICAL MODELING AND COMPARISON WITH EXPERIMENTAL DATA

In the first stage of numerical modeling, we calculated the parameters of the resonant laser plasma for the mechanism of heating electrons proposed in Refs. 5 and 6, which corresponds to the approximation $\nu_R = 0$ in Eqs. (8) and (10). Figure 12 presents the time dependence of the densities of the individual components of the laser plasma, as well as the electron temperature, for a detuning $\Delta\lambda = 2.1$ nm and a laser pulse intensity $\mathcal{I}_L = 66$ MW/cm^2 calculated for that case. The figure also presents the empirical temporal shape of the laser pulse, which was taken into account in the calculations. As is clear from this figure, the development of resonant laser breakdown is highly time-dependent. The electron temperature varies markedly during the laser pulse, rising from its initial value, which was assumed to equal the temperature of the cell ($\sim$0.06 eV), to 0.75 eV at the maximum of the lasing pulse. The high value of the temperature is associated with the still relatively low degree of ionization ($\sim$0.01) at that time. The rise of the electron density, and thus the large losses involved in the large ionizing the Na atoms lead to a subsequent drop in temperature to 0.45 eV by the end of the laser pulse (which is close to the analytically calculated

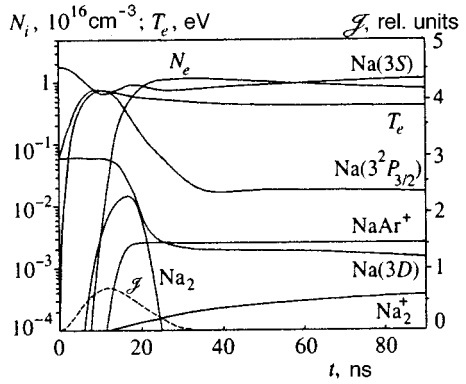

FIG. 12. Calculated plots of the time dependence of the concentrations of individual components of the resonant laser plasma and the electron temperature. The dashed line shows the experimentally measured temporal shape of the laser pulse. $\mathscr{J}_L = 66$ MW/cm^2, $\Delta\lambda = 2.1$ nm.

value $T_e \simeq 0.35$–0.4 eV in Ref. 4), the electron density reaching a maximum at the same time for the detuning $\Delta\lambda = 2.1$ nm. In the absence of radiation, the temperature continues to slowly decrease because of the cooling of the electrons in elastic collisions and dissociative recombination processes.

Note that in the analytic model in Ref. 6 and the numerical calculations in Ref. 11, the temperature was found to be independent of the time and significantly higher: 0.8 eV. The time-dependent character of breakdown typically also shows up in nonmonotonic behavior of the density of the ground-state sodium atoms. It decreases rapidly at first as a result of saturation of the resonant transition, then increases somewhat as N_e increases due to the rapid deexcitation of the resonant levels by electrons, and drops again toward the end of the laser pulse, reflecting the intense ionization of the medium. The high electron density causes Na_2 molecules to essentially vanish at the end of the laser pulse as a result of rapid dissociation by electron impact. The process leading to their formation is very slow[28,44] and was not taken into account in the present model. The main components of the resonant plasma in the afterglow are (see Fig. 12) electrons, Na^+ ions ($N(Na^+) \approx N_e$), and ground-state sodium atoms, the character of the drop in the theoretical $N_e(t)$ curve being consistent with experiment at large times equal to 100–250 ns (see Fig. 9).

Note that the line profiles that we measure and use to calculate N_e and T_e correspond to integration of the plasma fluorescence across the diameter of the laser beam. The associated error, however, is small, since the dominant contribution to the plasma emission comes from the narrow central zone of the plasma channel. An analysis of the experimental data obtained with the spatial resolution in Ref. 24, as well as a calculation of fluorescence in the spectral lines of the $3D-3P$ transition integrated over the diameter, utilizing the model described above, confirm this fact and show that the error associated with averaging is in the range 10–15%, which is within the overall error of the measurements.

Figure 7 presents the dependence of N_e on $\Delta\lambda$ calculated for the short-wavelength detunings $\Delta\lambda \geqslant 0.5$ nm. No calculations were performed for the near-resonance region $|\Delta\lambda| < 0.5$ nm because of the use of the static broadening mechanism in the model mentioned above and the neglect of the Doppler broadening in the calculation of the cross section of the resonant transition. No simulation was performed at long-wavelength detunings, since a correct comparison with experiment is impossible for them (see Sec. 3.1). As follows from Fig. 7, the calculations for $\Delta\lambda \geqslant 2$–$2.5$ nm agree well with the experimental data to within the measurement error. In the intermediate range of detunings equal to 1–2 nm the calculated values of N_e are significantly higher than the measured values. This also reflects the decrease in the effective intensity of the laser pulse in the region where the plasma parameters are recorded due to the influence of the instability and the absorption of the laser radiation considered in Sec. 3.1. As the calculations showed, at the small detunings $\Delta\lambda = 0.5$–1 nm practically complete ionization of the medium is achieved by the end of the laser pulse not only for the laser radiation intensities realized in the experiment, but also for significantly smaller intensities. The resonant plasma decays in the same manner; therefore, as is shown in Fig. 7, the calculated curve for $\mathscr{J} = 66$ MW/cm^2 practically coincides with the curve constructed for an intensity two times smaller at $\Delta\lambda = 0.5$ nm, deviating strongly from it at larger values of $\Delta\lambda$. This is reflected in the fact that even at the fairly large detuning $\Delta\lambda = 2.1$ nm both the experimental and calculated dependences of N_e on the intensity of the laser radiation have a tendency to achieve saturation (See Fig. 8). The relative agreement between the experimental data and the simulation results at small detunings is achieved, because at values of $\Delta\lambda$ close to resonance complete ionization of the sodium vapor is still achieved during the pulse even when the intensity of the laser radiation is reduced due to absorption. Figure 4 presents the theoretical dependence of the energy of the laser pulse absorbed in the sodium vapor on the detuning, which was calculated by integrating E_{dis} over the diameter of the cell with consideration of the measured $N(R)$ curve (see Fig. 2). A comparison of the experimental data with the theoretical curve in Fig. 4 reveals that they are in good agreement at large values of $\Delta\lambda$, but diverge significantly as resonance is approached. This is attributed to the appearance of intense absorption due to the development of instability under real experimental conditions. We note that, unlike N_e, the calculated values of the electron temperature agree with the experimental data to within the measurement error over the entire range of detunings (see Fig. 11). This is because T_e varies relatively slowly as the radiation intensity is varied at $0 \leqslant \Delta\lambda \leqslant 2$ nm (compare curves *2* and *3* in Fig. 11).

The decisive role of electron deexcitation of the resonant levels in the energy gain by electrons in the resonant plasma should cause the dependence of N_e on $\Delta\lambda$ to be determined mainly by the dependence of the population of the $3P$ levels $N(3P)$ on the detuning and, accordingly, by the degree of saturation of the resonant system by the laser radiation. The full width at half-height of the distribution of $N(3P)$ in the stationary case is specified by the expression

$$\Delta\lambda_{N(3P)} \simeq \frac{\lambda_0^2}{c}\Delta\nu_{1/2}\sqrt{\frac{G\mathscr{J}}{\mathscr{J}_s}}, \tag{11}$$

where $G = (1/2)(g_{3S}/g_{3P} + 1) = 2/3$ and c is the speed of

light. Setting $\Delta\nu_{1/2}\simeq 2.4\times 10^9$ Hz for $N_e\sim 10^{16}$ cm^{-3} and taking the relaxation time of the energy of the resonant system $T_1\simeq(\pi\Delta\nu_e)^{-1}\simeq 6\times 10^{-10}$ s, we can easily show that $\Delta\lambda_{N(3P)}\simeq 6$ nm. With consideration of the roughness of its evaluation, this value agrees with the experimentally measured value of 4 nm for the width of the $N_e(\Delta\lambda)$ distribution (see Fig. 7). We note that if the spontaneous decay time (16 ns) is used for T_1, $\Delta\lambda_{N(3P)}$ turns out to be five times greater. Thus, the quenching of the resonant levels by electrons strongly diminishes the saturation, as was noted back in Ref. 48. Nevertheless, since the saturation is still great near resonance, it can be assumed that calculations based on this model will also be valid in the range of detunings $\Delta\lambda<0.5$ nm regardless of the broadening mechanism, as is confirmed by the flattening of both the experimental and theoretical $N_e(\Delta\lambda)$ and $T_e(\Delta\lambda)$ curves as resonance is approached.

As was noted in the Introduction, a mechanism of resonant breakdown based on the hypothetical laser-induced associative ionization process $2\mathrm{Na}(3P)+h\nu_L\rightarrow \mathrm{Na}_2^+ + e$ [$k_{L,\mathrm{Al}}\simeq 5\times 10^{-10}$ cm^3/s for $\mathscr{I}_L>0.1$ MW/cm^2 (Ref. 8), which is an order of magnitude greater than the constant for ordinary associative ionization] was proposed in Ref. 8. In Ref. 9 experimentally obtained (under the conditions of exact resonance $\Delta\lambda=0$) laws for resonant breakdown were interpreted on the basis of this process, and the low electron temperature in the plasma afterglow $T_e\simeq 0.3$ eV, which contradicts the calculations in Refs. 6 and 11, was explained for the most part. However, as the data presented above show, even in the near-resonance region the calculated temperature can be fairly low. Taking into account that the density of the vapor in Ref. 9 was significantly higher ($N_0=0.8-4\times 10^{17}$ cm^{-3}), we should expect that the instability of the laser beam will be displayed even more strongly and that the uncertainty in the intensity of the radiation will increase accordingly. Therefore, it would be quite incorrect to perform any quantitative comparison of the experimental data with the results of rough evaluations. Nevertheless, to determine the contribution of laser-induced associative ionization to the final extent of ionization we performed calculations of resonant breakdown with consideration of this process in the present work. The corresponding curves are shown in Figs. 7 and 11. As expected, at small detunings the plasma parameters practically coincide with the values obtained in the preceding calculations, and at the large detunings $\Delta\lambda>2$ nm they only slightly (by 10–20% for N_e and by 2–3% for T_e) exceed them.

In a second series of calculations the heating of the electrons in the strong laser field was described by Eqs. (8) and (10), the Rabi frequency appearing in (8) and (10) being calculated in accordance with the experimentally measured time dependence of the intensity of the laser radiation. The theoretical plots of $N_e(\Delta\lambda)$ and $T_e(\Delta\lambda)$ for this case are also presented in Figs. 7 and 11. Confirming the statements in Sec. 4, these curves approach the curves calculated with neglect of the strong-field effects in the two limiting cases of small and large detunings (compare curves *2* and *4* in Figs. 7 and 11), deviating from them only at intermediate values of $\Delta\lambda$. However, these differences are small and do not exceed $\sim$30% for N_e, and the difference in T_e is even smaller. It is noteworthy that the difference for the calculated electron density can be significantly greater under other conditions. This is because in the present experiment the resonant transition is still under strong saturation when $\Delta\nu=\nu_R$. The equality $\mathscr{I}=\mathscr{I}_s$, as we know, is satisfied at the detunings $\Delta\nu_s=\nu_R\sqrt{T_1/T_2}$. For $N_e\simeq 10^{16}$ cm^{-3} the ratio T_1/T_2 is approximately equal to 5, and $\Delta\nu_s\simeq 2.3\nu_R$. Therefore, when $\Delta\nu=\nu_R$, the value of N_2 still differs little from N_1, and, as follows from (8) and (9), the difference $\Delta Q=Q_{12}-Q'_{12}$ is also small. The maximum effect will obviously be achieved when $T_2=T_1$.

The calculations showed that at the larger detunings $\Delta\lambda>2$–2.5 nm the calculated values of the electron density in the resonant plasma in the two cases differ from one another by an amount that is smaller than the measurement error, and, thus, the data from the present experiment do not permit an unequivocal determination of the differences between the two mechanisms described above for heating electrons in a strong laser field. The range of small detunings, where the effect is more pronounced, is inaccessible, as has been stressed repeatedly, to a correct comparison of the theory with the experimental data. We note that this circumstance has a fairly fundamental character, since the range of detunings in which developed instability of the laser radiation is observed is determined specifically by the condition $\Delta\nu\sim\nu_R$ (see, for example, Ref. 26).

We estimate the accuracy of the calculations, which is determined by the use of the diffusion approximation, the approximation of a Maxwellian electron energy distribution function, the errors in the reaction constants, the neglect of several elementary processes, including the excitation of the rovibronic levels of the molecules and molecular ions, whose constants are unknown, neglect of the possible role of small impurities in the buffer gas, etc., to be 20–30%. The experimental error is approximately the same. In addition, the experimental data on the electron density that we obtained are determined to a significant extent by the error in the calculation of the Stark broadening parameters, which, as a rule, also amounts to 10–20%. There are alternative calculations[22,27] of the broadening parameters,[49] which give values that are 30% smaller. The use of the data from Ref. 49 leads to a corresponding increase in the measured electron density. Taking all this into account, we can assume that counting on an increase in the accuracy of the experiment and the numerical simulation is very problematic, at least at the present time.

6. CONCLUSIONS

In the present work resonant laser breakdown in a Na vapor with a density $\sim 10^{16}$ cm^{-3} has been investigated experimentally and theoretically. Quantitative agreement between the experimental data for the parameters of a dense resonant plasma and the results of calculations based on a homogeneous model has been obtained for the first time at the large detunings $\Delta\lambda>2$–2.5 nm, where the effects of the resonant spatial instability of the laser beam and the absorption of radiation practically vanish. We note that the effects of resonant ionization must also be taken into account to

investigate the actual instability of laser beams. The influence of the quantum stimulated inverse bremsstrahlung processes, which do not reduce to collisions of the second kind in the general case, on the energy gain by electrons has been analyzed for the first time in reference to concrete experimental data. The numerical calculations showed that this effect can significantly increase the density of the resonant plasma at detunings $\Delta\nu$ of the order of the Rabi frequency. However, in this range of detunings the instability processes still strongly disturb the laser beam, precluding a correct comparison with experiment. At large values of $\Delta\nu$ the two cases differ little from one another, and the value of N_e calculated for both mechanisms lie within the range of the measurement error. It was also shown in this work that the previously proposed mechanism of resonant breakdown, which was associated with laser-induced associative ionization, produces only a small correction to the final degree of ionization of the resonant plasma (10–20%).

The considerable width of the resonant dependence of N_e on the detuning ($\sim$4 nm) discovered both in the calculations and in the experiment calls for taking into account the possible considerable ionization of the gaseous medium when intense laser radiation passes through it even under off-resonance conditions. We note that on the basis of the results of the present work the creation of conducting channels for transporting charged particles by resonant laser ionization (see, for example, Ref. 6) would be promoted by using large detunings, at which the shape of the laser beam is maintained, the absorption is relatively small, but the ionization of the medium is still significant. For example, the calculations showed that at a detuning $\Delta\lambda \simeq 2$ nm a laser beam with an energy of only 0.1 J is needed to create a homogenous plasma channel with $N_e \sim 0.5\times 10^{16}$ cm^{-3}, a diameter of 1 mm, and a length of about 10 m when the pulse duration is $\sim 10^{-8}$ s.

This work was performed with support from the Russian Fund for Fundamental Research (Grants Nos. 96-02-17390 and 96-02-17391). We thank O. B. Popovincheva, A. M. Popov, and V. S. Sopov'ev for some useful discussions of the kinetics of resonant ionization, as well as A. V. Brazhnikov for assisting in performing the experiments.

[1] T. V. Lucatorto and T. J. McIlrath, Appl. Opt. **19**, 3948 (1980).
[2] V. A. Kas'yanov and A. I. Starostin, in *Plasma Chemistry* [in Russian], B. M. Smirnov (ed.), Énergoatomizdat, Moscow (1989).
[3] I. M. Beterov, A. V. Eletskiĭ, and B. M. Smirnov, Usp. Fiz. Nauk **155**, 265 (1988) [Sov. Phys. Usp. **31**, 535 (1988)].
[4] A. V. Eletskiĭ, Yu. I. Zaĭtsev, and S. V. Fomichev, Zh. Éksp. Teor. Fiz. **94** (5), 98 (1988) [Sov. Phys. JETP **67**, 920 (1988)].
[5] R. M. Measures, J. Quant. Spectrosc. Radiat. Transfer **10**, 107 (1970).
[6] R. M. Measures and P. G. Cardinal, Phys. Rev. A **23**, 804 (1981).
[7] V. A. Kas'yanov and A. N. Starostin, Zh. Éksp. Teor. Fiz. **76**, 944 (1979) [Sov. Phys. JETP **49**, 476 (1979)].
[8] B. Carré, F. Roussel, P. Breger *et al.*, J. Phys. B: At. Mol. Phys. **14**, 4271 (1981).
[9] O. L. Landen, R. J. Winfield, D. D. Burgess *et al.*, Phys. Rev. A **32**, 2963 (1985).
[10] Yu. P. Raizer, *Laser-Induced Discharge Phenomena*, Consultants Bureau, New York (1977).
[11] R. M. Measures, N. Drewell, and P. G. Cardinal, Appl. Opt. **18**, 1824 (1979).
[12] W. L. Morgan, Appl. Phys. Lett. **42**, 790 (1983).
[13] A. M. Popov, O. B. Popovicheva, and T. V. Rakhmova, Fiz. Plazmy **18**, 643 (1992) [Sov. J. Plasma Phys. **18**, 336 (1992)].
[14] A. P. Gavrilyuk and N. R. Shaparev, Zh. Tekh. Fiz. **63** (6), 1 (1993) [Tech. Phys. **38**, 435 (1993)].
[15] D. V. Gaydarenko, A. G. Leonov, A. A. Panteleev *et al.*, Laser Phys. **3**, 151 (1993).
[16] D. I. Chekov, D. V. Gaydarenko, A. G. Leonov *et al.*, Opt. Commun. **105**, 209 (1994).
[17] R. Kunnemeyer and M. Kock, J. Phys. B: At. Mol. Phys. **16**, L607 (1983).
[18] D. V. Gaĭdarenko and A. G. Leonov, JETP Lett. **53**, 302 (1991).
[19] D. V. Gaidarenko and A. G. Leonov, Laser Phys. **2**, 901 (1992).
[20] M. A. Capelli, P. G. Cardinal, H. Herchen *et al.*, Rev. Sci. Instrum. **56**, 2030 (1985).
[21] A. G. Leonov, A. A. Panteleev, A. N. Starostin *et al.*, Zh. Éksp. Teor. Fiz. **105**, 1536 (1994) [JETP **78**, 827 (1994)].
[22] H. R. Griem, *Spectral Line Broadening by Plasmas*, Academic Press, New York (1974).
[23] Yu. P. Raizer, *Gas Discharge Physics*, Springer-Verlag, Berlin–New York (1991).
[24] M. A. Capelli and R. M. Measures, Appl. Opt. **26**, 1058 (1987).
[25] F. Javan and P. L. Kelly, IEEE J. Quantum Electron. **QE-2**, 470 (1966).
[26] D. Harter and R. Boyd, Phys. Rev. A **29**, 739 (1984).
[27] J. Grinberg, G. Couland, and Nguyen-Hoe, Phys. Lett. A **57**, 227 (1976).
[28] A. N. Klyucharev and M. L. Yanson, *Elementary Processes in Alkali-Metal Plasmas* [in Russian], Énergoatomizdat, Moscow (1988).
[29] A. V. Eletskiĭ and Yu. N. Zaĭtsev, Teplofiz. Vys. Temp. **27**, 456 (1989) [High Temp. (USSR) **27**, 350 (1989)].
[30] A. A. Radtsig and B. M. Smirnov, *Parameters of Atoms and Monatomic Ions. A Handbook* [in Russian], Énergoatomizdat, Moscow (1986).
[31] B. M. Smirnov, *Excited Atoms* [in Russian], Énergoatomizdat, Moscow (1982).
[32] J. Huennekens and A. Gallagher, Phys. Rev. A **28**, 238 (1983).
[33] C. G. Carrington, D. M. Stacey, and J. Cooper, J. Phys. B: At. Mol. Phys. **6**, 417 (1973).
[34] I. I. Sobelman, L. A. Vainshtein, and E. A. Yukov, *Excitation of Atoms and Broadening of Spectral Lines*, Springer-Verlag, Berlin (1981).
[35] S. Y. Chen and M. Takeo, Rev. Mod. Phys. **28**, 20 (1957).
[36] R. H. Chatham, A. Gallagher, and E. L. Levis, J. Phys. B: At. Mol. Phys. **13**, L7 (1980).
[37] D. H. Grandall, G. H. Dunn, A. Gallagher *et al.*, Astrophys. J. **191**, 789 (1979).
[38] L. M. Biberman, V. S. Vorobev, and I. T. Yakubov, *Kinetics of Nonequilibrium Low-Temperature Plasmas*, Consultants Bureau, New York (1987).
[39] L. Vriens and A. H. M. Smeets, Phys. Rev. A **22**, 910 (1980).
[40] M. Aymer, E. Luc-Koeing, and F. Combet-Farnoux, J. Phys. B: At. Mol. Phys. **9**, 1279 (1976).
[41] C. Laughlin, J. Phys. B: At. Mol. Phys. **11**, 1399 (1978).
[42] S. Geltman, J. Phys. B: At. Mol. Phys. **10**, 3057 (1977).
[43] *Physical Constants. A Handbook* [in Russian], Énergoatomizdat, Moscow (1986).
[44] R. Shuker, A. Gallagher, and A. V. Phelps, J. Appl. Phys. **51**, 1306 (1980).
[45] A. V. Eletskiĭ, *Physics Encyclopedia* [in Russian], Nauchn. Izd. Bol'shaya Ross. Éntsiklopediya, Moscow (1994), Vol. 4, p. 322.
[46] K. Kirby-Docken, C. J. Cerjan, and A. Dalgamo, Chem. Phys. Lett. **40**, 205 (1976).
[47] V. A. Kas'yanov and A. N. Starostin, Kvantovaya Elektron. **8**, 1050 (1981) [Sov. J. Quantum Electron. **11**, 625 (1981)].
[48] N. Ya. Shaparev, Zh. Éksp. Teor. Fiz. **80**, 957 (1981) [Sov. Phys. JETP **53**, 487 (1981)].
[49] M. S. Dimitrijevic and S. Sahal-Brechot, J. Quant. Spectrosc. Radiat. Transfer **34**, 149 (1985).

Translated by P. Shelnitz

Statistical theory of the diffusion of a passive tracer in a random velocity field

V. I. Klyatskin

Institute of Atmospheric Physics, Russian Academy of Sciences, 109017 Moscow, Russia; Pacific Oceanological Institute, Russian Academy of Sciences, Far-Eastern Branch, 690041 Vladivostok, Russia

A. I. Saichev

Nizhniĭ Novgorod State University, 603600 Nizhniĭ Novgorod, Russia

(Submitted 16 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1297–1313 (April 1997)

A passive tracer on the surface of an incompressible liquid behaves like a tracer in two-dimensional compressible flows, whose characteristic feature is the formation of cluster structures, i.e., compact regions of increased density surrounded by vast low-density regions. The cluster formation dynamics are studied, and statistical spatiotemporal characteristics of the density fields, which faithfully reflect the properties of the cluster structures, are calculated.
 [S1063-7761(97)01304-8]

1. INTRODUCTION

When a passive tracer moves in random incompressible flows, its density field becomes increasingly disconnected with time, even if it is spatially smooth at first, and the lines of constant density acquire a fractal character.[1,2] This process is accompanied by a decrease in the characteristic spatial scales of the density field until the influence of molecular diffusion becomes significant. However, when the density is averaged over an ensemble of realizations, the fine structure of the realizations of the density field vanishes. It can be revealed only by analysis of the probabilistic properties of the density field. This is applicable to an even greater degree to the realizations of the density of a passive tracer. The main feature distinguishing it from an incompressible tracer is that a passive tracer on the surface of an incompressible liquid behaves like a tracer in two-dimensional compressible flows. Accordingly, besides the characteristic features described above, a passive tracer exhibits some qualitatively new effects, primarily the formation of clusters of particles surrounded by vast low-density regions.

We note that the clustering of a passive tracer was probably first detected in Refs. 3–5, in which numerical simulation of the so-called Eole experiment was carried out within very simple equations describing the dynamics of the atmosphere. As part of this global experiment, 500 balloons of constant density were to be launched in Argentina in 1970–1971 and then allowed to spread out over the entire southern hemisphere at a height of approximately 12 km. The spatial positions of these balloons were to be determined daily with the aid of a system of satellites. Unfortunately, this experiment was never carried out.

The object of the statistical theory of the diffusion of a passive tracer is to determine the significant qualitative and quantitative features of typical realizations of the density of the tracer from known statistical properties of the density field or some of its functionals. This problem is especially important for the atmosphere and oceans, where experimentalists generally deal only with individual realizations of the fields under investigation, rather than with ensembles. In the present work the probabilistic properties of the density field of a passive tracer, which make it possible to indicate important characteristic features of the behavior of the density realizations in time and in space, are investigated in a delta-correlated approximation for a random velocity field. On the conceptual level this work is closely related to the work in Ref. 2 and can be regarded as a continuation of it. We note that there are essentially no publications on this subject in the scientific literature.

2. FORMULATION OF THE PROBLEM AND STATISTICAL MODEL OF A VELOCITY FIELD

The evolution of the density of a passive tracer moving in a velocity field $\mathbf{u}(\mathbf{r},t)$ is governed by the equation

$$\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial \mathbf{r}}\mathbf{u}(\mathbf{r},t)\right)\rho(\mathbf{r},t)=\mu\Delta\rho(\mathbf{r},t),\quad \rho(\mathbf{r},0)=\rho_0(\mathbf{r}). \tag{1}$$

Its right-hand side takes into account molecular diffusion with the diffusion coefficient μ, the total mass of the tracer being conserved during evolution, i.e.,

$$M=\int \rho(\mathbf{r},t)d\mathbf{r}=\int \rho_0(\mathbf{r})d\mathbf{r}=\text{const}.$$

The velocity field is assumed to be a random Gaussian field that is incompressible [$\text{div}\ \mathbf{u}(\mathbf{r},t)\equiv 0$], statistically isotropic in space, and stationary in time and has the correlation and spectral tensors ($\langle\mathbf{u}(\mathbf{r},t)\rangle\equiv 0$)

$$\langle u_i(\mathbf{r},t)u_j(\mathbf{r}',t')\rangle=B_{ij}(|\mathbf{r}-\mathbf{r}'|,t-t'),$$

$$B_{ij}(r,t)=\int d\mathbf{k}E_{ij}(\mathbf{k},t)\exp(i\mathbf{k}\cdot\mathbf{r}), \tag{2}$$

$$E_{ij}(\mathbf{k},t)=E(k,t)\left(\delta_{ij}-\frac{k_ik_j}{k^2}\right),\quad i,\ j=1,2,3.$$

Let the passive tracer move in the $z=0$ plane, i.e., let its density be represented in the form

$$\rho(\mathbf{r},t)=\tilde{\rho}(\mathbf{R},t)\,\delta(z),\quad \mathbf{R}=(x,y),\quad \mathbf{r}=(\mathbf{R},z). \tag{3}$$

 1063-7761/97/040716-09$10.00

Substituting this expression into (1) and integrating it over all z, we obtain the equation

$$\left(\frac{\partial}{\partial t}+\mathbf{U}(\mathbf{R},t)\frac{\partial}{\partial \mathbf{R}}\right)\tilde{\rho}(\mathbf{R},t)+\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}\tilde{\rho}(\mathbf{R},t)$$
$$=\mu\frac{\partial^2}{\partial \mathbf{R}^2}\tilde{\rho}(\mathbf{R},t). \tag{4}$$

Here $\mathbf{U}(\mathbf{R},t)$ is the projection of the field $\mathbf{u}(\mathbf{R},0,t)$ onto the $z=0$ plane, and the summation is performed over the repeated indices $\alpha=1, 2$. Since we shall study only the properties of a passive tracer, we shall omit the tilde in the notation for its density everywhere below.

Clearly, the field $\mathbf{U}(\mathbf{R},t)$ is also a compressible random field that is Gaussian, statistically isotropic in the $z=0$ plane, and stationary with time and has the correlation tensor

$$\langle U_\alpha(\mathbf{R},t)U_\beta(\mathbf{R}',t')\rangle=B_{\alpha\beta}(|\mathbf{R}-\mathbf{R}'|,t-t'),$$

where

$$B_{\alpha\beta}(\mathbf{R},t)=\int d\mathbf{k}_\perp\ \exp(i\mathbf{k}_\perp\cdot\mathbf{R})\int_{-\infty}^{\infty}dk_z E(\mathbf{k}_\perp^2+k_z^2,t)$$
$$\times\left(\delta_{\alpha\beta}-\frac{k_{\perp\alpha}k_{\perp\beta}}{k_\perp^2+k_z^2}\right).$$

We recall the general formula relating the correlation tensor to the solenoidal component $E^s(k_\perp,t)$ and the potential component $E^p(k_\perp,t)$ of the spectral tensor of an arbitrary compressible, statistically isotropic velocity field:

$$B_{\alpha\beta}(\mathbf{R},t)=\int d\mathbf{k}_\perp\ \exp(i\mathbf{k}_\perp\cdot\mathbf{r})\left[E^s(k_\perp,t)\left(\delta_{\alpha\beta}\right.\right.$$
$$\left.\left.-\frac{k_{\perp\alpha}k_{\perp\beta}}{k_\perp^2}\right)+E^p(k_\perp,t)\frac{k_{\perp\alpha}k_{\perp\beta}}{k_\perp^2}\right].$$

Comparing it with the preceding expression, we find that the solenoidal and potential components of the spectral tensor of the velocity field on the surface of an incompressible liquid are related to the spectral tensor of the liquid by the equalities

$$E^s(k_\perp,t)=\int_{-\infty}^{\infty}dk_z E(k_\perp^2+k_z^2,t),$$

$$E^p(k_\perp,t)=\int_{-\infty}^{\infty}dk_z E(k_\perp^2+k_z^2,t)\frac{k_z^2}{k_\perp^2+k_z^2}.$$

Henceforth, in calculating the statistical properties of the density of a passive tracer we shall assume that its velocity field $\mathbf{U}(\mathbf{R},t)$ is delta-correlated in time and has the correlation tensor

$$\langle u_\alpha(\mathbf{R},t)u_\beta(\mathbf{R}',t')\rangle=2B_{\alpha\beta}^{\text{eff}}(|\mathbf{R}-\mathbf{R}'|)\,\delta(t-t'), \tag{5}$$

where

$$B_{\alpha\beta}^{\text{eff}}(R)=\frac{1}{2}\int_{-\infty}^{\infty}B_{\alpha\beta}(R,t)dt=\int_0^{\infty}B_{\alpha\beta}(R,t)dt.$$

A discussion of such a model can be found, for example, in Refs. 6 and 7. We note further that because $\mathbf{U}(\mathbf{R},t)$ is statistically isotropic, the following relations hold:

$$B_{\alpha\beta}^{\text{eff}}(0)=\frac{1}{2}D_0\delta_{\alpha\beta},\quad \frac{\partial}{\partial R_\gamma}B_{\alpha\beta}^{\text{eff}}(0)=0, \tag{6}$$

$$-\frac{\partial^2}{\partial R_\gamma\partial R_\delta}B_{\alpha\beta}^{\text{eff}}(0)=\frac{1}{8}D^s[3\delta_{\alpha\beta}\delta_{\gamma\delta}-\delta_{\alpha\gamma}\delta_{\beta\delta}-\delta_{\alpha\delta}\delta_{\beta\gamma}]$$
$$+\frac{1}{8}D^p[\delta_{\alpha\beta}\delta_{\gamma\delta}+\delta_{\alpha\gamma}\delta_{\beta\delta}$$
$$+\delta_{\alpha\delta}\delta_{\beta\gamma}],$$

where

$$D_0=\int_0^{\infty}dt\int d\mathbf{k}_\perp[E^s(k_\perp,t)+E^p(k_\perp,t)]$$
$$=\frac{4}{3}\int_0^{\infty}dt\int d\mathbf{k}\,E(k,t),$$

$$D^s=\int_0^{\infty}dt\int d\mathbf{k}_\perp k_\perp^2E^s(k_\perp,t)=\frac{2}{3}\int_0^{\infty}dt\int d\mathbf{k}\,k^2E(k,t), \tag{6'}$$

$$D^p=\int_0^{\infty}dt\int d\mathbf{k}_\perp k_\perp^2E^p(k_\perp,t)$$
$$=\frac{2}{15}\int_0^{\infty}dt\int d\mathbf{k}\,k^2E(k,t).$$

In this paper we shall neglect the effects of molecular diffusion. Then Eq. (4) can be simplified and takes the form

$$\left(\frac{\partial}{\partial t}+\mathbf{U}(\mathbf{R},t)\frac{\partial}{\partial \mathbf{R}}\right)\rho(\mathbf{R},t)+\rho(\mathbf{R},t)\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}=0,$$
$$\alpha=1,\ 2. \tag{7}$$

This first-order partial differential equation can also be solved using characteristic curves. Introducing the characteristic curves $\mathbf{R}(t)$, which obey the equation

$$\frac{d}{dt}\mathbf{R}(t)=\mathbf{U}(\mathbf{R},t),\quad \mathbf{R}(0)=\boldsymbol{\xi}, \tag{8}$$

we pass from (7) to the ordinary differential equation

$$\frac{d}{dt}\rho(t)=-\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}\rho(t),\quad \rho(0)=\rho_0(\boldsymbol{\xi}). \tag{9}$$

The solutions of Eqs. (8) and (9) have a graphical geometric interpretation. They describe the evolution of the density in the vicinity of a fixed tracer particle, whose trajectory is specified by $\mathbf{R}=\mathbf{R}(t)$. Now, as is seen from (9), the density in compressible flows varies, increasing in regions of compression and decreasing regions of rarefaction of the medium. The solutions of the system consisting of (8) and (9) depend on the initial coordinate $\boldsymbol{\xi}$ of the particle:

$$\mathbf{R}(t)=\mathbf{R}(t|\boldsymbol{\xi}),\quad \rho(t)=\rho(t|\boldsymbol{\xi}). \tag{10}$$

The components of the vector $\boldsymbol{\xi}$, which uniquely defines the position of an arbitrary particle, are called its Lagrangian

coordinates. Now Eqs. (8) and (9) correspond to the Lagrangian description of the evolution of the density field. The relationship between the Eulerian coordinates $\mathbf{R}$ of a particle in a stationary coordinate system and its Lagrangian coordinates is given by

$$\mathbf{R}=\mathbf{R}(t|\boldsymbol{\xi}). \tag{11}$$

Solving for $\boldsymbol{\xi}$, we obtain a relation which expresses the Lagrangian coordinates in terms of the Eulerian coordinates:

$$\boldsymbol{\xi}=\boldsymbol{\xi}(t,\mathbf{R}). \tag{12}$$

Then, eliminating the dependence on $\boldsymbol{\xi}$ in the last equality in (10) with the aid of (12), we arrive at the Eulerian description of the density:

$$\rho(\mathbf{R},t)=\rho(t|\boldsymbol{\xi}(t,\mathbf{R})). \tag{13}$$

3. LIOUVILLE'S EQUATION

The statistical properties of a density field can be investigated most completely on the basis of Liouville's equation. In this section we derive it and discuss the structure of this equation for both the Lagrangian and Eulerian probability distributions of the density field. We first discuss the Lagrangian description.

3.1. The Lagrangian description

In the Lagrangian representation the behavior of a passive tracer is described by the ordinary differential equations (8) and (9). It is easy to go from them to the linear Liouville equation in the corresponding phase space. To derive it, we introduce the function

$$\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho|\boldsymbol{\xi})=\delta(\mathbf{R}(t|\boldsymbol{\xi})-\mathbf{R})\,\delta(\rho(t|\boldsymbol{\xi})-\rho), \tag{14}$$

which has been written in a form that explicitly takes into account the dependence of the solution of the original dynamical equations on the Lagrangian coordinates $\boldsymbol{\xi}$. Differentiating (14) with respect to time and using Eqs. (8) and (9), we arrive at an equation which is equivalent to the original problem in first-order partial derivatives, which is known as Liouville's equation:

$$\frac{\partial}{\partial t}\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho|\boldsymbol{\xi})=\left[-\frac{\partial}{\partial\mathbf{R}}\mathbf{U}(\mathbf{R},t)+\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}\frac{\partial}{\partial\rho}\rho\right]\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho|\boldsymbol{\xi}), \tag{15}$$

$$\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho|\boldsymbol{\xi})=\delta(\boldsymbol{\xi}-\mathbf{R})\,\delta(\rho_0(\boldsymbol{\xi})-\rho). \tag{15$'$}$$

As we know (see, for example, Refs. 6–8), to go from a Lagrangian to a Eulerian description, we must introduce the transformation matrix for going from Lagrangian to Eulerian coordinates:

$$j_{\alpha\beta}(t|\boldsymbol{\xi})=\frac{\partial}{\partial\xi_\beta}R_\alpha(t|\boldsymbol{\xi}),$$

Differentiating (8) with respect to the components of $\boldsymbol{\xi}$, we arrive at the equations for the elements of the transformation matrix:

$$\frac{d}{dt}j_{\alpha\beta}(t|\boldsymbol{\xi})=\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\gamma}j_{\gamma\beta}(t|\boldsymbol{\xi}),\quad j_{\alpha\beta}(t=0|\boldsymbol{\xi})=\delta_{\alpha\beta}.$$

Hence it follows that the determinant $j(t|\boldsymbol{\xi})=|j_{\alpha\beta}(t|\boldsymbol{\xi})|$ of this matrix obeys the equation

$$\frac{d}{dt}j(t|\boldsymbol{\xi})=\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}j(t|\boldsymbol{\xi}),\quad j(t=0|\boldsymbol{\xi})=1. \tag{16}$$

Since the field $j(t|\boldsymbol{\xi})$ is a quantitative measure of the degree of compression or expansion of the physically infinitesimal liquid particles, we shall call it the divergence. Comparing Eqs. (9) and (16), we see that

$$\frac{d}{dt}\{\rho(t)j(t)\}=0,$$

and, therefore,

$$\rho(t|\boldsymbol{\xi})\equiv\rho_0(\boldsymbol{\xi})/j(t|\boldsymbol{\xi}). \tag{17}$$

We include the field $j(t|\boldsymbol{\xi})$ in the treatment, i.e., we replace $\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho|\boldsymbol{\xi})$ by the function

$$\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=\delta(\mathbf{R}(t|\boldsymbol{\xi})-\mathbf{R})\,\delta(\rho(t|\boldsymbol{\xi})-\rho)\,\delta(j(t|\boldsymbol{\xi})-j). \tag{18}$$

Liouville's equation then obviously generalizes Eq. (15):

$$\frac{\partial}{\partial t}\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=\left[-\frac{\partial}{\partial\mathbf{R}}\mathbf{U}(\mathbf{R},t)+\frac{\partial U_\alpha(\mathbf{R},t)}{\partial R_\alpha}\left(\frac{\partial}{\partial\rho}\rho-\frac{\partial}{\partial j}j\right)\right]\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho,j|\boldsymbol{\xi}), \tag{19}$$

$$\Phi_{\mathrm{Lag}}(t=0;\mathbf{R},\rho,j|\boldsymbol{\xi})=\delta(\boldsymbol{\xi}-\mathbf{R})\,\delta(\rho_0(\boldsymbol{\xi})-\rho)\,\delta(j-1).$$

3.2. Eulerian description

To describe the statistical properties of the density field in the Eulerian description, we introduce a function analogous to (14):

$$\Phi_{\mathrm{Eul}}(t;\mathbf{R},\rho)=\delta(\rho(t,\mathbf{R})-\rho). \tag{20}$$

The equation for it is easily obtained either directly from (7) (see, for example, Ref. 7) or on the basis of Liouville,s equation (19) in the Lagrangian representation. In fact, taking (12) into account, we write

$$\delta(\mathbf{R}(t|\boldsymbol{\xi})-\mathbf{R})=\frac{1}{|\partial R_\alpha/\partial\xi_\beta|}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}(t,\mathbf{R}))=\frac{1}{j(t|\boldsymbol{\xi})}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}(t,\mathbf{R})).$$

Thus, with consideration of the relation (13) and the definition (20) we have

$$\Phi_{\mathrm{Lag}}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=\frac{1}{j}\,\delta(\boldsymbol{\xi}-\boldsymbol{\xi}(t,\mathbf{R}))\,\delta(j(t|\boldsymbol{\xi})-j)\Phi_{\mathrm{Eul}}(t,\mathbf{R};\rho). \tag{21}$$

Therefore,

$$\Phi_{\rm Eul}(t,\mathbf{R};\rho)=\int d\boldsymbol{\xi}\int j\Phi_{\rm Lag}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})dj. \tag{22}$$

Multiplying through Eq. (19) by j and integrating it over j and $\boldsymbol{\xi}$, we find the corresponding Liouville equation in the Eulerian representation:

$$\left(\frac{\partial}{\partial t}+\mathbf{U}(\mathbf{R},t)\frac{\partial}{\partial\mathbf{R}}\right)\Phi_{\rm Eul}(t,\mathbf{R};\rho)$$
$$=\frac{\partial U_\gamma(\mathbf{R},t)}{\partial R_\gamma}\frac{\partial}{\partial\rho}\rho\Phi_{\rm Eul}(t,\mathbf{R};\rho), \tag{23}$$
$$\Phi_{\rm Eul}(0,\mathbf{R};\rho)=\delta(\rho_0(\mathbf{R})-\rho). \tag{23'}$$

4. STATISTICAL ANALYSIS

Let us proceed to a statistical analysis of a chaotically moving passive tracer. We shall, first of all, discuss such important characteristics of tracer particles as their position in space and their density. The diffusion of a tracer in a random velocity field is described by Liouville's equation (19) in the Lagrangian representation and by Eq. (23) in the Eulerian representation. Averaging them over the ensemble of realizations of the velocity field $\{\mathbf{U}\}$ leads to equations for the one-point Lagrangian probability distribution

$$P(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=\langle\Phi_{\rm Lag}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})\rangle \tag{24}$$

and the one-point Eulerian probability distribution

$$P(t,\mathbf{R};\rho)=\langle\Phi_{\rm Eul}(t,\mathbf{R};\rho)\rangle. \tag{25}$$

The main problem in deriving these equations is in decoupling the correlations of the velocity field $\mathbf{U}(\mathbf{R},t)$ and the fields $\Phi_{\rm Lag}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})$ and $\Phi_{\rm Eul}(t,\mathbf{R};\rho)$, which are functionally related to it. The decoupling methods depend on the nature of the random field $\mathbf{U}(\mathbf{R},t)$. If it is Gaussian, the correlations are decoupled using the Furutsu–Novikov formula (see, for example, 6 and 7):

$$\langle U_\alpha(\mathbf{R},t)\Phi[\mathbf{U}]\rangle=\int d\mathbf{R}'\int dt'\langle U_\alpha(\mathbf{R},t)U_\beta(\mathbf{R}',t')\rangle$$
$$\times\left\langle\frac{\delta}{\delta U_\beta(\mathbf{R}',t')}\Phi[\mathbf{U}]\right\rangle, \tag{26}$$

which is valid for an arbitrary functional $\Phi[\mathbf{U}]$ of the Gaussian field $\mathbf{U}(\mathbf{R},t)$ and is essentially a formula for integration by parts in a functional space. When applied to the field $\mathbf{U}(\mathbf{R},t)$ (5), which is delta-correlated in time, the equality (26) can be simplified, and takes on the form

$$\langle U_\alpha(\mathbf{R},t)\Phi[\mathbf{U}]\rangle=\int d\mathbf{R}'B^{\rm eff}_{\alpha\beta}(|\mathbf{R}-\mathbf{R}'|)$$
$$\times\left\langle\frac{\delta}{\delta U_\beta(\mathbf{R}',t-0)}\Phi[\mathbf{U}]\right\rangle. \tag{27}$$

4.1. Lagrangian description

Averaging Eq. (19) over an ensemble of realizations of the random field $\mathbf{U}(\mathbf{R},t)$, using Eq. (27), and taking into account the equality

$$\frac{\delta}{\delta U_\beta(\mathbf{R}',t-0)}\Phi_{\rm Lag}(t;\mathbf{R},\rho,j|\boldsymbol{\xi})$$
$$=-\left[\frac{\partial}{\partial R_\beta}\delta(\mathbf{R}-\mathbf{R}')+\frac{\partial\delta(\mathbf{R}-\mathbf{R}')}{\partial R_\beta}\left(\frac{\partial}{\partial\rho}-\frac{\partial}{\partial j}j\right)\right]$$
$$\times\Phi_{\rm Lag}(t;\mathbf{R},\rho,j|\boldsymbol{\xi}),$$

we arrive at the following Fokker–Planck equation for the Lagrangian probability distribution (24):

$$\frac{\partial}{\partial t}P(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=\left\{\frac{1}{2}D_0\frac{\partial^2}{\partial\mathbf{R}^2}+D^p\left(\frac{\partial}{\partial\rho}\rho^2\frac{\partial}{\partial\rho}\right.\right.$$
$$\left.\left.-2\frac{\partial^2}{\partial\rho\partial j}\rho j+\frac{\partial^2}{\partial j^2}j^2\right)\right\}P(t;\mathbf{R},\rho,j|\boldsymbol{\xi}), \tag{28}$$
$$P(0;\mathbf{R},\rho,j|\boldsymbol{\xi})=\delta(\boldsymbol{\xi}-\mathbf{R})\,\delta(\rho_0(\boldsymbol{\xi})-\rho)\,\delta(j-1).$$

The solution of Eq. (28) is

$$P(t;\mathbf{R},\rho,j|\boldsymbol{\xi})=P(t;\mathbf{R}|\boldsymbol{\xi})P(t;j|\boldsymbol{\xi})\,\delta(\rho-\rho_0(\boldsymbol{\xi})/j), \tag{29}$$

where

$$P(t;\mathbf{R}|\mathbf{R}')=\exp\left\{\frac{1}{2}D_0t\,\frac{\partial^2}{\partial\mathbf{R}^2}\right\}\delta(\mathbf{R}-\mathbf{R}')$$
$$=\frac{1}{2\pi D_0t}\,\exp\left[-\frac{(\mathbf{R}-\mathbf{R}')^2}{2D_0t}\right] \tag{30}$$

is the probability distribution of the coordinates of a passive tracer particle, and

$$P(t;j|\boldsymbol{\xi})=\exp\left\{D^pt\frac{\partial^2}{\partial j^2}j^2\right\}\delta(j-1)$$
$$=\frac{1}{2j\sqrt{\pi\tau}}\,\exp\left\{-\frac{\ln^2(je^\tau)}{4\tau}\right\} \tag{31}$$

is the probability distribution of the divergence field in its vicinity. The dimensionless time $\tau=D^pt$ is used in (31) and everywhere below. We stress that the solution (29) signifies statistical independence of the coordinates $\mathbf{R}(t|\boldsymbol{\xi})$ and the divergence $j(t|\boldsymbol{\xi})$ in the vicinity of a particle with Lagrangian coordinates $\boldsymbol{\xi}$, and that the Gaussian distribution (30) corresponds to the standard Brownian motion of a particle with parameters

$$\langle\mathbf{R}(t|\boldsymbol{\xi})\rangle=\boldsymbol{\xi},$$
$$\sigma^2_{\alpha\beta}(t)=\langle[R_\alpha(t)-\xi_\alpha][R_\beta(t)-\xi_\beta]\rangle=D_0\delta_{\alpha\beta}t, \tag{32}$$

and the log-normal distribution (31) signifies that $\chi(t|\boldsymbol{\xi})=\ln j(t|\boldsymbol{\xi})$ is distributed according to a Gaussian law with the parameters

$$\langle\chi(t|\boldsymbol{\xi})\rangle=-\tau,\quad \sigma^2_\chi(t)=2\tau. \tag{32'}$$

Hence, we have the following expressions for the moments of the random divergence field, which also follow directly from Eq. (28):

$$\langle j^n(t|\boldsymbol{\xi})\rangle=\exp[n(n-1)\tau],\quad n=\pm1,\pm2,\ldots \tag{32'}$$

We stress that the mean divergence is constant, i.e., $\langle j(t|\boldsymbol{\xi})\rangle = 1$, and its higher moments increase exponentially with the time. The corresponding cumulative distribution function has the form

$$F(t;j) = \int_0^j P(t;j)\,dj = \Phi\left(\frac{\ln(je^{\tau})}{2\sqrt{\tau}}\right), \tag{33}$$

where

$$\Phi(z) = \frac{1}{\sqrt{\pi}}\int_{-\infty}^{z} \exp(-y^2)\,dy$$

is the standard error integral. A useful asymptotic formula for the probability that the mean divergence is exceeded, which is valid for $\tau \gg 1$, follows from (33):

$$P(j(t|\boldsymbol{\xi}) > 1) = \Phi\left(-\frac{\sqrt{\tau}}{2}\right) \approx \frac{1}{\sqrt{\pi\tau}}\,\exp\left(-\frac{\tau}{4}\right). \tag{34}$$

We note further that according to (32″) and (17) we have the following expression for the Lagrangian moments of the density:

$$\langle \rho^n(t|\boldsymbol{\xi})\rangle = \rho_0^n(\boldsymbol{\xi})\,\exp[n(n+1)\tau], \tag{35}$$

which signifies, in particular, an exponential increase in both the mean density and its higher moments in the Lagrangian representation. Here the joint probability distribution of the density and the divergence has the form

$$P(t;\rho,j|\boldsymbol{\xi}) = P(t;j|\boldsymbol{\xi})\,\delta(\rho - \rho_0(\boldsymbol{\xi})/j), \tag{36}$$

where $P(t;j|\boldsymbol{\xi})$ is described by (31). Integrating Eq. (36) over j, we obtain the Lagrangian probability distribution of the density:

$$P(t;\rho|\boldsymbol{\xi}) = \frac{1}{2\rho\sqrt{\pi\tau}}\,\exp\left\{-\frac{\ln^2(\rho e^{-\tau}/\rho_0(\boldsymbol{\xi}))}{4\tau}\right\}. \tag{37}$$

It can also be obtained as the solution of the Fokker–Planck equation following from (28):

$$\frac{\partial}{\partial t}P(t;\mathbf{R},\rho|\boldsymbol{\xi}) = \left\{\frac{1}{2}D_0\frac{\partial^2}{\partial \mathbf{R}^2} + D^p\frac{\partial}{\partial\rho}\rho^2\frac{\partial}{\partial\rho}\right\}P(t;\mathbf{R},\rho|\boldsymbol{\xi}), \tag{38}$$

$$P(0;\mathbf{R},\rho|\boldsymbol{\xi}) = \delta(\boldsymbol{\xi}-\mathbf{R})\,\delta(\rho_0(\boldsymbol{\xi})-\rho).$$

The paradoxical behavior of the statistical characteristics of the divergence and the density revealed above, which takes the form of a simultaneous increase in their moment functions with time, is attributable to the properties of the log-normal probability distribution.[9] Thus, we can introduce a certain ''typical realization'' $j^*(t)$ of the random process $j(t|\boldsymbol{\xi})$ such that in any time interval (t_1,t_2) the mean time with satisfaction of the inequality $j(t|\boldsymbol{\xi}) < j^*(t)$ coincides with the mean time with satisfaction of the reverse inequality $j(t|\boldsymbol{\xi}) > j^*(t)$, which equals $(t_2-t_1)/2$ (Fig. 1). The typical realization is defined as the solution of the equation $F(t;j^*(t)) = 1/2$, where $F(t;j)$ is a cumulative distribution function and for a log-normal distribution

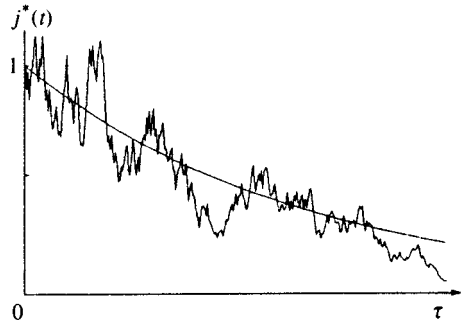

FIG. 1. Determination of the ''typical realization'' of the random process $j(t)$.

$j^*(t) = \exp\{-\langle \ln j(t|\boldsymbol{\xi})\rangle\}$. Thus, it follows from (32′) that the typical realization of the random divergence is an exponentially decaying function:

$$j^*(t) = \exp(-\tau). \tag{39}$$

There are upper estimates for realizations of the log-normal process $j(t|\boldsymbol{\xi})$.[9] For example, with the probability $p = 1/2$

$$j(t|\boldsymbol{\xi}) < 4\,\exp(-\tau/2)$$

in the time interval $t \in (0,\infty)$. Plots of this upper estimate and the realization $j(t|\boldsymbol{\xi})$ lying beneath it are shown in Fig. 2. Similarly, for density realizations we have the following typical realization and lower estimate:

$$\rho^*(t) = \rho_0\,\exp(\tau), \quad \rho(t|\boldsymbol{\xi}) > \rho_0\,\exp(\tau/2)/4.$$

We stress that the Lagrangian statistical properties of a passive tracer in compressible flows investigated above differ qualitatively from the properties of a tracer in incompressible media,[1,2,6,8] where $j(t|\boldsymbol{\xi}) \equiv 1$, while the density in the vicinity of a fixed particle is conserved: $\rho(t|\boldsymbol{\xi}) = \rho_0(\boldsymbol{\xi}) = \text{const}$. The estimates for a passive tracer presented above indicate that the statistics of the random processes $j(t|\boldsymbol{\xi})$ and $\rho(t|\boldsymbol{\xi})$ are shaped by the peaks of their realizations relative to the typical realizations.

At the same time, the probability distributions of the coordinates of the particles in the two cases of compressible and incompressible velocity fields are essentially the same.

The relative diffusion of two particles can be treated in a similar manner. For example, for the quantity $\mathbf{l}(t) = \mathbf{R}_1(t) - \mathbf{R}_2(t)$ we have the dynamical equation

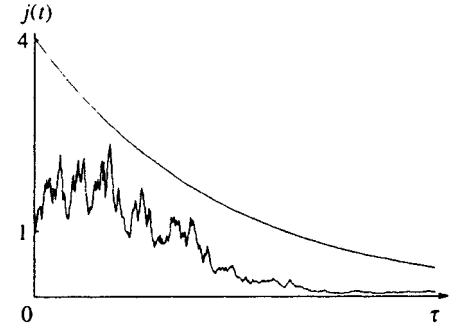

FIG. 2. Determination of the upper estimate of the random process $j(t)$.

$$\frac{d}{dt}\mathbf{l}(t)=\mathbf{U}(\mathbf{R}_1,t)-\mathbf{U}(\mathbf{R}_2,t),\quad \mathbf{l}(0)=\mathbf{l}_0,$$

which corresponds to the Fokker–Planck equation

$$\frac{\partial}{\partial t}P(t;\mathbf{l})=\frac{\partial^2}{\partial l_\alpha \partial l_\beta}D_{\alpha\beta}(\mathbf{l})P(t;\mathbf{l}),\quad P(0;\mathbf{l})=\delta(\mathbf{l}-\mathbf{l}_0),\qquad (40)$$

where $D_{\alpha\beta}(\mathbf{l})=2[B^{\text{eff}}_{\alpha\beta}(0)-B^{\text{eff}}_{\alpha\beta}(\mathbf{l})]$ is the structural matrix of the vector field $\mathbf{U}(\mathbf{R},t)$.

Equation (40) cannot be solved in the general case. However, if the initial distance between the particles $l_0 \ll l_{\text{cor}}$, where l_{cor} is the spatial correlation radius of the velocity field $\mathbf{U}(\mathbf{R},t)$, we can expand $D_{\alpha\beta}(\mathbf{l})$ in Taylor series. As a result, in a first approximation we obtain

$$D_{\alpha\beta}(\mathbf{l})=-\left.\frac{\partial^2 B^{\text{eff}}_{\alpha\beta}(\mathbf{l})}{\partial l_\alpha \partial l_\beta}\right|_{\mathbf{l}=0} l_\alpha l_\beta .$$

Using the representation (6), we can simplify the diffusion tensor $D_{\alpha\beta}(\mathbf{l})$ and write it in the form

$$D_{\alpha\beta}(\mathbf{l})=\frac{1}{8}[3D^s+D^p]\mathbf{l}^2\delta_{\alpha\beta}-\frac{1}{4}[D^s-D^p]l_\alpha l_\beta .\qquad (41)$$

Now, substituting (41) into (40), multiplying the two sides of the equation by l^n and integrating over $\mathbf{l}$, we obtain the closed equation

$$\frac{d}{dt}\langle l^n(t)\rangle=\frac{1}{8}n[n(D^s+3D^p)+2(D^s-D^p)]\langle l^n(t)\rangle,$$

whose solution corresponds to exponentially increasing functions of the time for all the moments ($n=1,2,\ldots$). Now the probability distribution for the random process $l(t)/l_0$ will be a log-normal distribution with the parameters

$$\left\langle \ln\left[\frac{l(t)}{l_0}\right]\right\rangle=\frac{1}{4}(D^s-D^p)t,$$

$$\sigma^2_{\ln[l(t)/l_0]}=\frac{1}{4}(D^s+3D^p)t.$$

Therefore, the typical realization for the distance between two particles will be the exponential function of the time

$$l^*(t)=l_0\ \exp\left\{\frac{1}{4}(D^s-D^p)t\right\},$$

which increases or decays, depending on the sign of the difference D^s-D^p. In particular, for the incompressible case ($D^p=0$) we have an exponentially increasing typical realization, which corresponds to exponentially rapid movement of the particles with small distances between them. This result is valid for the times

$$\frac{1}{4}D^s t\ll\ln\left(\frac{l_{\text{cor}}}{l_0}\right),$$

at which the expansion (41) is valid. In the other limiting case of a potential velocity field ($D^s=0$), the typical realization will be an exponentially decreasing function, i.e., the particles will tend to ''fuse.'' Taking into account that liquid particles are then compressed, we see that clusters, i.e., regions with compact concentrating of particles, which are located largely in rarefied regions, should form in this case.

Figure 3 presents the results of the numerical simulation of the evolution of a realization of an initially uniform distribution of particles in the random potential of a velocity field (according to the data in Ref. 10), which convincingly supports the conclusion regarding the formation of clusters. This picture is very reminiscent of the picture of the appearance of caustics of an optical field in randomly inhomogeneous media. This is no accident, since the dynamics of optical radiation in the approximation of geometric optics are also described by the stochastic equation (7), where $\mathbf{U}(\mathbf{R},t)$ is the gradient of the phase fluctuations of the wave (the parameter t plays the role of the distance traversed by the wave, and the density becomes the intensity of the optical wave). A similar pattern can be observed on the bottom of a pool or in shallow water on a sunny day when the surface of the water is gently agitated.

Returning to the analysis of the problem of the diffusion of a passive tracer in a random velocity field, we see that (6′) implies $D^s-D^p>0$, and therefore the closely arranged particles should move apart in an initial relatively short time interval. Nevertheless, as will be shown below, the presence of the potential component of the velocity field $\mathbf{U}(\mathbf{R},t)$ ensures cluster formation in the density field at large times.

4.2. Eulerian description

Above all, we note that in the case of a delta-correlated random velocity field, it is easy to go from the linear equation (4) to closed equations for both the mean density of a passive tracer and for high moments of the functions. For example, after averaging Eq. (4), we can use the Furutsu–Novikov equation (27) and the expression following from (4) for the variational derivative

$$\frac{\delta_\rho(\mathbf{R},t)}{\delta U_\alpha(\mathbf{R}',t-0)}=-\frac{\partial}{\partial R_\alpha}\delta(\mathbf{R}-\mathbf{R}')\rho(\mathbf{R},t),$$

to obtain the equation for the mean density of the tracer

$$\frac{\partial}{\partial t}\langle\rho(\mathbf{R},t)\rangle=\left(\frac{1}{2}D_0+\mu\right)\frac{\partial^2}{\partial\mathbf{R}^2}\langle\rho(\mathbf{R},t)\rangle.\qquad (42)$$

Under the condition $D_0\gg\mu$ ($\mu\ll\sigma_u^2 l_{\text{cor}}^2$, where σ_u^2 is the variance of the random velocity field and l_{cor} is its correlation radius), Eq. (42) coincides with the equation for the probability distribution (30) of the coordinates of a particle, and, therefore, the diffusion coefficient, which is analogous to the diffusion coefficient in an incompressible random velocity field, characterizes only the scales of the regions where the tracer is concentrated as a whole, and it contains no information on the local structure of the density realizations.

To describe the local behavior of the realizations of a passive tracer we must know the probability distribution of its density. When molecular diffusion is neglected, the equation for the Eulerian probability distribution of the density can easily be derived by multiplying Eq. (28) by j and integrating it over all possible values of j and ξ. We ultimately find the equation for the probability distribution (25):

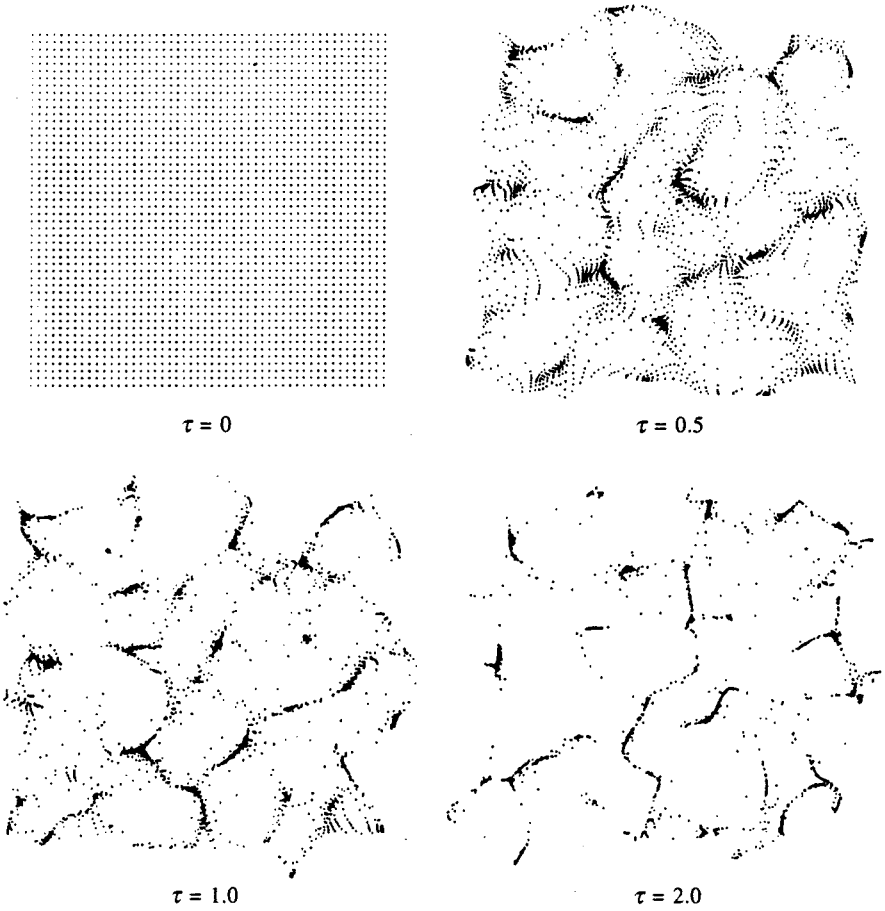

FIG. 3. Temporal evolution of an initially uniform distribution of particles in a realization of a potential random velocity field.

$$\left(\frac{\partial}{\partial t}-\frac{1}{2}D_0\frac{\partial^2}{\partial \mathbf{R}^2}\right)P(t,\mathbf{R};\rho)=D^p\frac{\partial^2}{\partial \rho^2}\rho^2 P(t,\mathbf{R};\rho), \quad (43)$$

$$P(0,\mathbf{R};\rho)=\delta(\rho_0(\mathbf{R})-\rho).$$

Hence it follows, in particular, that the moment functions of the density field are described by the equation

$$\left(\frac{\partial}{\partial t}-\frac{1}{2}D_0\frac{\partial^2}{\partial \mathbf{R}^2}\right)\langle\rho^n(\mathbf{R},t)\rangle=D^p n(n-1)\langle\rho^n(\mathbf{R},t)\rangle,$$

$$\langle\rho^n(\mathbf{R},0)\rangle=\rho_0^n(\mathbf{R}). \quad (44)$$

Its solution has the structure

$$\langle\rho^n(\mathbf{R},t)\rangle=\exp[n(n-1)\tau]\int d\mathbf{R}'P(t;\mathbf{R}|\mathbf{R}')\rho_0^n(\mathbf{R}'). \quad (45)$$

Thus, if the initial density of the tracer is identical everywhere, i.e., if $\rho_0(\mathbf{R})=\rho_0=\text{const}$, the probability distribution of the density does not depend on $\mathbf{R}$ and has a log-normal character with the cumulative distribution function

$$F(t;\rho)=\Phi\left(\frac{\ln(\rho e^{\tau}/\rho_0)}{2\sqrt{\tau}}\right). \quad (46)$$

In this case

$$\langle\rho(\mathbf{R},t)\rangle=\rho_0, \quad \langle\rho^n(\mathbf{R},t)\rangle=\rho_0^n\exp[n(n-1)\tau], \quad (47)$$

and the typical realization of the density field at any given point is

$$\rho^*(t)=\rho_0\exp(-\tau). \quad (48)$$

The Eulerian statistics of the density are shaped by the density fluctuations relative to this curve, attesting to the cluster character of the fluctuations of the density of the medium in arbitrary compressible flows and, in particular, in the case of a passive tracer under consideration.

5. ELEMENTS OF STATISTICAL TOPOGRAPHY

Below we shall discuss the one-point probability distribution of the tracer density in the Eulerian representation, which has already enabled us to draw several conclusions regarding the behavior of the realizations of a density field with time at fixed points in space. We now show that this distribution also makes it possible to elucidate some characteristic features of the spatiotemporal structure of the realizations of the density field. We note that important information on the spatial behavior of the realizations can be provided by an analysis of the level contours defined by the equality

$$\rho(\mathbf{R},t)=\rho=\text{const}, \quad (49)$$

the areas of the regions demarcated by these level contours

$$S(t,\rho)=\int\theta(\rho(\mathbf{R},t)-\rho)d\mathbf{R}, \quad (50)$$

and the total mass of the tracer contained in these regions

$$M(t,\rho)=\int\rho(\mathbf{R},t)\theta(\rho(\mathbf{R},t)-\rho)d\mathbf{R}. \quad (51)$$

Here $\theta(z)=1$ if $z>0$, and $\theta(z)=0$ if not (Heaviside step function). The statistical means of the areas and masses just described can obviously be expressed in terms of the solution of Eq. (43):

$$\langle S(t,\rho)\rangle=\int_{\rho}^{\infty}d\tilde{\rho}\int d\mathbf{R}P(t,\mathbf{R};\tilde{\rho}),$$

$$\langle M(t,\rho)\rangle=\int_{\rho}^{\infty}\tilde{\rho}d\tilde{\rho}\int d\mathbf{R}P(t,\mathbf{R};\tilde{\rho}). \tag{52}$$

After plugging in the solution of Eq. (43) and performing some relatively simple transformations, we find explicit expressions for the mean areas and masses of the particles within the level contours (49):

$$\langle S(t,\rho)\rangle=\int d\mathbf{R}\Phi\left(\frac{\ln[\rho_0(\mathbf{R})e^{-\tau}/\rho]}{2\sqrt{\tau}}\right), \tag{53}$$

$$\langle M(t,\rho)\rangle=\int \rho_0(\mathbf{R})d\mathbf{R}\Phi\left(\frac{\ln[\rho_0(\mathbf{R})e^{\tau}/\rho]}{2\sqrt{\tau}}\right).$$

Hence it is clear, in particular, that for $\tau\gg 1$ the area of the regions where the density is above the level ρ decreases exponentially with time according to the equation

$$\langle S(t,\rho)\rangle=\frac{1}{\sqrt{\pi\tau\rho}}\exp\left(-\frac{\tau}{4}\right)\int\sqrt{\rho_0(\mathbf{R})}d\mathbf{R}, \tag{54}$$

while the mass of the tracer contained in these regions

$$\langle M(t,\rho)\rangle=M-\sqrt{\frac{\rho}{\pi\tau}}\exp\left(-\frac{\tau}{4}\right)\int\sqrt{\rho_0(\mathbf{R})}d\mathbf{R} \tag{55}$$

tends monotonically to its total mass

$$M=\int\rho_0(\mathbf{R})d\mathbf{R}.$$

This again confirms the previous conclusion that the tracer particles tend to gather in clusters, i.e., compact regions of enhanced density surrounded by rarefied regions, with time.

The dynamics of cluster formation can be illustrated in the case in which the original passive tracer is uniformly distributed in the plane: $\rho_0(\mathbf{R})=\rho_0=\text{const}$. Here, by analogy with (52), the specific area of the regions within which $\rho(\mathbf{R},t)>\rho$ is

$$s(t,\rho)=\int_{\rho}^{\infty}P(t;\tilde{\rho})d\tilde{\rho}=\Phi\left(\frac{\ln(\rho_0e^{-\tau}/\rho)}{2\sqrt{\tau}}\right), \tag{56}$$

where $P(t,\rho)$ is the solution of Eq. (43) that does not depend on $\mathbf{R}$, and the specific mean mass of the tracer concentrated in these regions is described by the expression

$$m(t,\rho)=\frac{1}{\rho_0}\int_{\rho}^{\infty}\tilde{\rho}P(t;\tilde{\rho})d\tilde{\rho}=\Phi\left(\frac{\ln(\rho_0e^{\tau}/\rho)}{2\sqrt{\tau}}\right). \tag{57}$$

It follows from (56) and (57) that at long times their specific area decreases exponentially according to the expression

$$s(t,\rho_0)=\Phi\left(-\frac{\sqrt{\tau}}{2}\right)\approx\frac{1}{\sqrt{\pi\tau}}\exp\left(-\frac{\tau}{4}\right), \tag{58}$$

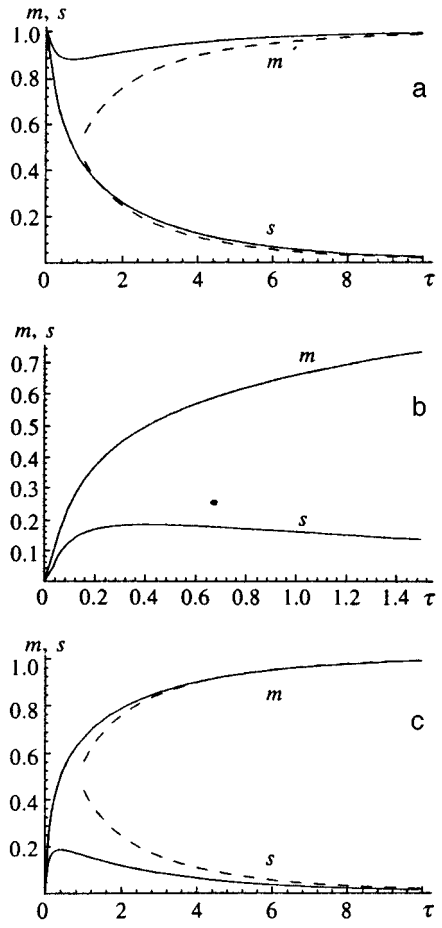

FIG. 4. Dynamics of the formation of ''cluster regions.''

regardless of the value of ρ/ρ_0, while essentially all the tracer particles gather within them:

$$m(t,\rho_0)=\Phi\left(\frac{\sqrt{\tau}}{2}\right)\approx 1-\frac{1}{\sqrt{\pi\tau}}\exp\left(-\frac{\tau}{4}\right). \tag{59}$$

The character of the temporal evolution of the formation of the cluster structure depends significantly on the value of ρ/ρ_0. For example, if $\rho/\rho_0<1$, we initially have $s(0,\rho)=1$ and $m(0,\rho)=1$. Then, because the passive tracer particles tend to move apart in the initial period, small regions with $\rho(\mathbf{R},t)<\rho$ form, which contain an insignificant part of the total mass. With the passage of time, these regions rapidly increase in size, their mass passes into the cluster regions, and the asymptotic dependences (58) and (59) are achieved quite rapidly (Fig. 4a). We note further that at the time $\tau^*=\ln(\rho_0/\rho)$ the area $s(t^*,\rho)=1/2$.

In the opposite case of $\rho/\rho_0>1$, we initially have $s(0,\rho)=0$ and $m(0,\rho)=0$. The initial movement of the particles produces small cluster regions with $\rho(\mathbf{R},t)>\rho$, which are essentially conserved with time and rapidly draw in a considerable part of the total mass (Fig. 4b). Subsequently, the areas of these regions decrease with the passage of time, but the mass contained in them increases according to the asymptotic dependences (58) and (59) (Fig. 4c).

6. CONCLUSIONS

We have thoroughly discussed the one-point probability properties of a passive tracer moving in a compressible chaotic velocity field. As was shown in this work, even the simplest statistical characteristics of the density and divergence fields enable one to draw useful conclusions regarding the fine structure of the density field and, in particular, regarding the clustering of the passive tracer. A more complete statistical topography of the tracer, including, for example, determination of the statistics of the lengths of level contours, the number of level contours, the mass of the particles contained in clusters, etc. requires knowledge of the statistical properties of the gradients of the density field and its higher derivatives.[2,8,11] We note that the conclusions of the present work are applicable not only to the idealized case of a passive tracer in the $z=0$ plane, but also to a real agitated surface $z=Z(\mathbf{R},t)$ on which a passive tracer is distributed. In fact, instead of (3) we then have

$$\rho(\mathbf{r},t)=\tilde{\rho}(\mathbf{R},t)\,\delta(Z(\mathbf{R},t)-z),$$

where the function $\tilde{\rho}(\mathbf{R},t)$ obeys the equation

$$\left(\frac{\partial}{\partial t}+\mathbf{u}(\mathbf{R},Z(\mathbf{R},t),t)\,\frac{\partial}{\partial \mathbf{R}}\right)\tilde{\rho}(\mathbf{R},t)$$
$$+\frac{\partial U_\alpha(\mathbf{R},Z(\mathbf{R},t),t)}{\partial R_\alpha}\tilde{\rho}(\mathbf{R},t)=0.$$

When the velocity field and the agitated surface are statistically independent, it is easy to see in the delta-correlated approximation of a velocity field that the statistical properties of the density do not depend on the agitation of the surface and coincide with the properties of the tracer density on an ideal plane studied above.

We thank E. M. Dobryshman, who directed our attention to the work in Refs. 3–5. The present work was carried out with the partial support of the Russian Fund for Fundamental Research (Projects Nos. 95-05-14247 and 96-05-65347).

[1] V. I. Klyatskin and W. A. Woyczynski, Zh. Éksp. Teor. Fiz. **108**, 1403 (1995) [JETP **81**, 770 (1995)].
[2] V. I. Klyatskin, W. A. Woyczynski, and D. Gurarie, J. Stat. Phys. **84**, 797 (1996).
[3] F. Mesinger and Y. Mintz, Technical Report No. 4, Department of Meteorology, University of California, Los Angeles (1970).
[4] F. Mesinger and Y. Mintz, Technical Report No. 5, Department of Meteorology, University of California, Los Angeles (1970).
[5] F. Mesinger, Mon. Weather Rev. **99**, 15 (1971).
[6] V. I. Klyatskin, Usp. Fiz. Nauk **164**, 531 (1994) [Phys. Usp. **37**, 501 (1994)].
[7] V. I. Klyatskin, *Stochastic Equations and Waves in Random Inhomogeneous Media* [in Russian], Nauka, Moscow (1980).
[8] A. I. Saichev and W. A. Woyczynski, in *Stochastic Models in Geosystems (IMA Vol. 85)*, Springer-Verlag, New York (1996), p. 359.
[9] V. I. Klyatskin and A. I. Saichev, Usp. Fiz. Nauk **162**, 161 (1992) [Sov. Phys. Usp. **35**, 231 (1992)].
[10] C. L. Zirbel and E. Cinlar, in *Stochastic Models in Geosystems (IMA Vol. 85)*, Springer-Verlag, New York (1996), p. 459.
[11] M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992).

Translated by P. Shelnitz

Dynamic fractal structure of emulsions due to motion and interaction of the particles: numerical simulation

V. V. Zosimov and D. N. Tarasov

Scientific-Research Institute of Applied Acoustics, 141980 Dubna, Moscow Region, Russia
(Submitted 9 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1314–1319 (April 1997)

The method of numerical simulation is used to study the geometrical structure of micro-emulsions in the plane. It is found that the interaction between the particles leads to the formation of a dynamic homogeneous fractal structure of the micro-emulsion. In the absence of any interaction between the particles the structure of the emulsion is homogeneous. The interaction energy of the particles at which the fractal inhomogeneity arises is close in magnitude to the interaction energy of the particles in real (e.g., aqueous) micro-emulsions. It is also found that the size of the inhomogeneities (correlation radius) depends on the particle density in the system and is largest for the density of the percolation transition. The numerical simulation data qualitatively coincide with the results of measurements in real micro-emulsions. © *1997 American Institute of Physics.* [S1063-7761(97)01404-2]

Fractal structures have become a popular object of study in recent years,[1,2] but are by no means exotic objects and may be encountered in completely ordinary situations. In a study of the mechanical properties of aqueous micro-emulsions the authors of Refs. 3 and 4 advanced the hypothesis that the geometrical structure of micro-emulsions contains inhomogeneities which are characterized by a fractal structure. In Ref. 5, using the method of rapid freezing, Ozhovan investigated the structure of a micro-emulsion of silicate glass in a vitrophobic liquid melt. Ozhovan discovered that for some fraction of the dispersed phase in some size range the geometrical structure of the micro-emulsion has inhomogeneities characterized by a fractal structure. The analysis of these data in Ref. 5 was based on the concept of a connecting cluster, arising during a percolation transition. In general, it is this connecting cluster that possesses fractal dimensionality. Near the percolation threshold the number of particles entering into this cluster is significantly less than the total number of particles in the system. The measurements reported in Ref. 5 allowed for all the particles in the system. Here in the case when there is no interaction between the particles, the system as a whole should be homogeneous. Thus, the fractal structure of a micro-emulsion cannot be explained by the percolation transition.

Other models of fractal structure formation in two-phase dispersive systems are based on a consideration of processes of particle aggregation for different laws of motion and interaction of the particles. To analyze these processes, numerical simulation methods have acquired the most widespread use. Such methods, along with their simplicity, give quite good results, agreeing with the experimental data to within 2 %.[6] What distinguishes processes of solid particle aggregation from the formation of inhomogeneities in a micro-emulsion is the following: although conditions exist in which solid particles have a small probability of forming a bond in a collision,[7] in the final count after some time clusters are formed which from then on do not change, i.e., the process is irreversible. A micro-emulsion exists in thermodynamic equilibrium, which excludes irreversible processes. Thus, inhomogeneities in micro-emulsions, as Ozhovan pointed out,[5] have a dynamic character, i.e., they are observed not only at certain spatial scales, but also on certain time scales.

The aim of the present paper is to explain the appearance of dynamic fractal structures in emulsions and their connection with the percolation transition. The problem is solved by numerical simulation.

A two-dimensional model which allows for all the above features is constructed as follows: 200 particles are placed in a cell with periodic boundary conditions. For a fixed number of particles their density varies due to variation of the size of the cell. The particles are modeled by solid disks of radius r_0 and mass m_0, and the interaction between the solid disks by a constant force acting over a distance of $0.1r_0$. The motion of the particles is simulated, as in the well-known method of molecular dynamics.[8] At the initial time an equilibrium distribution of particle velocities with temperature T is prescribed. At each subsequent time the velocities and forces acting between the particles are calculated and simultaneous mixing of all particles is effected. In contrast to the molecular dynamics method, to model the Brownian nature of the particle motion after a time t_0 the velocities of all particles are replaced by new values in accord with the distribution at the given temperature. The time t_0 is equal in order of magnitude to the time it takes a particle on average to move through the distance $0.1r_0$.

To determine the distribution of the particles in the system, the dependence $g(r)$ of the number of particles in a circle of radius r on this radius is found. At the initial time the particles are distributed in the cell as solid shells.[9] The function $g(r)$ is monitored as a function of time. The calculations are carried out until the values of $g(r)$ cease to vary for some set length of time; in this case it is assumed that the system has reached equilibrium.

Let us consider the behavior of the modeled system for different interaction energies between the particles. Let $E=0$, i.e., let the particles be modeled by solid disks without any interaction potential. In this series of runs the particle

 1063-7761/97/040725-04$10.00

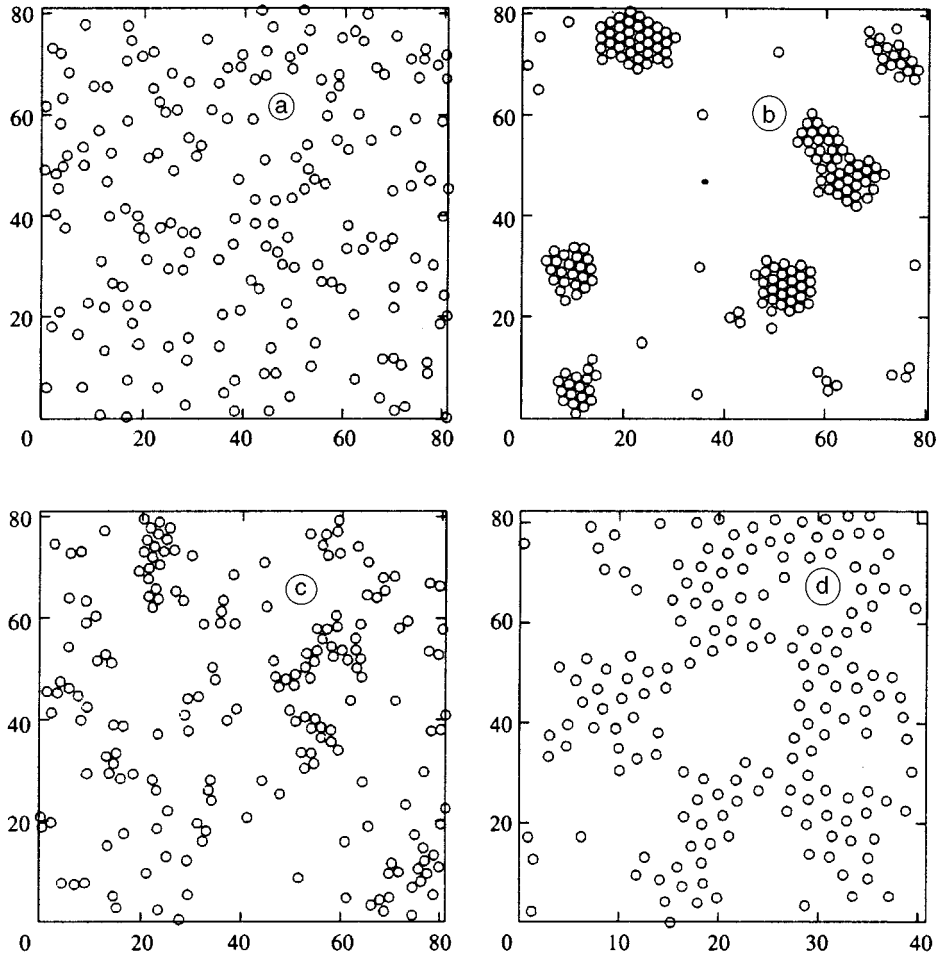

FIG. 1. System of particles after completion of simulation, $r_0=1$. a) $C=0.09$, $E=0$, b) $C=0.09$, $E=2.5$ kT, c) $C=0.09$, $E=1.3$ kT, d) $C=0.25$, $E=1.3$ kT.

concentration C in the system was varied from 0.01 to 0.5. It was found that, independent of the concentration, the function $g(r)$ does not vary with time after the commencement of particle motion and the values of $g(r)$ completely coincide with the values obtained theoretically by Ambartsumyan *et al.*[9] for the process of solid shells. In this case the geometrical structure is homogeneous and does not possess fractal properties. Thus, simple Brownian motion of solid spheres does not lead to fractal inhomogeneities in the system.

The introduction of a short-range model potential substantially alters the situation (Fig. 1). In this series of computer experiments the particle concentration remained constant and equal to 0.09. If the interaction energy between the particles satisfies $E>3kT$, then the particles form associates. A particle falling into such an associate has a very small probability of leaving it; for this reason with time each associate only increases in size as a result of collisions with free particles and other associates. As a result, at some time only one associate remains in the system, consisting of particles packed as they would be in a crystalline lattice. It turns out that after a time for a large enough interaction between the particles the system separates into two phases. Indeed, emulsions of liophobic liquids very often are unstable, and with time a separation of phases takes place.[10]

When the interaction energy between the particles satisfies $E<kT$ the function $g(r)$ almost coincides with the function $g(r)$ for $E=0$, and the system remains practically homogeneous.

For energies $kT<E<3kT$ inhomogeneities are formed in the system, that is to say, clusters. A particle, falling upon a cluster, after some time leaves it. After a certain time the

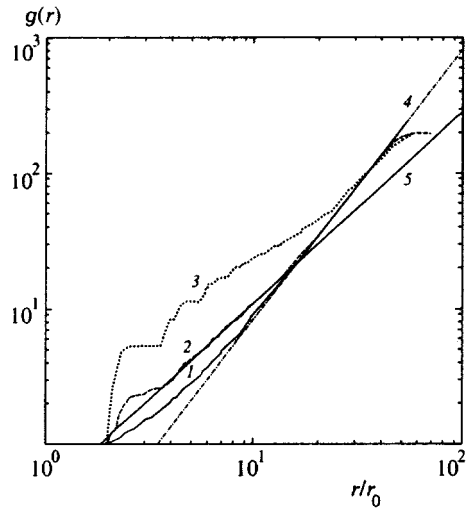

FIG. 2. Log–log plot of the function $g(r)$ for various interaction energies between the particles, $C=0.09$: *1)* $E=0$, *2)* $E=1.3$ kT, *3)* $E=2.5$ kT, *4)* $g(r)\propto r^2$ (homogeneous distribution), *5)* $g(r)\propto r^{1.42}$ (fractal distribution).

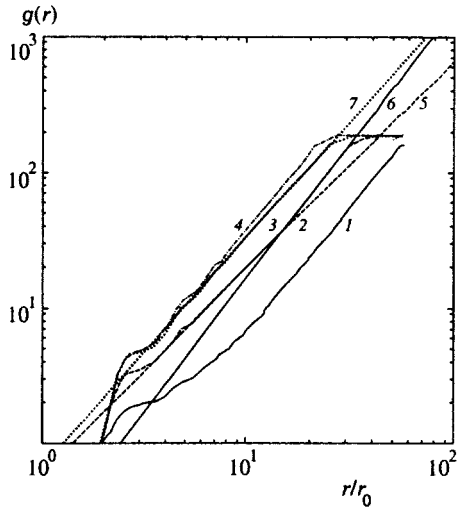

FIG. 3. Log-log plot of the function $g(r)$ for the interaction energies between the particles $E=1.3$ kT: *1)* $C=0.05$, *2)* $C=0.15$, *3)* $C=0.26$, *4)* $C=0.33$, *5)* $g(r) \propto r^{1.45}$ (fractal distribution), *6)* $g(r) \propto r^2$ (homogeneous distribution), *7)* $g(r) \propto r^{1.72}$ (fractal distribution).

clusters decay and new clusters form in a different place. In this case, for $3r_0 < r < 10r_0$ and $E=1.3kT$ we have $g(r) \propto r^{1.42}$ (Fig. 2). Consequently, the interaction between the particles leads to the result that the geometrical structure of the system becomes fractal.

The value $\sim(1-2)kT$ of the interaction energy between the particles in an emulsion is completely realistic. Thus, for example, it is generally assumed that an interaction exists between hydrophobic particles in water, due to a breakdown of the hydrogen bonds in water caused by the hydrophobic particles and equal in order of magnitude to the energy of the hydrogen bond between the water molecules.[11] The energy of the hydrogen bond between two water molecules is $E_H \approx 5$ kcal/mole, which at room temperature $T=300$ K gives $E_H \approx 8kT$, which coincides in order of magnitude with the energies at which fractal inhomogeneities are formed in the simulated system.

In the next series of computer experiments we varied the particle concentration from $C=0.06$ to $C=0.4$ while keeping the interaction energy between the particles constant and equal to $1.3kT$. The function $g(r)$ is plotted in Fig. 3 for various concentrations. It can be seen that at low concentrations up to $C=0.08$ the dependence $g(r)$ corresponds to the homogeneous distribution that obtains in the model of an ideal gas or ideal solution. At concentrations above 0.08 regions of inhomogeneity begin to be formed which have a fractal dimension D: $g \propto r^D$, where the correlation function is described by a power law only for $4r_0 < r < \xi$, where ξ is the correlation radius. The dependence of D and ξ on the particle concentration in the system is shown in Fig. 4.

A decrease of the correlation radius ξ with further increase of the particle concentration is characteristic for many other cases where fractal inhomogeneities are present in the system, e.g., in the case of solutions of polymers[12] or in cluster–cluster aggregation models.[13] This phenomenon is explained by the fact that the concentration of the inhomogeneity–clusters increases as the particle concentration is increased. At a particle concentration equal to the critical concentration of the percolation transition, the clusters begin to overlap. As a result, the correlation radius ξ decreases.

In the case of polymer solutions the correlation radius ξ for concentrations below the concentration of the percolation transition (the critical concentration) does not change and is equal to the diameter of the coiled-up macromolecule.[12] In the case of cluster–cluster aggregation, percolation depends on the size of the simulation region. If the simulation region is enlarged without limit, the critical concentration of the percolation transition decreases without limit.[13] In our case, below the critical concentration an abrupt decrease of the correlation radius takes place. This is because the action of the attractive forces and Brownian motion balance out at a certain cluster size and this size depends on the particle concentration.

As the concentration is increased, the fractal dimension of the inhomogeneities increases and tends to the Euclidean dimension of space (Fig. 4a), which is characteristic for cluster–cluster aggregation models.[6,13] In polymer solutions the fractal dimension of the inhomogeneities remains constant as the concentration is increased.[12] Reference 5 does not present any data on the variation of the fractal dimension of the inhomogeneities, which is understandable in view of the small range of concentrations at which the measurements were made.

To summarize, by means of computer simulation we have established that in emulsions dynamic inhomogeneities with fractal structure can indeed exist. The size of these inhomogeneities is largest near the critical concentration of the

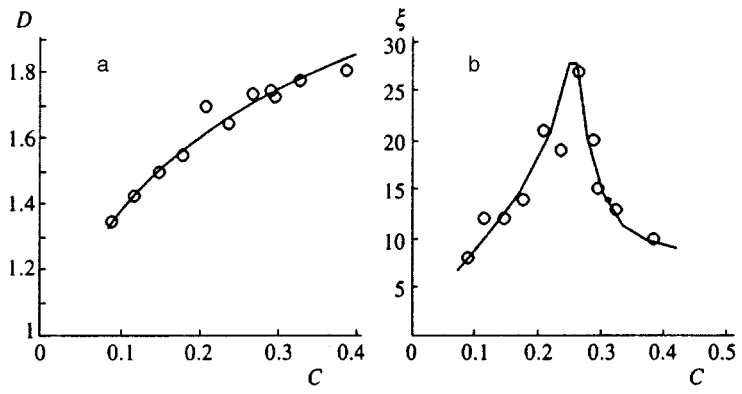

FIG. 4. Fractal dimension of the inhomogeneities D (a) and correlation radius ξ (b) plotted as functions of the particle concentration at $E=1.3kT$.

percolation transition and falls off abruptly both with increase and with decrease of the particle concentration.

The fractal structure of the system owes its existence not to the percolation transition (at the percolation transition only a connecting cluster possesses fractal dimensionality), but to attraction between the particles. When the interaction between the particles is decreased, the fractal structure of the system disappears. The interaction energy at which fractal inhomogeneities exist corresponds in order of magnitude to the interaction energy in real emulsions. The magnitude of the fractal dimension and its growth with increase of the concentration agree with the results obtained in cluster–cluster aggregation models.

Although it would be difficult to compare the results of this simulation with real emulsions due to the fact that we have used here a two-dimensional model and the real interaction forces between the particles of the emulsion are unknown, these results coincide qualitatively with the results obtained in Ref. 5 for a micro-emulsion of silicate glass in a vitrophobic liquid melt by the method of rapid freezing.

ACKNOWLEDGMENTS

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project No. 95-02-06084).

[1] B. M. Smirnov, Usp. Fiz. Nauk **163**, 51 (1993) [Phys. Usp. **36**, (1993)].
[2] V. V. Zosimov and L. M. Lyamshev, Usp. Fiz. Nauk **165**, 361 (1995) [Phys. Usp. **38**, (1995)].
[3] P. Argyrakis, G. Duportail, and P. Lianos, J. Chem. Phys. **95**, 3808 (1991).
[4] G. Duportail, J.-C. Brochon, and P. Lianos, J. Phys. Chem. **96**, 1460 (1992).
[5] M. I. Ozhovan, Zh. Éksp. Teor. Fiz. **104**, 4021 (1993) [JETP **77**, 939 (1993)].
[6] B. M. Smirnov, Usp. Fiz. Nauk **149**, 177 (1986) [Sov. Phys. Usp. **29**, 481 (1986)].
[7] P. Meakin, J. Chem. Phys. **83**, 3645 (1985).
[8] B. Ya. Simkin and I. I. Sheĭkht, *Quantum-chemical and Statistical Theory of Solutions. A Computational Approach* (Ellis Harwood, London, 1995).
[9] R. V. Arbatsumyan, I. Mekke, and D. Shtoyan, *Introduction to Stochastic Geometry* [in Russian] (Nauka, Moscow, 1989).
[10] A. V. Dumanskiĭ, *Colloid Science* [in Russian] (Goskhimizdat, Moscow–Leningrad, 1948).
[11] G. C. Pimentel and A. L. McClellan, *The Hydrogen Bond* (Freeman, San Francisco, 1960).
[12] P.-G. de Gennes, *Scaling Concepts in Polymer Physics* (Cornell Univ. Press, Ithaca, 1979).
[13] H. F. van Gareren, W. H. Dokter, T. P. Beelen *et al.*, J. Chem. Phys. **102**, 480 (1994).

Translated by Paul F. Schippnick

Acoustics of a relativistic superfluid

Yu. V. Vlasov

L.D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, 117334 Moscow, Russia
(Submitted 2 August 1996)
Zh. Éksp. Teor. Fiz. **111**, 1320–1327 (April 1997)

The velocity of propagation of first and second sound at arbitrary temperatures is calculated here for the first time according to the currently accepted covariant theory of superfluidity. Low-intensity shock waves are considered. In the low-temperature limit the results agree with already known expressions obtained by another method. © *1997 American Institute of Physics.* [S1063-7761(97)01504-7]

1. INTRODUCTION

It so happened historically in the covariant theory of superfluidity[1–4] that the main attention was given to fundamental problems while applied problems remained in the shadow. This was particularly the case with the study of the propagation of sound in a relativistic superfluid, which was only begun quite recently. The speeds of first, second, and fourth sound were first calculated by Vil'chinskiĭ,[5] using a simple, but as is well known,[1,2,4,6] invalid model (of the same type as the Israel theory), in which the superfluid and normal components were treated as ideal fluids. Therefore, the results of Ref. 5 in fact pertain not to a superfluid, but to noninteracting gases. In point of fact[1,2] the interaction of the two components leads to the result that the ordinary hydrodynamics of an ideal fluid is no longer applicable, and the more elaborate formalism developed by Lebedev, Khalatnikov, and Carter is required. Results serving as confirmation were obtained by Carter and Langlois[7] for the phonon equation of state. However, the speed of second sound at arbitrary temperature, when the spectrum of elementary excitations deviates from the phonon spectrum, has not yet been reliably calculated. We will attempt to fill this gap below by using the recently derived theory of shock waves in a superfluid.[8] Although there is no need to rederive the results of Ref. 8 here, we will, as an introduction to the subject, present some basic results which will help us to derive, for example, the intermediate formulas previously omitted as unimportant.

In this paper we use the system of units customary in relativity theory with $\hbar=c=1$, in which the metric tensor is $g^{\mu\nu}=\mathrm{diag}\{-1,1,1,1\}$.

2. A RELATIVISTIC SUPERFLUID

In the covariant theory of superfluidity[1–4] the conservation laws of the particle number current

$$\nabla_\rho n^\rho=0 \tag{1}$$

and the entropy current

$$\nabla_\rho s^\rho=0 \tag{2}$$

are written in combination with the conservation law of the energy–momentum (pseudo-) tensor

$$\nabla_\rho T^\rho_\nu=0, \tag{3}$$

where

$$T^\rho_\nu=n^\rho\mu_\nu+s^\rho\Theta_\nu+\Psi g^\rho_\nu, \tag{4}$$

and the one-forms conjugate with the corresponding one-vectors are called the chemical potential covector μ_ρ and temperature covector Θ_ρ. The pressure function Ψ related to the Lagrangian L by the Legendre transformation

$$\Psi=L-s^\rho\Theta_\rho, \tag{5}$$

depends on three invariants, specifically the chemical potential in the superfluid reference system (this is the terminology of Carter, Lebedev, and Khalatnikov),

$$\mu^2=-\mu_\rho\mu^\rho, \tag{6}$$

the effective temperature

$$\Theta^2=-\Theta_\rho\Theta^\rho \tag{7}$$

and the product

$$z^2=-\Theta^\rho\mu_\rho. \tag{8}$$

Thus, there exist four reference systems which are associated with the dynamical variables n^ρ, μ_ρ, s^ρ, and Θ_ρ. If we were working with an ideal two-component fluid, we would have two reference systems. But in the case of a relativistic superfluid, this is anything but the case: e.g., Eckart's reference system does not coincide with the superfluid reference system since the vectors n^λ and μ^λ are not collinear. As a matter of fact, only two of these four variables are independent, and the two others are expressed in terms of them by the formulas

$$n^\rho=F\mu^\rho+H\Theta^\rho, \quad s^\rho=H\mu^\rho+G\Theta^\rho, \tag{9}$$

where the coefficients F, G, and H are calculated with the help of the pressure function:

$$F=\frac{1}{\mu}\frac{\partial\Psi}{\partial\mu}, \quad G=\frac{1}{\Theta}\frac{\partial\Psi}{\partial\Theta}, \quad H=\frac{1}{2z}\frac{\partial\Psi}{\partial z}, \tag{10}$$

and are components of the inverse metric tensor in the space of dynamical variables.[2] Theories of the Israel type correspond to a conformally planar ''metric,'' and may be incorrectly viewed as equivalent to the current theory at low temperatures, for then $H\to 0$. Below, in the instance of acoustic waves we will see just how wrong such a viewpoint is.

In place of the dependence on the quantity z we may introduce a dependence on the relative velocity w between the superfluid component and normal component by means of the formula[7,8]

$$z^2=\frac{H}{G}\,\mu^2w^2+\sqrt{\left(\frac{H}{G}\,\mu^2w^2\right)^2+\Theta^2\mu^2}. \tag{11}$$

3. SHOCK WAVES

The general theory of discontinuities in a relativistic superfluid was considered in Ref. 8. To treat the propagation of acoustic and weak shock waves, we will proceed from the general formalism. Let us consider a planar shock wave propagating along the x^1 axis, with unit normal $\lambda^\rho=(0,1,0,0)$ to the hypersurface of the discontinuity. Here it is convenient to represent the vectors and covectors in the form

$$n^\rho=n(\cosh\,\varphi,\sinh\,\varphi,0,0),\quad \mu_\rho=\mu(-\cosh\,\psi,\sinh\,\psi,0,0),$$
$$s^\rho=s(\cosh\,\alpha,\sinh\,\alpha,0,0),\quad \Theta_\rho=\Theta(-\cosh\,\beta,\sinh\,\beta,0,0). \tag{12}$$

We assume that the medium in front of the discontinuity surface is at rest, and we write the equations in the reference system accompanying the shock wave. Thus, ahead of the front the relative velocity will be equal to zero, and

$$\alpha_-=\beta_-=\varphi_-=\psi_-. \tag{13}$$

The latter expression defines the velocity of the medium $v=\tanh\,\psi_-$ ahead of the shock wave front; obviously, this velocity is equal to the velocity of propagation of the discontinuity in the reference system in which the medium ahead of the front is at rest.

Conservation laws (1)–(3) lead to the following relations on the discontinuity (the brackets denote the difference, the subscripts + and − denote the states in front of and behind the discontinuity:[8]

$$[n^\rho\lambda_\rho]=0,\quad [s^\rho\lambda_\rho]=0,\quad [T^\rho_\nu\lambda_\rho]=0. \tag{14}$$

Hence we have

$$n^1_+=n^1_-,\quad s^1_+=s^1_-, \tag{15}$$

i.e., taking Eqs. (12) and (13) into account

$$n_+\sinh\,\varphi_+=n_-\sinh\,\varphi_-,\quad s_+\sinh\,\alpha_+=s_-\sinh\,\alpha_-. \tag{16}$$

The last of relations (14) defines two forms of strong discontinuities,[8] specifically first-sound discontinuities, which are characterized by the condition

$$\mu^0_+=\mu^0_-,$$

and second-sound discontinuities, which satisfy the equality

$$\Theta^0_+=\Theta^0_-.$$

Equations (12) and (16) lead to the relations

$$\mu_+\,\cosh\,\psi_+=\mu_-\,\cosh\,\psi_-,$$
$$\Theta_+\,\cosh\,\beta_+=\Theta_-\,\cosh\,\beta_-. \tag{17}$$

The latter relations together with relations (16) allow one, with the help of formulas (9), to obtain expressions for v in terms of the thermodynamic quantities for the two types of discontinuities:

$$n_-v=F_+\sqrt{\mu_-^2-\mu_+^2+v^2\mu_+^2}+H_+\sqrt{\Theta_-^2-\Theta_+^2+v^2\Theta_+^2}, \tag{18}$$

$$s_-v=G_+\sqrt{\Theta_-^2-\Theta_+^2+v^2\Theta_+^2}+H_+\sqrt{\mu_-^2-\mu_+^2+v^2\mu_+^2}. \tag{19}$$

The solutions of Eqs. (18) and (19) respectively define discontinuities analogous to ordinary shock waves and temperature discontinuities in superfluid helium.[9] It should also be noted that the quantities μ and Θ grow[8] (this is valid at least for the phonon equation of state).

4. FIRST AND SECOND SOUND

In order to obtain the velocity of propagation of first and second sound from Eqs. (18) and (19), we replace the expressions for the parameters behind the discontinuity front by the finite difference $\mu_+-\mu_-=\Delta\mu$ and drop the subscript $-$. We obtain the system of equations

$$F(v^2-1)\Delta\mu+\mu v^2\,\frac{\partial F}{\partial\mu^2}\,\Delta\mu^2+\mu v^2\,\frac{\partial F}{\partial\Theta^2}\,\Delta\Theta^2$$
$$+\mu v^2\,\frac{\partial F}{\partial z^2}\,\Delta z^2+H(v^2-1)\Delta\Theta+\Theta v^2\,\frac{\partial H}{\partial\mu^2}\,\Delta\mu^2$$
$$+\Theta v^2\,\frac{\partial H}{\partial\Theta^2}\,\Delta\Theta^2+\Theta v^2\,\frac{\partial H}{\partial z^2}\,\Delta z^2=0, \tag{20}$$

$$G(v^2-1)\Delta\Theta+\Theta v^2\,\frac{\partial G}{\partial\mu}\,\Delta\mu+\Theta v^2\,\frac{\partial G}{\partial\Theta^2}\,\Delta\Theta^2$$
$$+\Theta v^2\,\frac{\partial G}{\partial z^2}\,\Delta z^2+H(v^2-1)\Delta\mu+\mu v^2\,\frac{\partial H}{\partial\mu}\,\Delta\mu$$
$$+\mu v^2\,\frac{\partial H}{\partial\Theta^2}\,\Delta\Theta^2+\mu v^2\,\frac{\partial H}{\partial z^2}\,\Delta z^2=0. \tag{21}$$

For the quantity Δz^2, with the help of relation (11) we write down the expansion

$$\Delta z^2=\Theta\Delta\mu+\mu\Delta\Theta+(H/G)\mu\Theta w^2, \tag{22}$$

substitution of which into Eqs. (20) and (21) in the linear approximation leads to the biquadratic equation

$$\left\{v^2-F\left[F+\mu\,\frac{\partial F}{\partial\mu}+\Theta\,\frac{\partial H}{\partial\mu}+\Theta\left(\mu\,\frac{\partial F}{\partial z^2}+\Theta\,\frac{\partial H}{\partial z^2}\right)\right]^{-1}\right\}$$
$$\times\left\{v^2-G\left[G+\Theta\,\frac{\partial G}{\partial\Theta}+\mu\,\frac{\partial H}{\partial\Theta}+\mu\left(\Theta\,\frac{\partial G}{\partial z^2}\right.\right.\right.$$
$$\left.\left.\left.+\mu\,\frac{\partial H}{\partial z^2}\right)\right]^{-1}\right\}-v^2\left\{H\left[H+\mu\,\frac{\partial F}{\partial\Theta}+\Theta\,\frac{\partial H}{\partial\Theta}+\mu\left(\mu\,\frac{\partial F}{\partial z^2}\right.\right.\right.$$
$$\left.\left.\left.+\Theta\,\frac{\partial H}{\partial z^2}\right)\right]^{-1}\right\}\left\{H\left[H+\mu\,\frac{\partial H}{\partial\mu}+\Theta\,\frac{\partial G}{\partial\mu}+\Theta\left(\Theta\,\frac{\partial G}{\partial z^2}\right.\right.\right.$$
$$\left.\left.\left.+\mu\,\frac{\partial H}{\partial z^2}\right)\right]^{-1}\right\}=0 \tag{23}$$

with solutions defining the velocity of propagation of first and second sound.

This biquadratic equation is solved approximately by splitting it into a pair of quadratic equations, one for first sound,

$$u_{\mathrm{I}}^2 = F\left[F + \mu\,\frac{\partial F}{\partial \mu} + \Theta\,\frac{\partial H}{\partial \mu} + \Theta\left(\mu\,\frac{\partial F}{\partial z^2} + \Theta\,\frac{\partial H}{\partial z^2}\right)\right]^{-1} = F\,\frac{\partial \mu}{\partial n} \tag{24}$$

and one for second sound,

$$u_{\mathrm{II}}^2 = G\left[G + \Theta\,\frac{\partial G}{\partial \Theta} + \mu\,\frac{\partial H}{\partial \Theta} + \mu\left(\Theta\,\frac{\partial G}{\partial z^2} + \mu\,\frac{\partial H}{\partial z^2}\right)\right]^{-1} = G\,\frac{\partial \Theta}{\partial s}. \tag{25}$$

This approximate solution is valid given certain conditions (about this see below), and in the most general case first and second sound are, as one says, ''coupled,'' i.e., Eq. (23) does not split into Eqs. (24) and (25). At the same time, formulas (24) and (25) do not in principle impose any restrictions on the equation of state and thus generalize the particular case considered in Ref. 7.

It is interesting to compare Eqs. (24) and (25) with the analogous nonrelativistic relation.[9] Toward this end, we express the quantity G in terms of the determinant of the Carter tensor:[3,7] $G=B/K$, where B has the dimensions of μ/n (and at low temperature coincides with this quantity). On the other hand, it is well known[7] that the effective ''energy density'' of the normal component ρ_N is linked with the entropy density in the superfluid reference system s_s by the relation $\rho_N=\Phi^2 K s_s^2$, where Φ^2 is called a dilaton[3] and the effective density of the superfluid component ρ_s is expressed in terms of the chemical potential μ_N in the normal reference system as $\rho_s=\Phi^2\mu_N^2$. Therefore in place of formula (7) we may write

$$u_{\mathrm{II}}^2 = B\,\frac{s_s^2}{\mu_N^2}\,\frac{\rho_s}{\rho_N}\,\frac{\partial \Theta}{\partial s},$$

which has some superficial resemblance to the well-known nonrelativistic expression.[9] However, the analogy with the nonrelativistic case cannot be complete for the reason that in the covariant theory[1–4] one cannot introduce conserved normal and superfluid currents. It can be shown that formula (8) goes over to the corresponding nonrelativistic expression:[9] the ''energy density'' of the normal and superfluid components acquires the ordinary meaning of a mass density, the entropy density at nonrelativistic velocities will be the same in any reference system, and the chemical potential tends to a constant value equal to the rest energy of the coupled particles.

To check expressions (24) and (25) in the low-temperature limit, where the elementary excitations are phonons, they may be compared with the corresponding result of Carter and Langlois,[7] obtained by the extremely cumbersome Hadamard method. But in the high-temperature case we no longer have any known result against which to check expression (25): in principle it cannot coincide with the speed of second sound calculated by Vil'chinskiĭ.[5] Thus, we will show that relation (25) gives the result of Carter and Langlois in the low-temperature limit.

At low temperatures the Lagrangian is given by the expression[4]

$$\Lambda\{\mu,z,\Theta\} = \Lambda\{\mu\} - \psi\{\mu,z,\Theta\}, \tag{26}$$

where the second term, corresponding to the contribution of the elementary excitations, is written explicitly as

$$\psi\{\mu,z,\Theta\} = \frac{1}{4}\,c_s\left(\frac{3}{4\underline{\hbar}}\right)^3 (G^{\nu\lambda}\Theta_\nu\Theta_\lambda)^2, \quad \underline{\hbar}=0.99\hbar. \tag{27}$$

Here the projective tensor is

$$G^{\nu\lambda} = g^{\nu\lambda} + \left(1-\frac{1}{c_s^2}\right)\frac{\mu^\nu\mu^\lambda}{\mu^2} \tag{28}$$

and c_s is the speed of first sound. With the help of relations (5), (9), (10), and (26)–(28), by differentiating the Lagrangian (9), we readily calculate the coefficients of interest

$$G = -c_s\left(\frac{3}{4\underline{\hbar}}\right)^3\left[-\Theta^2 + \left(1-\frac{1}{c_s^2}\right)\frac{z^4}{\mu^2}\right],$$
$$H = c_s\left(1-\frac{1}{c_s^2}\right)\left(\frac{3}{4\underline{\hbar}}\right)^3\frac{z^2}{\mu^2}\left[-\Theta^2 + \left(1-\frac{1}{c_s^2}\right)\frac{z^4}{\mu^2}\right]. \tag{29}$$

Next, with the help of relations (9) and (10) we calculate the entropy density

$$s = c_s\left(\frac{3}{4\underline{\hbar}}\right)^3\left[-\Theta^2 + \left(1-\frac{1}{c_s^2}\right)\frac{z^4}{\mu^2}\right]^2 \times \sqrt{\Theta^2 - \left[1-\left(\frac{1}{c_s^2}\right)^2\right]\frac{z^4}{\mu^2}}. \tag{30}$$

Equations (18) and (19) and consequently also (24) and (25) pertain to the case in which the relative velocity between the components ahead of the wavefront is equal to zero. This corresponds to $z^2=\mu\Theta$. Then substituting expressions (29) and (30) into Eq. (25) gives

$$u_{\mathrm{II}}^2 = c_s^2/3, \tag{31}$$

which exactly coincides with the result of Carter and Langlois,[7] even though the calculational technique was completely different. The agreement of the results obtained by two independent methods gives a basis for confidence in the validity of both methods.

Similar calculations using the Israel model of noninteracting components with the Lagrangian (26), (27) leads to the incorrect result $u_{\mathrm{II}}^2=1/3$, whereas expression (25) agrees with the nonrelativistic value of the speed of second sound.[9] Finally, it should be noted that for the phonon excitation spectrum the speed of second sound is given by the same expression as for superfluid helium.

It only remains for us to convince ourselves that the speed of first sound (24) corresponds to the value of c_s determined by Carter. Toward this end, we make use of the

relation $F=\Phi^2(1+A^2/K)$ and the asymptotic behavior of the dilaton Φ^2, the true anomaly A, and the Carter determinant K at low temperature:[4] $\Phi^2=n/\mu+O(\Theta^4)$, $A^2/K=O(\Theta^4)$. It follows that

$$u_{\rm I}^2=\frac{n}{\mu}\frac{\partial\mu}{\partial n}\equiv c_s^2, \tag{32}$$

i.e., Eq. (24) is also found to be in complete agreement with current theory, which, it may be noted, could easily have been foreseen.

Let us turn our attention to the low-temperature asymptotic limit $H/G\sim O(\Theta)$. In the Israel theory we have $H\equiv 0$. But if we neglect (without justification) the coefficient H, we obtain the same incorrect result $u_{\rm II}^2=1/3$. When using the hydrodynamics of an ideal fluid in the description of relativistic superfluid systems it should be borne in mind what such a simplified theory leads to.

5. WEAK DISCONTINUITIES

The logical conclusion of our calculations is a study of weak discontinuities. In other words, we will try to ascertain whether the discontinuities appear in front of the sound wave or behind it. Ordinary shock waves (always) propagate faster than sound waves. But in superfluid helium temperature discontinuities can also arise on the trailing edge of a wave of second sound.[9] Let us see how the matter stands in a relativistic superfluid.

In Eqs. (18) and (19) we must keep not only linear terms, but also second order terms, proceeding from the approximate expressions

$$\begin{aligned} nv=&(F+\Delta F)[\{2\mu\Delta\mu+(\Delta\mu)^2\}(v^2-1)+v^2\mu^2]^{1/2}\\ &+(H+\Delta H)[\{2\Theta\Delta\Theta+(\Delta\Theta)^2\}(v^2-1)\\ &+v^2\Theta^2]^{1/2},\\ sv=&(H+\Delta H)[\{2\mu\Delta\mu+(\Delta\mu)^2\}(v^2-1)+v^2\mu^2]^{1/2}\\ &+(G+\Delta G)[\{2\Theta\Delta\Theta+(\Delta\Theta)^2\}(v^2-1)+v^2\Theta^2]^{1/2} \end{aligned}$$

and the more accurate [than (22)] expansion

$$\begin{aligned} \Delta z^2=&\Delta\Theta\mu+\Theta\Delta\mu+\frac{1}{2}\left(\Delta\Theta^2\frac{\mu}{\Theta}+2\Delta\mu\Delta\Theta\right.\\ &\left.+\frac{\Theta}{\mu}\Delta\mu^2\right)+\frac{1}{2}\left(\frac{H}{G}\mu^2w^2\right)^2\frac{1}{\Theta\mu}+\frac{H}{G}\mu^2w^2. \end{aligned}$$

Skipping intermediate steps, we present the final results for the case of the phonon equation of state:

$$v_{\rm I}^2=u_{\rm I}^2+\left\{2(1-u_1^2)-u_{\rm I}^4\frac{\mu^2}{n}\frac{\partial^2 n}{\partial\mu^2}\right\}\frac{\Delta\mu}{\mu}, \tag{33}$$

$$v_{\rm II}^2=u_{\rm II}^2+\frac{3-16c_s-9c_s^2}{6}\frac{\Delta\Theta}{\Theta}. \tag{34}$$

Equation (33) shows that if the speed of first sound u_1 is small, the corresponding discontinuity always propagates ahead of the sound wave. We may obtain a more explicit understanding of formulas (33) and (34) for specific instances of equations of state. Specifically, for the polytropic equation of state we have[10] $u_1^2=(1-m/\mu)/\gamma$, where m is the rest mass of the paired particles and $\gamma>1$ is the polytrope index. Taking (32) into account, we may convince ourselves that $v>u_1$.

It follows from (34) that if the speed of first sound c_s is small, temperature discontinuities are formed on the leading edge of the wave of second sound, which corresponds to the results for superfluid helium in the low-temperature range.[9] However, for $c_s\geqslant 0.97$ the temperature discontinuities fall behind the wave of second sound, this being a specific property of a relativistic superfluid.

Finally, let us turn our attention to the almost obvious fact that substituting the speed of first sound (24) into system (18) expanded out to quadratic terms in the jumps of the various quantities at the discontinuity shows that the quantities $\Delta\Theta$ and w^2 should be of a higher than first order in $\Delta\mu$. And in a wave of second sound the jump $\Delta\mu$ and the quantity w^2 are quantities of higher than first order relative to $\Delta\Theta$.

6. CONCLUSIONS

To summarize, in the foregoing we have calculated the speeds of first and second sound in a relativistic superfluid at arbitrary temperature and have also discussed where the discontinuities arise in waves of first and second sound. The main result consists in Eq. (23) and its consequences for first (24) and second sound (25). Equation (23), in contrast to the Israel model and the result of Ref. 5, does not in general split into Eqs. (24) and (25). The speed of first sound (24) at low temperature is given by the well-known expression (32) which is not in any doubt. The expression for the speed of second sound (25), valid at arbitrary temperatures, in the particular case of the low-temperature limit gives the result (31) found previously by Carter and Langlois.[7] At the same time, the equation obtained here (25) pertains specifically to a relativistic superfluid and not to the Israel model, by means of which the speed of second sound was previously calculated. Thus, the expressions for the speeds of first and second sound obtained here generalize well-known results which are either particular cases of the low-temperature limit[7] or pertain to the approximate model,[5] which cannot be applied to a description of a relativistic superfluid. We were compelled toward this latter conclusion by the incorrect expression for the speed of second sound found in the case of the phonon equation of state.

It is also important to mention that Eq. (23) is universal in the sense that it applies to any two-component system, e.g., in the hydrodynamical description of nuclear matter.

As for the calculational technique, we employed what is unquestionably the most universal method possible, since it was based on the global formulas (18) and (19) for shock waves, which in the limit of infinitesimally small amplitude lead to the velocities of propagation of first and second sound. This approach allowed us to also consider weak discontinuities [Eqs. (33) and (34)].

Finally, the present formalism can be constructed in terms of temperature and pressure as is done in nonrelativistic theory,[9] but this would only complicate the calculations and not increase the physical clarity of the exposition. On the

contrary, operating with the temperature and chemical potential allowed us to write down all the equations, so to speak, in symmetrical form, i.e., to obtain Eqs. (20) and (21). It may be noted in passing that such ''symmetry'' falls away only when we turn to specific equations of state, which can be clearly seen from the difference between formulas (31) and (32) or formulas (33) and (34).

ACKNOWLEDGMENTS

I am grateful to I. M. Khalatnikov for conversations which stimulated me toward a generalization of the nonrelativistic results of Ref. 9 by means of the relativistic theory expounded in Refs. 1 and 2.

[1] V. V. Lebedev and I. M. Khalatnikov, Zh. Éksp. Teor. Fiz. **82**, 1601 (1982) [Sov. Phys. JETP **55**, (1982)].
[2] B. Carter and I. M. Khalatnikov, Phys. Rev. D **45**, 4536 (1992).
[3] B. Carter and I. M. Khalatnikov, Class. Quantum Grav. **11**, 2013 (1994).
[4] B. Carter and I. M. Khalatnikov, Rev. Math. Phys. **6**, 277 (1994).
[5] S. I. Vil'chinskiĭ, Zh. Éksp. Teor. Fiz. **106**, 1430 (1994) [JETP **79**, 772 (1994)].
[6] T. S. Olson, Phys. Lett. A **149**, 71 (1990).
[7] B. Carter and D. Langlois, Phys. Rev. D **51**, 5855 (1995).
[8] Yu. V. Vlasov, submitted to Phys. Rev. D.
[9] I. M. Khalatnikov, *Introduction to the Theory of Superfluidity* (Addison, Redwood, 1989), p. 75; I. M. Khlalatnikov, Zh. Éksp. Teor. Fiz. **23**, 253 (1952).
[10] B. Carter and D. Langlois, Phys. Rev. D **52**, 4640 (1995).

Translated by Paul F. Schippnick

Molecular-dynamics simulation of evaporation of a liquid

V. V. Zhakhovskiĭ*) and S. I. Anisimov

L.D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 9 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1328–1346 (April 1997)

The molecular-dynamics method is used to investigate high-temperature evaporation of a simple liquid. The interaction of the atoms is described by a Lenard–Jones 6–12 potential. The simulation shows that fluctuations of the binding energy in the surface layer play an important role in evaporation, thanks to which a significant contribution to the evaporated flux comes from atoms whose kinetic energy is of the same order of magnitude as the mean thermal energy. Such a mechanism of evaporation differs substantially from the traditional one [Ya. I. Frenkel', *Kinetic Theory of Liquids* (Clarendon Press, Oxford, 1946)] based on the assumption that only those particles evaporate that have energies of the order of the binding energy, i.e., much larger than the mean thermal energy. The structure of the transitional layer between the bulk gas and liquid phases is investigated. Potential energy fluctuations and pairwise correlation functions in the bulk phases and transitional layer are calculated. The velocity distribution function of the atoms for evaporation into vacuum is found. © *1997 American Institute of Physics.* [S1063-7761(97)01604-1]

1. INTRODUCTION

The problem of evaporation of condensed matter is one of the classical problems of physics. It is of fundamental interest and is important for many applications in scientific research and engineering. The intense study which this problem has received in recent years has been facilitated by the development of efficient methods of producing high energy densities based on the use of laser radiation and high-power particle beams. The extensive experimental material obtained in this way is usually interpreted on the basis of greatly simplified theoretical models (see, for example, Refs. 1–3). In this connection, processes taking place in the gas phase, where kinetic effects play an important role,[4–6] have been studied in a more fundamental way than processes in the condensed phase and in the interphase transitional layer. Instead of analyzing the latter, most studies make do with the simple single-particle model, which treats evaporation as the escape of the fastest atoms from a potential well whose depth is equal to the mean binding energy U_0. In order to calculate the flux of evaporated material, the velocity component v_z normal to the phase boundary is integrated over the tail of the Maxwellian distribution with $v_z > \sqrt{2U_0/m}$. Such a model is obviously an oversimplification. In reality, for atoms located in the surface layer and making the main contribution to the evaporated flux, the binding energy U is not a fixed quantity, but depends on the immediate environment of the given atom. It is equal to U_0 only in order of magnitude. The following circumstance is also important. A surface atom with binding energy U before the transition to the gas phase completes on average $\exp(U/kT)$ oscillations, i.e., it remains in the bound state for a time of the order of $\tau \exp(U/kT)$, where τ is equal in order of magnitude to the inverse Debye frequency. At temperatures $kT \ll U$ this time is much greater than the characteristic time of variation of the binding energy of the atom U, i.e., than the time, of order τ, required for a noticeable rearrangement to take place in the immediate environment of the surface atom. Thus, evaporation must be considered as a collective process. As a consequence of fluctuations of the binding energy, atoms or groups of atoms are formed in the surface layer such that the energy per particle required to tear any of them free is significantly less than U_0 and their lifetime in the bound state in the surface layer is of order τ. These atoms with kinetic energy on the order of the mean thermal energy make a substantial contribution to the evaporated flux. Here we may make the following remark. If we assume that the main contribution to the evaporated flux comes from particles belonging to the tail of the Maxwellian distribution, then it may be expected that evaporation will lead to a depletion of fast particles from the surface layer of the condensed phase. In this case the complicated question arises of the mechanism by which fast particles fill the tail. This question has been discussed a number of times, but it has not obtained a satisfactory answer. The fluctuation mechanism of evaporation discussed here is free from this difficulty.

At temperatures less than critical, the boundary between the liquid and gas phases is well defined: the density varies in the direction of its maximum gradient over a length on the order of the interatomic distance in the liquid. This boundary has a complicated shape which varies with time. Long-wavelength perturbations of the boundary include the well-known capillary waves with dispersion law $\omega^2 = \gamma k^3/\rho$, where γ is the surface tension and ρ is the density of the liquid.[12] Curving of the phase boundary as the wave propagates leads to variation of the binding energy of the atoms by a magnitude of the order of $\gamma k n^{-1/3}$, where n is the particle number density in the liquid. This quantity becomes comparable to the mean binding energy U_0 at wavelengths of the order of the interatomic distance. Numerical calculations[11] show that at sufficiently high temperatures perturbations of the interphase boundary with wavelength of the order of

 1063-7761/97/040734-12$10.00

$(5-8)d$, where d is the mean interatomic distance, have a considerable amplitude. Perturbations of the boundary increase the surface-averaged thickness of the transitional layer, which is also much greater than the thickness of the boundary. The excitation of short-wavelength surface modes inside the transitional layer leads to the formation on the phase boundary of segments with high local curvature which make a significant contribution to the evaporated flux. Note that at wavelengths severalfold greater than the interatomic distances, the motions of some neighboring atoms are correlated; therefore, along with the evaporation of individual atoms one may also expect large clusters to break free from the surface. Such clusters have been observed in molecular-dynamics calculations.[11]

The real structure of the surface layer of an evaporating liquid should have a substantial effect on the interaction of the atoms of the vapor with the surface of the liquid. In studies of surface phenomena at the interface between condensed matter and a gas, calculations of such quantities as accommodation coefficients and reflection coefficients of the atoms have traditionally occupied an important place. Usually such calculations have a model character and are typically carried out for a flat stationary interface. In light of the complicated, time-dependent structure of the transitional layer between a liquid and a gas, the applicability of the results of such studies to a real liquid–vapor boundary raises doubts. Processes in the transitional layer determine the character of the velocity distribution of the evaporated atoms, and discontinuities in the gas-dynamic quantities in the Knudsen layer near the phase boundary depend, in turn, on this distribution. To solve the Boltzmann equation for the Knudsen layer, the distribution function of the atoms emitted by the surface is usually assumed to be Maxwellian (with temperature equal to the temperature of the liquid and particle number density equal to the density of the saturated vapor at this temperature) for $v_z<0$ (the z axis is directed from the liquid to the gas).[4–10] This boundary condition was never rigorously derived. It is probably correct in the case of a planar phase boundary, but is not obvious in general. A molecular-dynamics simulation would give some idea of how close it is to reality for a real boundary.

In the present study we have simulated the evaporation of a liquid by the molecular-dynamics method and have undertaken an effort to consider the above-mentioned questions of the mechanism of intense evaporation, the structure of the transitional layer between a liquid and its vapor, and the velocity distribution of the evaporated atoms. The paper is organized as follows. Section 2 describes the model and formulates some methodological questions. Section 3 addresses simulation of equilibrium between the liquid and vapor in a closed space, and Sec. 4—non-equilibrium evaporation into vacuum.

2. THE MODEL

We consider a system of N particles located in a parallelepiped with dimensions $L\times L\times 2L$, called from here on the molecular-dynamics cell. The interaction between the ith and jth particles is described by the Lenard–Jones potential

$$u(r_{ij})=4\varepsilon[(\sigma/r_{ij})^{12}-(\sigma/r_{ij})^{6}], \tag{1}$$

where r_{ij} is the distance between the particles, and ε and σ are the potential parameters. In order to separate the system into a liquid phase and a gas phase along the z axis, we introduce an additional potential acting on the system from the direction of one of the faces of the parallelepiped, the floor of the molecular-dynamics cell $z=-L$:

$$u_b(z_i)=4\varepsilon_b[(\sigma/(L+z_i))^{12}-(\sigma/(L+z_i))^{6}]. \tag{2}$$

To simulate evaporation into vacuum and into a closed space, we introduce a repulsive potential at the opposite face of the cell $z=L$

$$u_c(z_i)=4\varepsilon_c[\sigma/(L-z_i)]^{12}. \tag{3}$$

This potential plays the role of the lid of the molecular-dynamics cell. For a closed lid we have $\varepsilon_c=\varepsilon$, and for an open lid $\varepsilon_c=0$. Periodicity conditions with period L are imposed on the system in the xy plane. In this way, for each particle with radius vector $\mathbf{r}_i$ located in the main cell an infinite series of images with radius vectors $\mathbf{r}_i+\mathbf{n}L$ is introduced, where $\mathbf{n}$ is an integer-valued vector in the xy plane: $\mathbf{n}=\{l,m,0\}$. The distance between the ith and jth particles is defined here by the rule

$$r_{ij}=\min(|\mathbf{r}_i-\mathbf{r}_j+\mathbf{n}L|),$$

where the minimum is taken over all vectors $\mathbf{n}$.

The molecular-dynamics simulation of the system is based on the solution of the equations of motion of the particles making up the system. As the equations of motion it is convenient to take Newton's equations, which do not contain first derivatives in time

$$m\ddot{\mathbf{r}}_i=\mathbf{f}_i(\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_N), \quad i=1,2,\ldots,N, \tag{4}$$

where the forces $\mathbf{f}_i$ are determined by differentiating the potentials (1)–(3).

Special numerical methods have been developed to solve such equations (mainly in connection with the problem of celestial mechanics).[13] In the present study, to integrate the equations of motion (4) we applied (apparently for the first time in the practice of molecular-dynamics calculations) the explicit eighth-order Störmer method. This method was chosen in place of the simpler schemes traditionally used in molecular-dynamics simulations[14] because of the necessity of calculating the trajectories of the system for an extended interval of time since the establishment of thermodynamic equilibrium in a two-phase system is much slower than in a one-phase system. The application of a higher-order method makes it possible while preserving the prescribed accuracy to use a several times greater time step and thereby speed up the calculation and reduce the accumulation of errors. Note that this statement is valid for sufficiently dense systems, where the number of particle–neighbors calculated is large and the main fraction of the time ($>90\%$) is spent calculating the interparticle distances and forces. Therefore, an increase in the order of this explicit integration method only insignificantly increases the computer time expended per integration step (less than 10% in our case) and allows us to substantially increase the size of the integration step (by roughly a factor of four). Thus, in the case of a large number

of neighbors (dense systems, long-range potentials) the use of the explicit eighth-order Störmer method speeds up the calculation in comparison with ordinary methods by roughly a factor of $3-4$. A significant increase in the amount of computer memory used is not a real problem at the present time. The calculational formula used to integrate the equations of motion does not depend on the particle velocities and has the form

$$x_{n+1}=2x_n-x_{n-1}+\frac{h^2}{60480}\sum_{j=0}^{7}a_j\ddot{x}_{n-j}. \tag{5}$$

Here n is the number of the time step, h is the length of the time step, and the quantities a_j are given in the following table.

j	0	1	2	3	4	5	6	7
a_j	88324	−121797	245598	−300227	236568	−117051	33190	−4125

To simplify the presentation, formula (5) is written down only for the x component of the vector $\mathbf{r}$. The form of formula (5) is the same for the y and z components. Obviously, more than just this one formula is needed for the molecular-dynamics simulation since it is also required to know the velocities $\dot{\mathbf{r}}_i$ of all the particles at each time step to calculate the physical quantities which depend on the particle velocities. We derived and implemented the following formula:

$$\dot{x}_{n+1}=\frac{1}{2h}(x_{n+1}-x_{n-1})+\frac{h}{3}(\ddot{x}_{n+1}+2\ddot{x}_n). \tag{6}$$

The main difference between formula (6) and the usual finite-difference estimates of the first derivative consists in the use of second derivatives to enhance the accuracy. This approach is natural since it uses values of the accelerations already calculated in order to evaluate formula (5). It should be emphasized that formula (6) is not used in the process of integrating the equations of motion (4) by the method (5). Therefore, the error of estimate (6) affects only values of the velocity-dependent physical quantities such as the kinetic energy calculated at the nth step.

To carry out the calculations, it is convenient to use dimensionless variables. It is customary to take σ as the unit of length, ε as the unit of energy, and the quantity $\tau=\sigma\sqrt{m/48\varepsilon}$ as the unit of time. In what follows, if the dimensions are not explicitly stated, we will use these quantities. For example, the temperature in dimensionless units is expressed as

$$T=\frac{16}{N}\sum_{i=1}^{N}v_i^2=16\langle v^2\rangle.$$

To transform to dimensioned quantities, it is necessary to use the parameters of the Lenard–Jones potential of the specific material involved. For example, for argon we may use $\sigma=3.405\times10^{-8}$ cm, $\varepsilon=1.653\times10^{-14}$ erg$=119.8$ K, therefore the unit of time is $\tau=3.114\times10^{-13}$ s.

To start the integration of Eqs. (4), it is necessary to assign the initial coordinates and velocities of all the particles. Toward this end, the particles are arranged at the nodes of a simple cubic lattice inside a cube of dimensions $L\times L\times L$ resting on the floor of the molecular-dynamics cell. The particle velocities were chosen to be equal in magnitude to $v=\sqrt{T/16}$ and randomly distributed in direction. Note that the results of the calculation do not depend on how the initial conditions are prescribed. As the molecular-dynamics simulations showed, a Maxwellian distribution is established much faster in a two-phase system than equilibrium between the phases is reached. In order to bring the system into an equilibrium state with prescribed temperature, at the initial step of the simulation it is necessary to scale the velocities and damp the oscillations of the center of mass along the z axis. This initial step is followed by a shorter intermediate step in which the system is not subjected to any external forces and enters a stationary state. Here a small change in the temperature takes place. The criterion for the completion of both steps is constancy of the temperature along the z axis and smallness of the fluctuations of the center of mass of the entire system and of the gas and liquid phases separately. Note that the z component of the total momentum of the system is not conserved; therefore fluctuations of the position of the center of mass of the entire system are inevitable and reflect the fluctuations of the particle number density.

The main difficulty of a molecular-dynamics simulation, as is well known, has to do with the fact that to calculate the forces acting on a particle it is necessary at each time step to calculate all the distances between the particles. The number of these distances is equal to $N(N-1)/2$; therefore the number of calculations grows as N^2. The most successful way of avoiding quadratic growth of the calculation time with increase of the number of particles was proposed by Verlet.[15] It consists of defining the matrix of nearest neighbors, which is only used to calculate the interparticle distances. With this goal, a cutoff radius of interaction r_c and horizon radius r_g are introduced. The particles located a distance less than r_g from the given particle are taken to be its nearest neighbors. Only particles from the interaction sphere $r_{ij}<r_c$ interact with the given particle. The particles located between the two spherical surfaces with radii r_c and r_g form a buffer which exchanges particles with the interaction sphere. The matrix of nearest neighbors is replenished every k time steps. So that even the fast particles will not be able to make it past the buffer in the time kh, where h is the time step, Verlet[15] proposed choosing the horizon radius equal to

$$r_g=r_c+sk\langle v\rangle h, \tag{7}$$

where $\langle v\rangle$ is the mean-square velocity of the particles and the parameter $s\approx5.3$. We have introduced two new elements into the classical algorithm. First, the particles j moving out beyond the horizon radius of the ith particle during the time kh were excluded from amongst its nearest neighbors until the matrix of nearest neighbors is updated again. This procedure made it possible to speed up the program by roughly 10% and use formula (8). Second, we did not take r_g to be a

TABLE I. Parameters of closed two-phase systems.

T	0.7250 ± 0.0058	0.7952 ± 0.0063	0.897 ± 0.007	0.991 ± 0.008
n_l	0.8060	0.7722	0.7164	0.6582
n_g	0.004886	0.009655	0.02367	0.04687
u_l	-5.423 ± 0.5796	-5.146 ± 0.6070	-4.713 ± 0.6463	-4.285 ± 0.6845
u_g	-0.05292 ± 0.1364	-0.09212 ± 0.1741	-0.2282 ± 0.2851	-0.4263 ± 0.4013
ϵ_l	1.087 ± 0.8874	1.190 ± 0.9721	1.346 ± 1.099	1.486 ± 1.214
ϵ_g	1.107 ± 0.8991	1.232 ± 1.002	1.329 ± 1.077	1.474 ± 1.200

constant quantity as was done in Ref. 15, but rather assumed it to depend on the velocity of the particle and the structure of its immediate environment. We assumed that $r_g(i)$ increases with growth of the velocity of the particle v_i, decreases as the number of particles N_i inside the cutoff radius of the ith particle increases, and decreases as the time of the next update is approached. These conditions are satisfied by the empirically chosen dependence

$$r_g(i,m)=r_c+(k-m)\langle v\rangle h\left(3+\frac{v_i^2}{\langle v^2\rangle}\right)\frac{4}{1+CN_i/r_c^3}, \quad (8)$$

where $m=0,\ldots,k-1$ is the number of time steps that have passed since the last replenishment, $k-m$ is the number of time steps until the next update, v_i is the speed of the ith particle, and C is an empirical constant. The maximum value of C for which no cases of penetration of the buffer by outlying particles are observed was identified by preliminary calculations. In our calculations we adopted the value $C=0.88$. Using the parameters of our model $N_i\approx 100$ and $r_c=3.2$, for $m=k/2$ and $v_i=\langle v\rangle$ we obtain $s\approx 2.2$. This reduces the number of neighbors (for $k=16$, $\langle v\rangle=1/4$, $h=1/32$) in comparison with formula (7) by a factor of roughly 1.5. On the basis of this estimate we may conclude that defining the horizon radius by formula (8) substantially increases the efficiency of the algorithm in comparison with its original definition.[15] Moreover, as a consequence of its sensitivity to the local value of the particle the number density formula (8) offers definite advantages for simulating two-phase systems. It should be noted that formula (8) is not optimal from the point of view of minimizing the number of nearest neighbors. This question requires special study, and we will not touch on it further in this paper.

3. EQUILIBRIUM LIQUID–GAS SYSTEM

In this section we present results of a molecular-dynamics simulation of an equilibrium two-phase Lenard–Jones system. The system was brought into equilibrium with the prescribed temperature. Its physical characteristics were then calculated by averaging along a phase trajectory on the constant energy surface, i.e., we simulated the microcanonical ensemble. As has already been mentioned, we are using molecular-dynamics units. In all of the cases the number of particles in the main cell N was equal to 8000, and the cell had dimensions $L\times L\times 2L$ with $L=25.1984$. The integration step was $h=0.03125$ and the cutoff radius was $r_c=3.2$. We give for guidance the dimensioned values of the molecular-dynamics units in the case of argon (atomic mass 40): unit velocity $v_0=1093.3$ m/s, unit particle number density 1.6825 g/cm^3. The units of length, time, and energy were given above. Some results of the calculations are summarized in Table I. For the liquid and gas phases (indicated by the subscripts l and g, respectively) the table gives the values of the density n_l and n_g, mean potential energy per particle u_l and u_g, and mean kinetic energy ϵ_l and ϵ_g at four temperatures (the numbers after the $\pm$ sign are the values of the mean-square fluctuations). As can be easily seen, the kinetic energy fluctuations satisfy the relation

$$\langle(\langle\epsilon\rangle^2-\epsilon^2)\rangle=\frac{2}{3}\langle\epsilon^2\rangle,$$

with high accuracy, which indicates that the system is in thermodynamic equilibrium. The small value of the temperature fluctuations ($\approx 0.8\%$), a consequence of the large number of particles, may also be noted. This justifies the adopted formulation of the problem, based on the use of an *NVE* ensemble without artificial procedures for maintaining constant temperature. The time dependence of the ''instantaneous'' temperature, defined by relation (7), is shown in Fig. 1.

In order to determine the dependence of the mean value of a physical quantity f on the coordinate z, the molecular-dynamics cell was subdivided into 512 layers of thickness $\delta z=2L/512$ and the values of f (which depends on the coordinates and velocities of the particles) were summed over all the particles falling within the given layer. The result was then divided by the total number of particles falling within

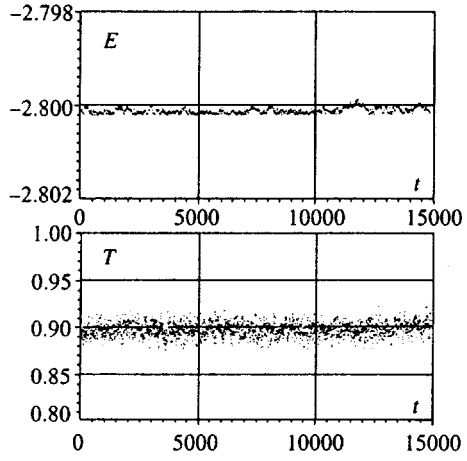

FIG. 1. Time dependence of the total energy and instantaneous temperature in a closed system with $T=0.897$.

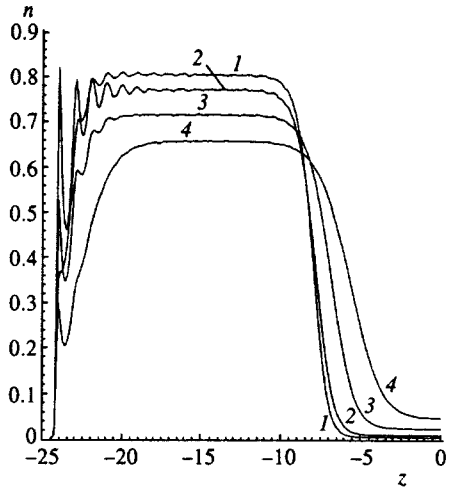

FIG. 2. Spatial profiles of the particle number density in a closed system at different temperatures: *1*—$T=0.725$, *2*—$T=0.795$, *3*—$T=0.897$, *4*—$T=0.991$.

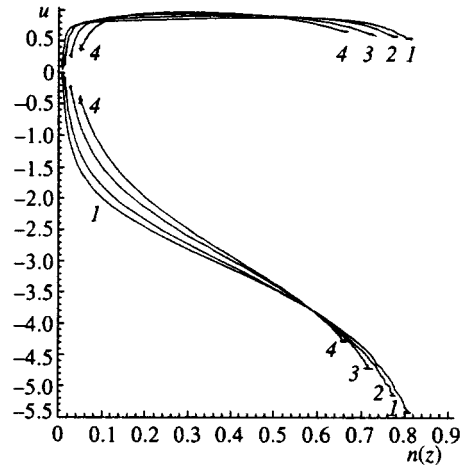

FIG. 4. Mean potential energy per particle $\overline{u}$ and its mean-square fluctuation δu as functions of the particle number density at different temperatures. Notation of the curves is the same as in Fig. 2.

the layer. Profiles of the particle number at different temperatures calculated in this way are shown in Fig. 2. As was noted in the previous Section, near the floor of the cell the system is acted upon by forces with the potential (2). At $T=0.725$ the constant ε_b was set equal to 0.5, and at higher temperatures it was set equal to 1. This potential affects the density profile only near the floor of the cell and has no effect on the bulk properties of the phases. Note that in the transitional layer between the liquid and gas the density profile turns out to be monotonic. Thus, our calculation does not support the hypothesis advanced some time ago by Fisher[16] of the existence of density oscillations in the transitional layer.

Figure 3 plots the mean potential energy per particle and the mean-square fluctuation of this quantity as functions of z. We used the standard recipe in the calculation of these quantities: half of the interaction energy was assigned to each of the pair of interacting particles. The calculation was carried out according to the scheme described above. The marked growth of the potential energy (binding energy) fluctuations in the interphase region draws attention to itself. It is precisely these fluctuations, as was noted in the Introduction, that increase the probability that particles go from the liquid to the gas. Figure 4 plots the mean potential energy and its fluctuations as functions of the particle number density. The curves corresponding to different temperatures differ little in the region of densities corresponding to the transitional layer. Therefore, as a rough estimate it may be assumed that the potential energy in the transitional layer is a function of the local density. More detailed information about the binding energy of the atoms in the liquid, gas, and transitional layer may be obtained by calculating the potential energy distribution function from the molecular-dynamics simulation data. Examples of such calculations for different temperatures are shown in Fig. 5 ($T=0.795$) and Fig. 6 ($T=0.991$). The figures show the unit-normalized potential energy probability distribution functions for the equilibrium system for several different cross sections of the molecular-dynamics cell. The energy was averaged over layers of thickness 0.2, parallel to the xy plane. The z coordinate of each of the layers is indicated in the caption. Curve *1* in both figures corresponds to the homogeneous liquid, and curve *8*—to the gas, and the remaining curves—to the transitional layer. For all of the cross sections the potential energy distribution function is nonzero at negative energies. In the gas phase it exhibits features associated with the formation of diatomic and larger clusters. In the high-temperature case the potential energy distributions in the cross sections close to the ''gas'' end of the transitional layer exhibit similar behavior (see curves *5–7* in Fig. 6). This can be interpreted as an indication of the formation of clusters inside the transitional layer. The mean potential energy per atom in the gas phase is negative (which is also apparent from Figs. 3 and 4). In the transitional layer the distribution broadens noticeably, corresponding to growth of the amplitude of the binding energy fluctuations. The shape of the distribution function in the transitional layer and in the gas phase differs substantially from a Gaussian.

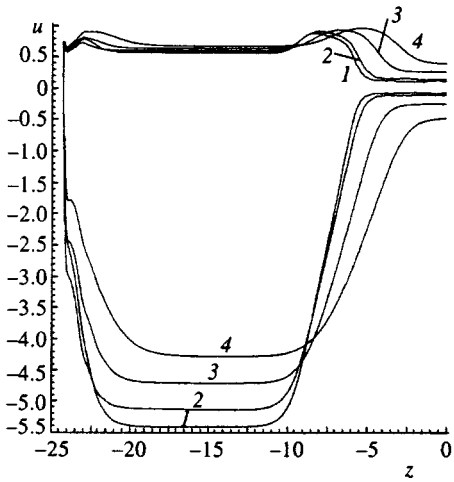

FIG. 3. Spatial profiles of the mean potential energy per particle $\overline{u}$ and its mean-square fluctuation δu in a closed system at different temperatures. The curves are labeled as in Fig. 2.

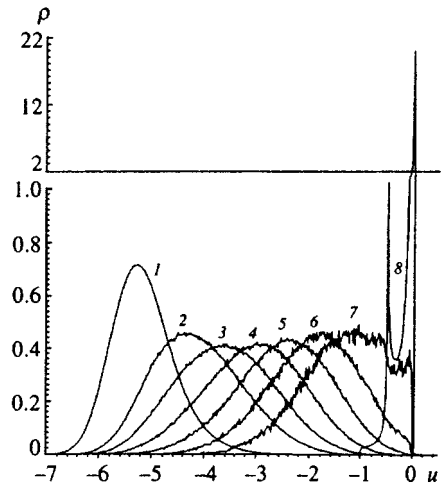

FIG. 5. Normalized distributions of the potential energy $\rho(u)$ for a system in equilibrium with $T=0.795$, calculated for different layers during the transition from liquid (curve *1*) to gas (curve *8*). Coordinates of the layers: *1*—the layer $-17.0<z<-14.0$, *2*—the layer $-9.0<z<-8.8$, *3*—the layer $-8.5<z<-8.3$, *4*—the layer $-8.0<z<-7.8$, *5*—the layer $-7.5<z<-7.3$, *6*—the layer $-7.0<z<-6.8$, *7*—the layer $-6.5<z<-6.3$, *8*—the layer $-1.0<z<18.0$.

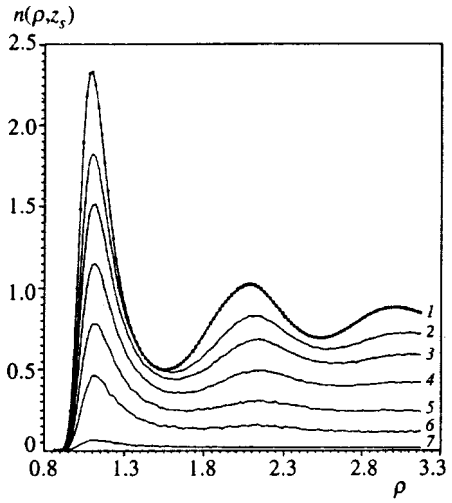

FIG. 7. Pairwise radial distributions in an equilibrium system at $T=0.795$, plotted vs the position of the layer z_s during the transition from liquid (curve *1*) to gas (curve *7*): *1*—the layer $-17.0<z<-14.0$, *2*—the layer $-9.0<z<-8.8$, *3*—the layer $-8.5<z<-8.3$, *4*—the layer $-8.0<z<-7.8$, *5*—the layer $-7.5<z<-7.3$, *6*—the layer $-7.0<z<-6.8$, *7*—the layer $-1.0<z<18.0$.

Here we must note than in our earlier paper (Ref. 11) the value $u/2$ is erroneously used in the figures instead of u.

The two-particle distribution function $n(\mathbf{r}_i,\mathbf{r}_j)$ is an important source of information about the structure of the transition layer. Obviously, in a homogeneous phase it depends only on the interatomic distance $r_{ij}=|\mathbf{r}_i-\mathbf{r}_j|$. In our case, in which the system varies in the z direction, $n(\mathbf{r}_i,\mathbf{r}_j)$ is a function of three variables, which we may choose as the z coordinates z_i and z_j of the two particles and the radial distance $\rho_{ij}=\sqrt{(x_i-x_j)^2+(y_i-y_j)^2}$ between their projections on the xy plane. The two-particle distribution functions in these variables were calculated using the results of the molecular-dynamics simulation. In order to get an idea of how the function $n(\rho_{ij},z_i,z_j)$ varies as one goes from the liquid to the gas phase, we divided the transitional region into a series of layers z_s of thickness 0.2 parallel to the xy plane, and the coordinate z_i was chosen inside one of these layers. Distributions of two types were investigated: radial, depending on the variable ρ_{ij} and calculated for the condition $|z_i-z_j|<0.2$, and axial, depending on the variable z_j-z_i and calculated for the condition $\rho_{ij}<0.2$. Distributions of both types were also constructed for the bulk phases. The distributions were calculated for different temperatures from $T=0.725$ to $T=0.991$. Figures 7 and 8 plot the radial distributions in various layers z_s, calculated for the two extreme temperature values. The coordinates of the layers are indicated in the captions. It can be seen that as one goes from the liquid to the gas phase, the first thing one notices is the destruction of the higher-order coordination spheres. Here the positions of the maxima hardly change at all. This means that the most probable distance between some isolated particle and the particles of the nth coordination sphere tends to remain fixed in an equilibrium two-phase system. Thus, as one goes from the liquid to the gas a tendency is manifested to preserve close-range order.

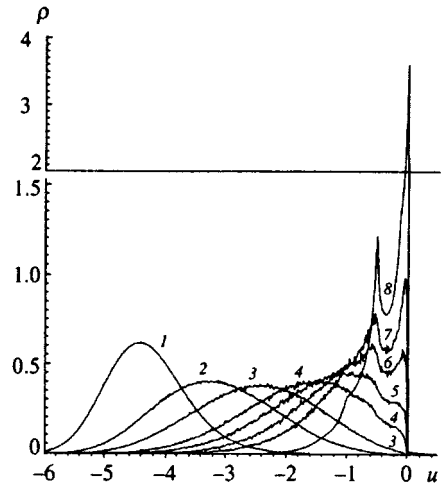

FIG. 6. Normalized distributions of the potential energy $\rho(u)$ for a system in equilibrium with $T=0.991$, calculated for different layers during the transition from liquid (curve *1*) to gas (curve *8*). Coordinates of the layers: *1*—the layer $-15.0<z<-14.0$, *2*—the layer $-6.5<z<-6.3$, *3*—the layer $-5.5<z<-5.3$, *4*—the layer $-4.5<z<-4.3$, *5*—the layer $-4.0<z<-3.8$, *6*—the layer $-3.5<z<-3.2$, *7*—the layer $-3.0<z<-2.8$, *8*—the layer $3.0<z<18.0$.

Examples of axial distributions are shown in Figs. 9 and 10. Although the level of the fluctuations in the axial distributions is higher than in the radial ones due to poor statistics, the main features of the transition from liquid to gas are preserved in this case. The curves numbered *1* in all the figures correspond to the homogeneous liquid phase, for which the radial and axial distributions should obviously coincide. Calculation indeed demonstrates complete agreement between the two distributions. This can be seen from Figs. 9 and 10, in which the square symbols indicate the radial distributions for the liquid phase. On the basis of a study of the pairwise correlation functions in the transitional region be-

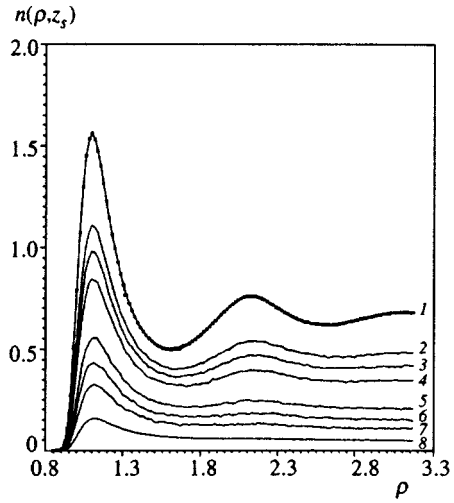

FIG. 8. Pairwise radial distributions in an equilibrium system at $T=0.991$, plotted vs the position of the layer z_s during the transition from liquid (curve *1*) to gas (curve *8*): Coordinates of the layers: *1*—the layer $-15.0<z<-14.0$, *2*—the layer $-6.5<z<-6.3$, *3*—the layer $-6.0<z<-5.8$, *4*—the layer $-5.5<z<-5.3$, *5*—the layer $-4.5<z<-4.3$, *6*—the layer $-4.0<z<-3.8$, *7*—the layer $-3.5<z<-3.2$, *8*—the layer $3.0<z<18.0$.

tween the liquid and the vapor, we may draw the following conclusions about the structure of the liquid surface layer. As a consequence of preservation of short-range order, the local density (defined on a scale of $2-3$ unit lengths) in the interphase region takes only two values corresponding to the densities of the bulk phases—liquid or gas. This means that there exists a well-defined phase boundary between the liquid and the gas, whose thickness is equal to roughly $1-2$ molecular-dynamics units. The amplitude of the fluctuations in the position of this boundary defines the mean density profile along the z axis and the mean thickness of the interphase region, which increases with the temperature and reaches roughly eight molecular-dynamics units at $T\approx 0.9$.

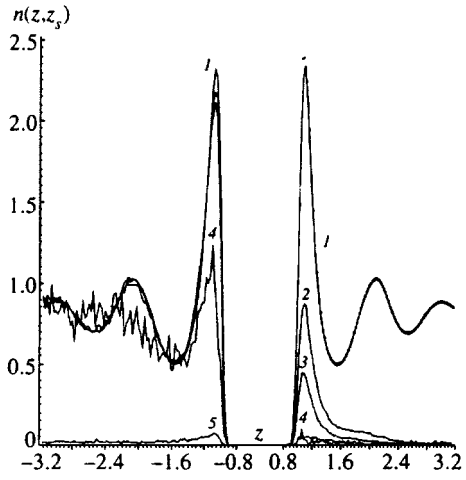

FIG. 9. Pairwise axial distributions in an equilibrium system for $\rho_{ij}<0.2$, plotted vs the position of the layer z_s $(z=z_j-z_i)$ at $T=0.725$. Squares on curve *1* correspond to the radial distribution in the liquid: *1*—the layer $-17.0<z<-14.0$, *2*—the layer $-9.0<z<-8.8$, *3*—the layer $-8.5<z<-8.3$, *4*—the layer $-6.5<z<-6.3$, *5*—the layer $-1.0<z<18.0$.

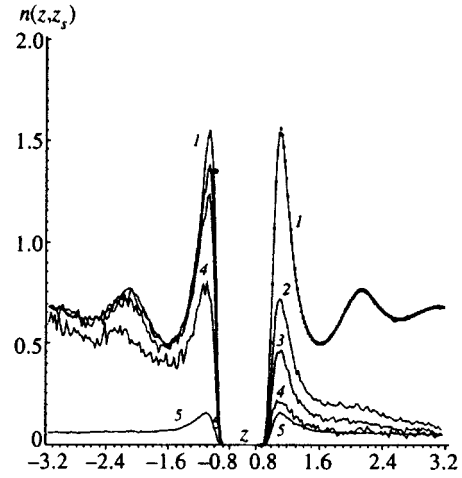

FIG. 10. Pairwise axial distributions in an equilibrium system for $\rho_{ij}<0.2$, plotted vs the position of the layer z_s $(z=z_j-z_i)$ at $T=0.991$. Squares on curve *1* correspond to the radial distribution in the liquid: *1*—the layer $-15.0<z<-14.0$, *2*—the layer $-6.5<z<-6.3$, *3*—the layer $-5.5<z<-5.3$, *4*—the layer $-4.0<z<-3.8$, *5*—the layer $3.0<z<18.0$.

We made an attempt to apply the formalism of temporal correlation functions to study the density fluctuations in the transitional layer. Toward this end, three regions were distinguished in the molecular-dynamics cell, corresponding to the bulk liquid and gas phases and to the transitional layer. For each of these regions we analyzed the density fluctuations with wave vectors $\mathbf{q}_{l,m}=(2\pi/L)(l,m,0)$ $(l=0,1,2,3,$ $m=0,1,2,3)$ lying in the xy plane and calculated the dynamic structure factor. We calculated the spatial Fourier transform of the density operator at the time t for each region:

$$n(\mathbf{q},t)=\frac{1}{N_t}\sum_{j=1}^{N_t}\exp(i\mathbf{q}\cdot\mathbf{r}_j(t)),$$

where N_t is the number of particles observed in the region at the time t and the subscript j numbers the particles. The time series $n(\mathbf{q},t)$ was written to disk. After the termination of the calculation, the fast Fourier transform was used to estimate the spectral density function of the stationary process $n(\mathbf{q},t)$, where this spectral density function is the desired dynamic structure factor:

$$S(q,\omega)=\frac{N}{2\pi}\int dt\langle n^*(\mathbf{q},0)n(\mathbf{q},t)\rangle\exp(-i\omega t), \qquad (9)$$

where $F(q,t)=N\langle n^*(\mathbf{q},0)n(\mathbf{q},t)\rangle$ is the autocorrelation function of the density fluctuations. Using the finite Fourier transform, we may rewrite this expression in the form

$$S(q,\omega)=\lim_{T\to\infty}\frac{N}{2\pi T}\left|\int_0^T dt\; n(\mathbf{q},t)\exp(-i\omega t)\right|^2. \qquad (10)$$

Here N is the mean number of atoms in the region. As a consequence of the statistical invalidity of estimating the spectral density from a single realization of the random process in question,[17] we divided the entire time interval into N_d segments in order to reduce the error of the estimate,

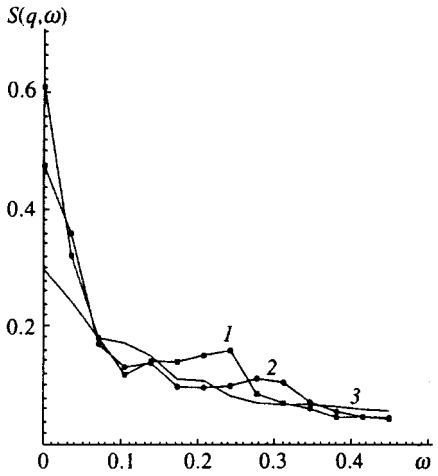

FIG. 11. Dynamic structure factor $S(q,\omega)$ of the liquid phase in equilibrium with the vapor at $T=0.725$ as a function of ω for different values of q ($N_d=81$): *1*—0.2493, *2*—0.3526, *3*—0.7480.

which is proportional to $1/\sqrt{N_d}$. We did this without using spectral windows. Details of the technique can be found in Ref. 18.

An example of the dynamic structure factor calculated according to the described technique is shown in Figs. 11 and 12. At the longest wavelengths and low temperatures (0.725, 0.897) collective excitations are observed whose velocity turns out to be close to the speed of sound in argon ≈ 1 (1000 m/s) near the triple point.

Because of the variability of the mean density $n(z)$ in the transitional layer, it is difficult to interpret $S(q,\omega)$ as the dynamic structure factor of the interphase region. However, if we use the existence of a well-defined boundary between the liquid and the vapor, mentioned above, then it is possible to attempt to link the function $S(q,\omega)$ with the spectrum of the fluctuations of this boundary. A similar analysis was carried out for the static case by Triezenberg and Zwanzig,[19] which examined the effect of density fluctuations in the transitional layer on the fluctuations of the separating Gibbs surface and on the surface tension. The formula which they obtained links the Fourier component of the height of the Gibbs surface with the Fourier components of the fluctuations of the local density $n(\mathbf{q},z)$:

$$z_0(\mathbf{q})=\int dz\; n(\mathbf{q},z)/(n_l-n_g), \tag{11}$$

where the vector $\mathbf{q}$ lies in the xy plane. The formalism developed by Triezenberg and Zwanzig[19] permits one to generalize to the dynamic case. It can be shown that for time-dependent density fluctuations the Fourier component of the height of the separating Gibbs surface is related to the Fourier component of the density fluctuations by a relation analogous to (11):

$$z_0(\mathbf{q},t)=\int dz\; n(\mathbf{q},z,t)/(n_l-n_g). \tag{12}$$

If we now substitute relation (12) in Eq. (10), we obtain

$$S(q,\omega)=(n_l-n_g)\lim_{T\to\infty}\frac{1}{2\pi NT}\left|\int_0^T dt\; z_0(\mathbf{q},t)\times\exp(-i\omega t)\right|^2.$$

Thus, the dynamic structure factor which we have calculated for the transitional layer coincides with the power spectrum of the fluctuations of the separating Gibbs surface. For wave vectors lying in the interval $0.2493\leqslant|\mathbf{k}|\leqslant 0.899$, this spectrum corresponds to high-power low-frequency noise. A study of the dynamics of the separating surface by the above-described method would require an extension of the wave-vector interval.

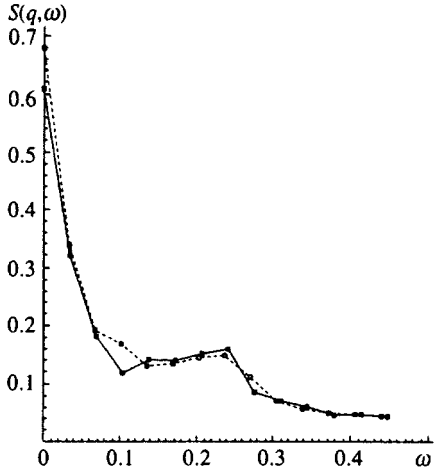

FIG. 12. Dynamic structure factor $S(q,\omega)$ of the liquid at $T=0.725$ as a function of ω for $q=0.2493$ ($N_d=81$): the solid curve corresponds to the equilibrium system, the dashed curve to vapor in vacuum.

4. EVAPORATION INTO VACUUM

In order to simulate evaporation of a liquid into vacuum, the potential at the upper face of the molecular-dynamics cell is set equal to zero (''open lid''). In this case the atoms can freely leave the cell and equilibrium states are not attained. It is possible, however, to bring the system into a stationary nonequilibrium state with prescribed constant temperature T_s on the surface of the liquid phase. This is achieved by transferring the particles that have left the molecular-dynamics cell to the near-floor region of the cell with new velocities chosen such that the total kinetic energy is conserved and the x and y components of the total momentum remain equal to zero. In such a stationary state it is possible to average the characteristics of the system over time and compare the results with the characteristics of the equilibrium system with $T=T_s$.

The maximum flux density of the atoms evaporated from the surface of a condensed body is given by the Hertz formula[20]

$$j_m=n_s\sqrt{\frac{k_BT}{2\pi m}},$$

where n_s is the density of the vapor in equilibrium with the liquid at the temperature T. The ratio of the real flux density of the evaporated atoms j to their maximum value j_m is

TABLE II.

T_s	j_{max}	j	α
0.725	0.00023956	0.0001908	0.79646
0.795	0.00049577	0.0003932	0.79311
0.897	0.00129090	0.0010280	0.79636

called the evaporation coefficient. Table II displays the results of calculations of the fluxes j_m and j and the evaporation coefficient $\alpha = j/j_{\max}$ for three temperatures of the liquid surface.

The near-constancy of α with temperature draws attention to itself. It is possible that the value $\alpha \simeq 0.8$ is universal for simple liquids.

In the simulation of evaporation into vacuum a significant change in the density profiles is observed. First, the mean vapor density is decreased severalfold below its equilibrium value. Second, the density of the bulk gas and liquid phases becomes inhomogeneous. Figure 13 shows that the density of the liquid grows linearly with z. It can be seen from Fig. 14 that the density of the vapor falls with z, also according to a linear law. Table III lists values of the density gradients of the liquid and gas phases for three temperatures, and also values of the gradients of the mean kinetic and potential energy per particle in the liquid.

Knowing the gradients of the thermodynamic quantities and the corresponding fluxes, it is easy to find the transport coefficients. For the liquid phase this problem is of great interest; however, we will not consider it here.

As was done in the equilibrium state, for the case of evaporation into vacuum we calculated spatial profiles of the mean potential energy per particle and its mean-square fluctuations from the molecular-dynamics simulation data. We also determined the potential energy distribution functions for the bulk phases and different cross sections in the transitional layer. Comparison with the equilibrium case shows that the differences are not large and show up mainly in the

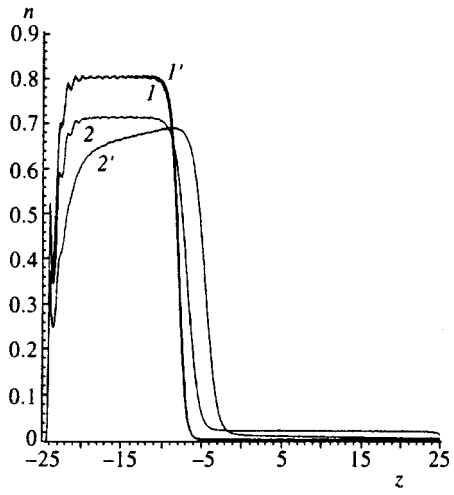

FIG. 13. Particle number density at different temperatures as a function of z: *1*—$T = 0.725$ (equilibrium system), *1′*—$T_s \approx 0.725$ (evaporation into vacuum), *2*—$T = 0.897$ (equilibrium system), *2′*—$T_s \approx 0.9$ (evaporation into vacuum).

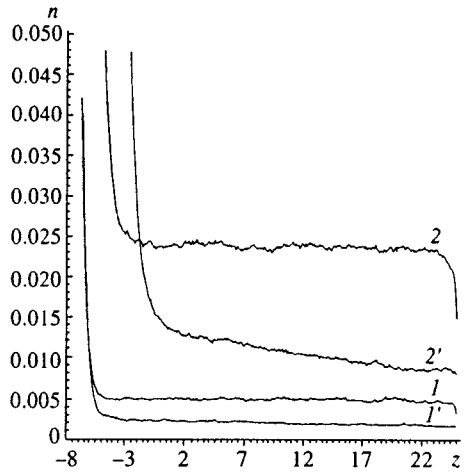

FIG. 14. As in Fig. 13, but for particle density in vapor only.

gas phase. To first order it may be assumed that the potential energy in the transitional layer at a given temperature is a function of the local density. We are led to the same conclusion by an examination of the two-particle distribution functions in the non-equilibrium system, from which it follows that evaporation into vacuum has only an insignificant effect on the pairwise distribution functions at the same densities $n(z_s)$.

Molecular-dynamics simulation of evaporation into vacuum allows us to determine a very important characteristic of the evaporation process–the velocity distribution function of the evaporated particles near the phase boundary. This function, which serves as a boundary condition for gas-kinetic calculations,[21] is usually assumed to be Maxwellian for particles moving away from the surface ($v_z > 0$) and equal to zero for particles moving toward the surface ($v_z < 0$). No serious theoretical derivation of this boundary condition is known to us. Figures 15 and 16 present results of a calculation of the particle distribution functions over the transverse velocity v_x in the bulk phases and in the transitional layer for different temperatures. All the distribution functions are quite close to Maxwellian with zero mean velocity and temperature falling somewhat as one goes from liquid to gas. Spatial profiles of the mass velocities and temperatures are shown in Figs. 17 and 18. The distribution of the particles over the longitudinal component of the velocity, v_z, has a more complicated character. As can be seen from Figs. 19 and 20, inside the liquid the distribution is Maxwellian. Figure 21 shows that the number of particles with negative velocities falls almost monotonically in the transitional layer and in the gas phase. However, even at the upper

TABLE III. Parameters of open two-phase systems.

T_s	dn_l/dz	dn_g/dz	$d\bar{u}_l/dz$	$d\bar{\epsilon}_l/dz$
0.725	+0.00058	−0.000030	−0.005	−0.0020
0.795	+0.00150	−0.000057	−0.010	−0.0038
0.897	+0.00440	−0.000210	−0.034	−0.0110

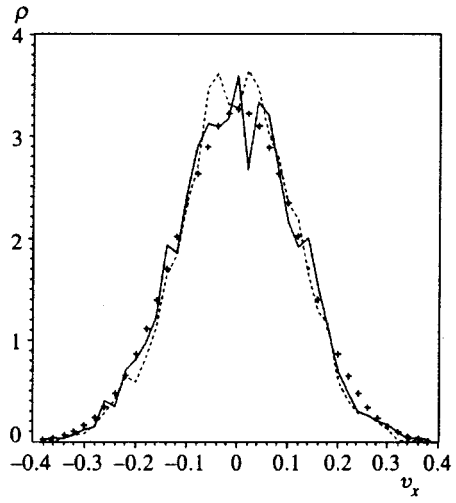

FIG. 15. The particle velocity distribution $\rho(v_x)$ inside the z_s layers for $T=0.725$: +—liquid layer $-13<z<-11$, solid curve—transitional layer $-5<z<-3$, dashed curve—gas layer $23<z<25$.

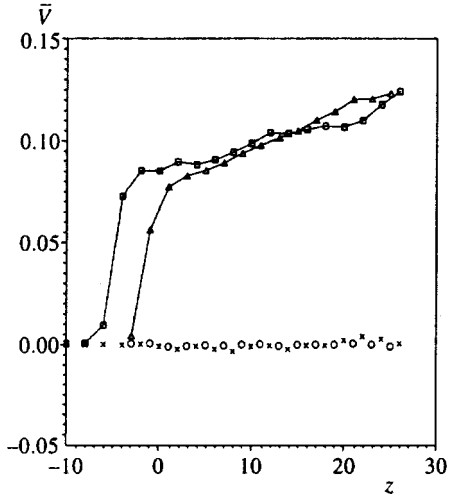

FIG. 17. Mean particle velocities $\overline{V}$ inside the z_s layers at $T_s=0.725$: □—$\overline{V_s}$, ×—$\overline{V_x+V_y}$; at $T_s=0.897$: △—$\overline{V_s}$, ○—$\overline{V_x+V_y}$.

boundary of the molecular-dynamics cell the fraction of particles with negative velocities amounts to around 8.1% for $T_s=0.725$ and around 10.7% for $T_s=0.897$.

The existence of two mechanisms of formation of particles with negative velocities in the evaporated flux may be supposed.

1. As a consequence of the low particle density in the gas, the particles of the gas are found almost always in a region in which the long-range attractive field of the surrounding particles is acting. Since the vapor density is observed to decrease with distance from the liquid surface, the mean field in which a particle moves should create a mean force directed toward the liquid. Obviously, this mechanism can operate far from the interphase boundary and in the absence of collisions.

2. Since evaporation takes place from an interphase boundary perturbed by surface waves, individual segments of which can make large angles with the xy plane, the interaction of the evaporated particles with these fluctuations can give rise to a ''tail'' of particles with negative velocities. It is clear that this mechanism can lead to the appearance of a significant returning force, but only near the interphase boundary at distances of the order of $1-2$ thicknesses of the transitional layer.

In Fig. 21 two spatial regions are clearly distinguishable, apparently corresponding to these two mechanisms of formation of a negative tail in the particle velocity distribution.

From Figs. 17 and 18 (which show the first and second moments of distributions 15, 16 and 19, 20—the profiles of the mass velocity and effective temperature, respectively) it can be seen that inside the molecular-dynamics cell the ''temperatures'' of the v_x and v_z distributions differ markedly. Strictly speaking, they cannot even be compared, since the v_z distribution function differs substantially from a Maxwellian. The ''transverse temperature,'' corresponding to nearly Maxwellian distributions over the velocity compo-

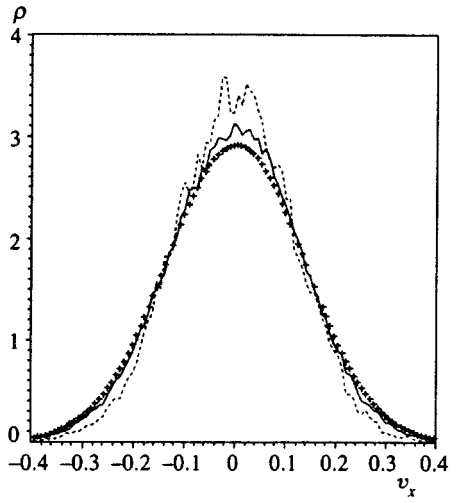

FIG. 16. The particle velocity distribution $\rho(v_x)$ inside the z_s layers for $T=0.897$: +—transitional layer $-6<z<-2$, solid curve—transitional layer $-2<z<0$, dashed curve—gas layer $22<z<24$.

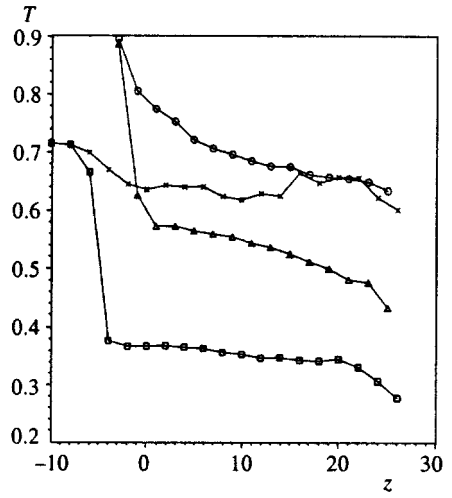

FIG. 18. Temperature inside the z_s layers at $T_s=0.725$: □—$T_{V_z-\overline{V_z}}$, ×—$T_{V_{x,y}}$; at $T_s=0.897$: △—$T_{V_z-\overline{V_z}}$, ○—$T_{V_{x,y}}$.

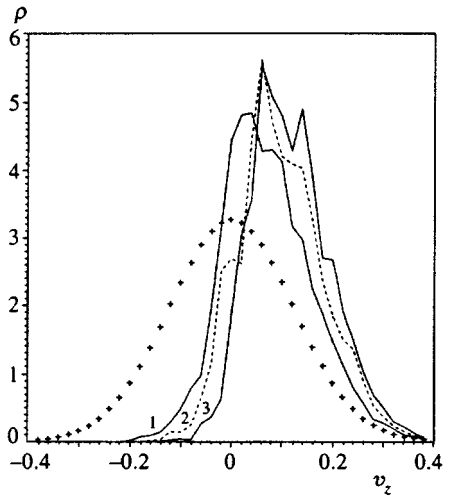

FIG. 19. Particle velocity distributions $\rho(v_z)$ plotted as a function of z_s at $T_s=0.725$: +—liquid layer $-13<z<-11$; *1*—transitional layer $-5<z<-3$; *2*—gas layer $9<z<11$; *3*—gas layer $23<z<25$.

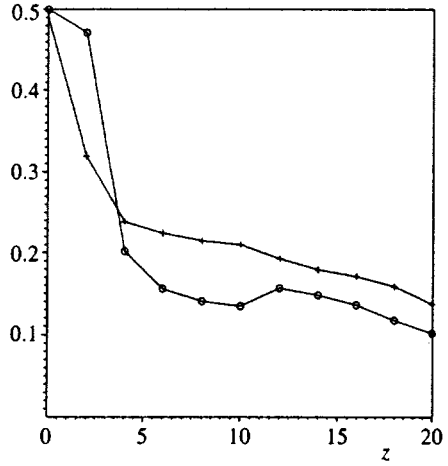

FIG. 21. Fraction of particles with negative velocity inside the z_s layers as a function of separation from the liquid boundary: ○—$T_s=0.725$; +—$T_s=0.897$.

nents v_x and v_y varies quite smoothly in the transitional layer and in the gas phase. The ''longitudinal temperature,'' corresponding to the distribution over v_z and having the meaning of the width of the distribution, falls abruptly in the transitional layer (on a scale of $5-8$ molecular-dynamics units) and then varies smoothly in the gas phase, remaining $1.5-2$ times lower than the ''transverse temperature.'' Simultaneous with the steep variation of the ''longitudinal temperature'' a finite z component of the mass velocity arises on the same scale of the thickness of the transitional layer. Of course, the decrease of the width of the v_z distribution and the simultaneous appearance of a z component of the mass velocity are a result of the transformation of the original Maxwellian distribution into an asymmetric distribution in which the fraction of particles with $v_z<0$ is radically diminished. This transformation takes place on the scale of the transitional layer. The fraction of particles with negative z component of their velocity falls with distance from the liquid surface and should become negligibly small at distances much greater than the thickness of the transitional layer. Thus, in the case of a low-density vapor the use of a semi-Maxwellian distribution as a boundary condition in the solution of gas-dynamic problems is a reasonable approximation.

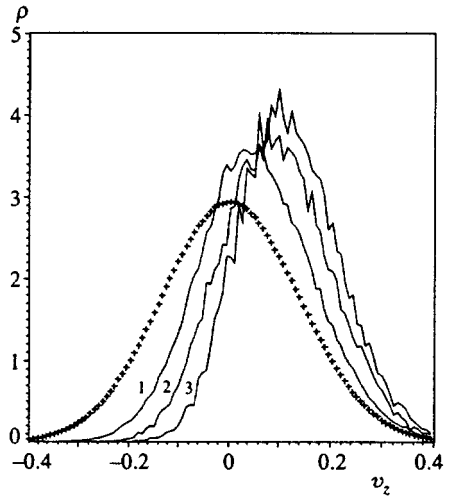

FIG. 20. Particle velocity distributions $\rho(v_z)$ plotted as a function of z_s at $T_s=0.897$: +—liquid layer $-6<z<-2$; *1*—transitional layer $-2<z<0$; *2*—gas layer $8<z<10$; *3*—gas layer $22<z<24$.

ACKNOWLEDGMENTS

This work was carried out with the support of the Russian Fund for Fundamental Research (Grant No. 95-02-4535a).

*) e-mail: zvv@hedric.msk.su

[1] Ya. I. Frenkel', *Kinetic Theory of Liquids* (Clarendon Press, Oxford, 1946).
[2] O. Knake and I. N. Stranskiĭ, Usp. Fiz. Nauk **68**, 261 (1959) [Sov. Phys. Usp. **2**, (1959)].
[3] J. P. Hirth and G.M. Pound, *Condensation and Evaporation: Nucleation and Growth Kinetics* (Pergamon Press, Oxford, 1963).
[4] S. I. Anisimov, Zh. Éksp. Teor. Fiz. **54**, 338 (1968) [Sov. Phys. JETP **27**, (1968)].
[5] M. N. Kogan, Ann. Rev. Fluid Mech. **5**, 383 (1973).
[6] S. I. Anisimov and A. Kh. Rakhmatulina, Zh. Éksp. Teor. Fiz. **64**, 869 (1973) [Sov. Phys. JETP **37**, 441 (1973)].
[7] V. I. Zhuk, Izv. Akad. Nauk SSSR, Mekh. Zhidk. Gazov, No. 2, 97 (1976).
[8] T. Ytrehus, in *Rarefied Gas Dynamics*, edited by J. L. Potter (AIAA, New York, 1977).
[9] C. Cercignani, in *Rarefied Gas Dynamics*, edited by P. Camparque (CEA, Paris, 1979), Vol. 1, p. 141.
[10] C. Cercignani, in *The Boltzmann Equation and its Applications* (Springer-Verlag, New York, 1988).
[11] S. I. Anisimov and V. V. Zhakhovskiĭ, JETP Lett. **57**, 99 (1993).
[12] L. D. Landau and E. M. Lifshitz, *Fluid Dynamics*, 2nd ed. (Pergamon Press, Oxford, 1987).
[13] E. Hairer, S. P. Norsett, and G. Wanner, *Solving Ordinary Differential Equations V.1: Nonstiff Problems* [Russian translation] (Springer, Berlin, 1987).
[14] D. W. Heermann, *Computer Simulation Methods in Theoretical Physics* (Springer-Verlag, Berlin, 1986).
[15] L. Verlet, Phys. Rev. **159**, 98 (1967).

[16] I. Z. Fisher, *Statistical Theory of Liquids* (University Press, Chicago, 1964).
[17] A. N. Shiryaev, *Probability* [in Russian] (Nauka, Moscow, 1989).
[18] J. S. Bendat and A. G. Piersol, *Random Data: Analysis and Measurement Procedures*, 2nd & rev. ed. (Wiley, New York, 1986).
[19] D. G. Triezenberg and R. Zwanzig, Phys. Rev. Lett. **28**, 1183 (1972).
[20] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pt. 1, 3rd. ed. (Pergamon Press, Oxford, 1980).
[21] S. I. Anisimov (Ed.), *Physical Kinetics and Transport Processes of Phase Transitions* [in Russian] (Nauka i Tekhnika, Minsk, 1980).

Translated by Paul F. Schippnick

Partial suppression of hydrodynamic mixing in profiled shells

N. A. Inogamov

L. D. Landau Institute of Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 24 October 1996)
Zh. Éksp. Teor. Fiz. **111**, 1347–1368 (April 1997)

The problems of stability and mixing are important in the physics of high-energy densities. Ablation-induced acceleration of foils and compression of liners entail loss of symmetry and the development of instability. The most destructive instability is the fundamental f^- mode, which conserves the pressure in Lagrangian particles. A means has been proposed to eliminate this dangerous mode, based on special profiling of the mass distribution among the subshells. The presence of this mode has led to novel proposals for limiting the degree of instability and optimization of the shells by profiling in the important case of very large density ratios at the ablation front. The solution is based on a class of new polytropes with an inverted density profile and a negative polytrope index N. In this class the density ρ of the material does not decrease towards the boundary with the vacuum, as for ordinary polytropes with $N>0$, but rather increases. This permits modeling multilayer distributions of ρ typical of inertial confinement fusion systems in which the high-density subshells form an inner core surrounding a low-pressure cavity, and the outer layers are made from low-density materials (plastic, foam-type materials, composites). It is emphasized that the distributions are self-similar, and hence both the linear and the turbulent dynamics are scale-invariant. The spectral problem of perturbations in an incompressible fluid has a hidden symmetry. Isospectral deformations of the density profile $I\{\rho_0(y)\}$ are known that leave the spectrum unchanged. It is of interest to apply the transformation I to the invariant $f^{\pm}$ modes, since they are not tied to any specific profile of $\rho_0(y)$. This paper analyzes a new type of invariant mode obtained in this way. [S1063-7761(97)01704-6]

1. INTRODUCTION

The program of laser-induced inertial confinement fusion has been developing for more than 20 years.[1–7] In that time there have been vast changes in the technology of high-power laser systems, and the technology of making targets has advanced in parallel. To obtain high compression, the laser-generated pulse (duration and shape) and the structure of the target must be well fitted to one another. Present-day methods allow one to make targets with any density distribution $\rho(r)$ (Refs. 8 and 9) by depositing coatings of controlled thickness of a broad spectrum of materials. It is proposed to use these capabilities to make profiles $\rho(r)$ that are optimum with respect to instability. In these targets the development of gasdynamic perturbations is slow relative to the fastest possible development. As is well known,[2–7] the Rayleigh–Taylor or interchange instability is the main obstacle to achieving the ignition threshold and success in the entire program of controlled thermonuclear fusion.

The optimum targets consist of a stack of subshells with a density ρ_i and thickness d_i, where $1<i<I$, in which the density ρ falls off and the thickness increases with increasing radius r according to a specified rule . The number I is large ($I\gg 1$) and therefore the relative jumps in the density are small

$$2|\rho_{i+1}-\rho_i|/(\rho_{i+1}+\rho_i)\ll 1.$$

With this condition a density distribution with small steps can be approximated by a continuous function. These targets can be either thin or thick, since on the one hand the thickness of the outer shell, d_I, and the entire structure is large, but on the other hand, the minimum densities are reached in the very thin inner subshells. Accordingly, the effective aspect ratio $R_{\rm eff}/\Delta R_{\rm eff}$ is intermediate between large values ~ 100 and small values ~ 1.

The isobaric Rayleigh-Taylor mode is very important in the theory of instability. This is the solution for which the velocity field $\mathbf{v}$ satisfies the condition $\nabla\cdot\mathbf{v}=0$. As a result, the pressure is conserved in the Lagrangian particles during the motion. This result is particularly obvious for a compressible medium, since the condition $\nabla\cdot\mathbf{v}=0$ implies conservation of volume, i. e., the particles are not compressed. It is also called the condition of incompressibility, and it is the most important property of the solution. For this reason, the condition is called the isobaric condition. It is satisfies identically the isobaric boundary condition, which specifies constant pressure at the contact surface. This condition is also called the free-boundary condition. Its growth rate $\sigma=\sqrt{gk}$ is the largest in the class of all possible unstable modes. Because of the invariance of the Lagrangian pressure the mode is insensitive to stratification and even to the equations of state in the subshells or particles, which can be arbitrary and different for the individual particles. It cannot be eliminated by profiling. These now well-known properties were established in Ref. 10, where it was also shown that the isobaric gravitational waves are closely related to trochoidal waves.

It thus appears desirable to extend the Cowling classification used in astrophysics.[11,12] According to this scheme,

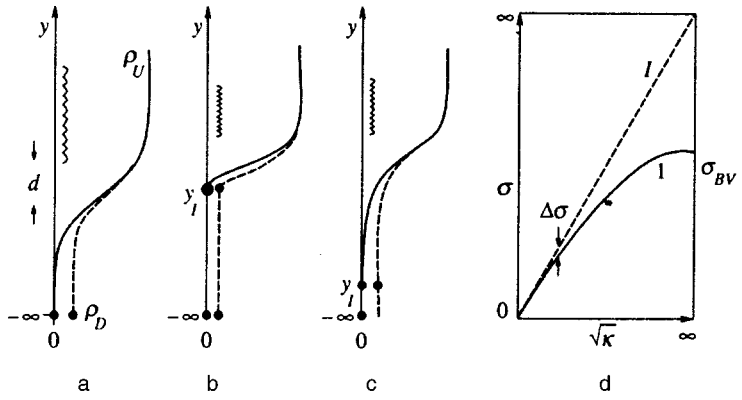

FIG. 1. The question as to whether or not the f^- mode exists in the spectrum and where it is localized relative to the valuable high-density region, shown by the wavy line.

there are p, $g^{\pm}$ and f modes. These are acoustic pressure waves (p modes), stable ($g^+, \omega^2>0$) and unstable ($g^-, \omega^2<0$) gravitational waves, and the fundamental f mode. The extended classification consists of p, g^+, g^-, f^+, and f^- modes.[1)] The f mode is now called the f^+ mode. The isobaric properties of this classification of modes has been known for a long time (see the work of Gerstner, cited in Ref. 10). The isobaric Rayleigh–Taylor mode will also be called the fundamental f^- mode.

Previous proposals for profiling in an incompressible fluid[13,14] involved the smoothing of the density jump

$$\rho(y)=\rho_D+\theta(y)(\rho_U-\rho_D)\to\rho_{\mathrm{CONT}}(y),$$

where $\rho_{\mathrm{CONT}}(y)$ is a monotonic continuous function, such as $\tan^{-1}y$ or erf$y$, with $\rho_{\mathrm{CONT}}(-\infty)=\rho_D$ and $\rho_{\mathrm{CONT}}(\infty)=\rho_U$. In the smoothed case the unbounded growth rate,

$$\sigma=\sqrt{(1-\mu)/(1+\mu)gk},$$

$$\mu=\rho_D/\rho_U=(1-A)/(1+A),$$

is cut off as $k\to\infty$ by the Brunt–Väisälä frequency

$$\sigma_{BV}=\sqrt{g\ \max\left[d\ \ln\rho_0/dy\right]}.$$

We note in passing that in a compressible medium it is equal to

$$\sigma_{BV}=\sqrt{(g/\gamma)(-d\ \ln s_0/dy)},$$

where s_0 is the hydrostatic entropy profile and γ is the adiabatic index. Because of this limitation in turbulent mixing, a time delay $\sim 1/\sigma_{BV}$ occurs in the mixing. Asymptotically at large scales the effect of the smoothing becomes insignificant.

Let us consider the growth rates and the localization of the gravitational modes. A distinct set of eigenmodes is associated with each profile $\rho(y)$. For the case depicted in Fig. 1a, we have

$$\sigma^2\to(1-\mu)/(1+\mu)gk \quad \text{for } kd\to 0,$$

$$\sigma\to\sigma_{BV} \quad \text{for } kd\to\infty.$$

The function $\sigma(k)$ is given by curve *1* in Fig. 1d. The dashed curve *1* in this figure corresponds to the f^- mode. For $\mu\to 0$ we have $\Delta\sigma(k)/\sigma(k)\to 0$ as $kd\to 0$, and curves *1* and *I* are tangent at the origin. It is important to emphasize that in this case, even for $\mu\to 0$ the value of $\sigma(\infty)$ remains finite, $\sigma(k)\to\sigma_{BV}\simeq\sqrt{g/d}$, $kd\to\infty$, where d is the width of the transition region; see Fig. 1. This is due to the lower boundary condition, according to which the perturbations are required to be bounded as $y\to-\infty$, and the fact that the zero of the function $\rho(y)$ for $\mu=0$ is located at infinity and the derivative falls off faster than the function, so that $\rho_y'(y)/\rho(y)\to 0$ as $y\to-\infty$. Therefore, the f^- mode does not appear in the spectrum of this profile for $\mu=0$.

In the case shown in Fig. 1b, $\sigma_{BV}\to\infty$ for $\mu\to 0$, and curve *1* tends to curve *I*. Mixing of the payload, indicated by the wavy line, occurs because of the growth of the f^- mode.

Most of the drop in the accelerating pressure is across this layer. Therefore most of the kinetic energy built up through acceleration is accumulated in it.

In the case intermediate between the two shown in Fig. 1c, the f^- mode is present in the spectrum for $\mu=0$. Nevertheless, this presence is merely formal, since the main contribution to the mixing of the payload is associated with the mode with a bounded growth rate $\simeq\sqrt{g/d}$, as in the case of Fig. 1a. Although the growth rate of the f^- mode is larger, nevertheless, because the point y_I is far off, this mode, which decays exponentially with distance from the point y_I, is spatially separated from the valuable layer of immediate interest and has little effect on its mixing.

Consequently, with the optimum profile it is proposed to separate the f^- mode (Fig. 2b) from the payload (shown by the wavy line in Fig. 2a) by the extended low-density tail (the section LDT in Fig. 2a). The high-density subshells, which are the most important for the compression process,

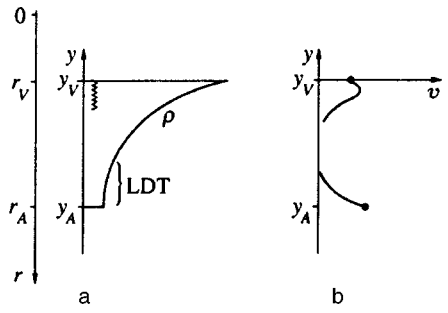

FIG. 2. a — Smoothed profile of a multilayer shell bounded by isobaric boundaries y_V and y_A. b — Localization of the f^- and g_0^- modes near the y_A and y_V boundaries, respectively.

are situated on the payload, next to the boundary r_V or $y_V=0$. The mode responsible for the mixing in this part is localized near the boundary y_V (see Fig. 2b). Therefore two different regions of instability and mixing arise near the boundaries r_V and r_A (or $y_V=0$ and y_A).

It should be mentioned that in the long-wavelength case,

$$k<a^2g/4b^2v_A^2=(a^2/4b^2)M_A^{-2}/h, \quad h=c^2/g,$$

the boundary condition at the ablation front can be approximated by the isobaric condition. Here v_A is the velocity of the ablation front with respect to the cold material and c is the speed of sound in the cold material. An estimate follows from the well-known Takabe formula,[15]

$$\sigma=a\sqrt{gk}-bkv_A,$$

$a\approx 0.9$, $b\approx 3$, and the Mach number $M_A=v_A/c$ of the front is small compared with unity. It is also noted that the mixing in the inner layers of a complex shell is of interest in itself, regardless of the situation on the ablation front. This question is of particular interest in the case of complex large targets and long laser pulses.

The optimum profiles are self-similar in the sense that all the subshells hold equal status with respect to the growth of instability. This means that within a single profile, regions with enhanced instability are eliminated. In these regions the density variation is peaked towards the boundary of the inner cavity, r_V. These profiles are polytropic, and belong to the class of power-law functions, $\rho\propto(-y)^N$. The parameter N is called the polytrope index. For these polytropes the index N is negative. Therefore $\rho\to\infty$ as $y\to 0$. The ordinary polytropes studied in connection with geophysical and astrophysical applications[16–20] satisfy $N>0$ and $\rho\to 0$ as $y\to 0$.

In self-similar profiles there are no dimensional scales. Therefore the spectral characteristics of the linear theory are scale-invariant, and self-similar formulas are generated. The expression for σ is obtained by dimensional analysis and is equal to $\sigma_m=\sqrt{A_m gk}$, where A is a dimensionless number, and m labels the countable set $\{m\}$ of discrete eigenvalues. Qualitatively the formula for σ is found to be the same as in the simple case of a discontinuity in an incompressible fluid. For an instability, the mode of interest is that with the largest growth rate. It corresponds to the ''ground state,'' with $m=0$. The dependence of σ on N was studied briefly in Ref. 21. In the present paper not only are the eigenvalues calculated, but also the eigenfunctions; the locations of their maxima are determined; a study is made on how a finite pressure in the inner cavity and a finite density of the compressible multilayer shell affect the final result, as well as the effect of a rigid wall; also included in the analysis is the role of the equation of state, which is important, since the subshells are made out of different materials; and the structure of a new kind of self-similar Rayleigh–Taylor turbulence is studied. The structure remains self-similar in a compressible medium. Previously, self-similarity was known only in the incompressible case. It should also be noted that the dependence of A_0 ($A_0=\sigma_0^2/gk$) on the variable parameter N can be used for optimization. A decrease in N improves the one-dimensional dynamics of the shell and reduces the amount of energy E necessary for ignition. On the other hand, a result of this reduction is that instability begins to develop faster, and the one-dimensional modeling becomes inadequate for the real situation. Therefore there must exist an intermediate value of N corresponding to a lower ignition threshold. This is of no small importance in the situation where this threshold is still to be achieved, the more so in that it only requires changes in the targets, rather than expensive upgrading of the parameters of the laser system

2. PRINCIPLES OF STABILIZATION

The system of gasdynamic equations is

$$D_t\rho+\rho\nabla\cdot\mathbf{v}=0, \quad \rho D_t\mathbf{v}+\nabla p-\rho\mathbf{g}=0, \quad D_t s=0,$$

$$D_t=\partial_t+(\mathbf{v}\cdot\nabla), \quad s=s(p,\rho).$$

Linearizing it in the standard way about the unperturbed state, we obtain the equation[16]

$$(p_L)''_{yy}-\frac{\rho'_{0y}}{\rho_0}(p_L)'_y-k^2Wp_L=0,$$

$$W=1-\frac{\rho'_0}{\rho_{0y}}\frac{g}{\sigma^2}+\left(\frac{\sigma^2}{gk}-\frac{gk}{\sigma^2}\right)\frac{g}{kc_0^2},$$

$$p_L=p-\frac{g}{i\omega}\rho_0 v, \tag{2.1}$$

for the perturbed Lagrangian pressure p_L, which is valid for an arbitrary equation of state. Here and below, $g=|\mathbf{g}|$, $c^2=(\partial p/\partial\rho)_S$, and p is the perturbed Eulerian pressure. If the equation $s=p/p^\gamma$ is satisfied, then the function W in Eq. (2.1) can be written as

$$k^2W=k^2+\left(\frac{\sigma}{c_0}\right)^2+\frac{k^2}{\gamma}\frac{g}{\sigma^2}\frac{s'_{0y}}{s_0}. \tag{2.1$'$}$$

Otherwise

$$\rho_0\propto(-y)^N, \quad p_0\propto(-y)^{N+1}, \quad c_0=\sqrt{\gamma g(-y)/(N+1)},$$

$$s_0\propto(-y)^\theta, \quad \theta=1-N(\gamma-1), \tag{2.2}$$

and Eq. (2.1) becomes

$$\eta(p_L)''_{\eta\eta}-N(p_L)'_\eta-(\eta-2a-N)p_L=0,$$

$$a=-\frac{1}{2}\left(N-\frac{N+1}{\gamma}\Sigma^2+\frac{\theta}{\gamma\Sigma^2}\right),$$

$$\eta=ky, \quad \Sigma^2=\frac{\sigma^2}{gk}. \tag{2.3}$$

With the substitution $p_L=e^\eta u$ and $\eta=-z/2$, it is transformed into the Kummer equation

$$zu_{zz}+(-N-z)u_z-au=0$$

(see Ref. 22, p. 504), and the substitution $p_L=\eta^{N/2}\psi$ and $\eta=z/(2a+N)$ converts Eq. (2.3) into the time-independent Schrödinger equation

$$\psi_{zz}-(-E+U)\psi=0$$

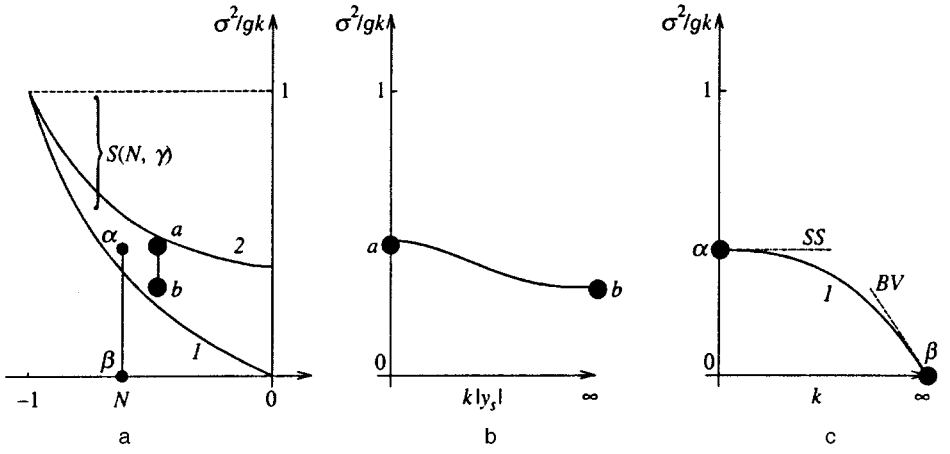

FIG. 3. a — Partial stabilization involves a decrease in the growth rate by the amount $S(N,\gamma)$ relative to the growth rate of the f^- mode. The stabilization is retained when the adiabatic index is varied (b) and when the density of the shell is finite at the boundary with the inner cavity (c).

for a particle in a Coulomb field with a total angular momentum l equal to $N/2$, energy $E=-(2a+N)^{-2}$, and a potential $U=-1/z+l(l+1)/z^2$.

We shall assume that the upper and lower mixing zones (Fig. 2b) are spatially separated. Then to describe g-perturbations in the upper zone near the surface $y_V=0$ we require that the solution of Eq. (2.3) go to zero as $y\to-\infty$. It is expressed in terms of the Kummer U functions[22] and is equal to

$$p_L=e^{\eta}U(a,-N,-2\eta),\quad M(a,b,z)=\sum_{j=0}^{\infty}\frac{a_j}{b_j}\frac{z^j}{j!},$$

$$a_j=a(a+1)\ldots(a+j-1),\quad a_0=1, \tag{2.4}$$

$$U=-\frac{\pi}{\sin N\pi}\left[\frac{M(a,-N,-2\eta)}{\Gamma(a+N+1)\Gamma(-N)}-(-2\eta)^{N+1}\frac{M(a+n+1,\,N+2,-2\eta)}{\Gamma(a)\Gamma(N+2)}\right]. \tag{2.4$'$}$$

Now we must consider the condition on the upper boundary. We assume that this is an isobaric boundary. Then its velocity and displacement are nonvanishing, but the perturbation of the Lagrangian pressure vanishes on it. Let us consider first the simplest case, where we can neglect the pressure in the inner cavity. Then the unperturbed pressure p_0 vanishes on the boundary y_V, i. e., according to the distributions (2.2) the point y_V is at the origin ($y_V=0$). We require $p_L(y_V)=0$ and $y_V=0$. Analyzing the equilibrium equation, we find that the index N is bounded from below: $N>-1$. We see that to satisfy the condition $p_L(0)=0$ we must get rid of the first term in the expression for U given by (2.4)$'$. To do so, we require that it occur at one of the poles $z_m=-m$, $m=0,1,2,\ldots$ of the gamma function $\Gamma(z)$ in the denominator of the first fraction in Eq. (2.4)$'$. Consequently, we have

$$a_m+N+1=-m.$$

From this result and the definition of a in Eq. (2.3) it is easy to obtain the dispersion equation. We discard the branches belonging to the modes p and f^+, while the g^+ modes are absent. Among the g^- modes, we are interested in the mode $m=0$ with the largest growth rate

$$\Sigma_0^2=\sqrt{\beta^2+\theta/(N+1)}-\beta,\quad \beta=\gamma(N+2)/2(N+1).$$

In the incompressible case, $\gamma\to\infty$, this formula simplifies and takes the form

$$(\Sigma_0^2)_{\text{rigid}}=-N/(N+2).$$

Compressibility (for the same profiles of ρ and p) enhances the instability, $\Delta(N,\gamma)=\Sigma_0^2-(\Sigma_0^2)_{\text{rigid}}>0$. The function $\Delta(N,\gamma)$ increases monotonically for fixed N as γ decreases, i. e., with increasing compressibility. It reaches the largest value, equal to $\Delta(N,1)$, at $\gamma=1$. The relative role of the compressibility varies with N. As N decreases, the gap

$$\Delta(N,1)=(\Sigma_0^2)_{\text{soft}}-(\Sigma_0^2)_{\text{rigid}}=\frac{N+2}{2(N+1)}\left(\sqrt{1+4\frac{N+1}{(N+2)^2}}-1\right)+\frac{N}{N+2} \tag{2.5}$$

between the soft case ($\gamma=1$, curve *2*, Fig. 3a) and the hard case ($\gamma=\infty$, curve *1*, Fig. 3a) is reduced. It goes to zero at $N=-1$.

For any fixed γ the growth rate σ decreases with increasing N. For $N=-1$ it increases to its maximum value, $\sqrt{gk}$, coinciding with the growth rate of the f^- mode. An expansion of Σ_0^2 at the point of the maximum growth rate, $N=-1$, $\Sigma^2=1$, is

$$\Sigma_0^2=1-2\,\delta N+(2+4/\gamma)(\delta N)^2+O[(\delta N)^3],$$

where $\delta N=N+1$. As one can see, the term linear in δN does not depend on γ. Therefore the expansion of $\Delta(N,1)$ at this point begins with the term quadratic in δN. The region of present interest, $N<0$, $\Sigma^2>0$, is a square. It is bounded by the upper and left limits. The upper limit is the growth rate $\sigma^2=\sqrt{gk}$ of the f^- mode, and the left limit is the condition $N>-1$, which follows from the requirement that the mass be finite near the point $y=0$.

Reduced growth rates are found within the gap $\Delta(N,1)$ defined by Eq. (2.5) As we see, the polytropic growth rates are shifted downward by the amount $S(N,\gamma)$ compared to the growth rate of the f^- mode (see Fig. 3a). This is how the stabilizing action of the profiling works.

3. THE ROLE OF THE EQUATIONS OF STATE IN THE SUBSHELLS

This paper presents an analysis of the spectral properties of polytropes inverted in density. The results are used for modeling the growth of instability in multilayer targets. The question arises: how applicable is this approach for the real situation? The target consists of a large number of subshells made of different materials. Can one describe the development in a unified manner when the equation of state of the material is first, nonideal, and second, different in the different subshells?

We shall find an answer to these questions. Targets with a power-law profile of the density ρ are formed in practice by selecting a sequence of materials that increase in density and specifying the thickness of the subshells. It is assumed that they are in an effective gravitational field. This is a standard approximation used by many investigators. If this assumption is valid, then the equation of hydrostatics, $p_y' = -g\rho$ is also valid. This means that if the pressure p_V in the cavity is neglected, then even the pressure profile will follow a power law. According to the equations of hydrostatics, the exponents of p and ρ differ by unity. Consequently the ratio p/ρ is a linear function of the position.

The basic model is Eq. (2.1). It is valid for any equation of state. In order to arrive at the solutions, given in Sec. 2, inertial and thermodynamic conditions must be satisfied. The inertial conditions (the power-law profile of ρ, linearity of the ratio p/ρ) are already satisfied. It remains to consider the thermodynamic condition, which says that c^2 must be a linear function of position.

Let us find this condition. We write the relation between the thermodynamic functions c^2 and p/ρ. By definition we have

$$c^2 = \gamma\frac{p}{\rho}, \quad c^2 = \left(\frac{\partial p}{\partial \rho}\right)_S, \quad \gamma = \left(\frac{\partial \ln p}{\partial \ln \rho}\right)_S. \tag{3.1}$$

In the case of a nonideal equation of state the exponent γ in adiabatic processes is a function of a single thermodynamic variable, for example, the density:

$$\gamma = \gamma(\rho).$$

In the various subshells this function will be $\gamma^j(\rho)$, where the index j labels the subshell. This is a hydrostatics problem, i.e., the distribution is stationary. In equilibrium the densities of the subshells are ρ_j and are fixed. Therefore the exponents depend only on the label

$$\gamma^j(\rho) = \gamma^j(\rho_j) = \gamma^j.$$

The ratio p/ρ in Eq. (3.1) is linear. Therefore for this approach to be applicable, γ must not depend on j; that is, the value of γ must be the same in the subshells.

The values of the exponents γ vary somewhat for different materials. The effect of these variations on the main result of this work is small. The main result is a partial stabilization of the instability (Fig. 3a). This involves a reduction in the growth rate σ by the shift S. The conclusion that σ is reduced also holds when the exponents γ are the same in all the subshells. Here the role of the inertial conditions are more important than the thermodynamic conditions.

To show this, let us consider the two-shell case. We set $\gamma = \gamma_U$ for $y_S < y < 0$ and $\gamma = \gamma_D$ for $y < y_S$. Then the growth rate σ for $k|y_S| \gg 1$ is determined by the exponent γ_U, and for $k|y_S| \ll 1$ by the exponent γ_D. In the intermediate region, $k|y_S| \sim 1$, σ interpolater monotonically (see Fig. 3b). This result is explained by the fact that the small-scale perturbations are localized inside the outer layer, and this layer is unimportant for long-wavelength perturbations.

This discussion implies that the growth rate in the problem with a variable exponent $\gamma(y)$ remains in the gap between curves *2* and *1* (see the points a and b in Figs. 3a and *3*b). Consequently the growth rate σ remains reduced.

4. MULTILAYER PROFILE WITH FINITE DENSITY

Let us investigate how the result changes if the pressure in the inner cavity is taken into account and we give up the condition that the density of the shell material is infinite at the boundary with the cavity. As before, we place the origin of the coordinate system $y = 0$ at the point where the extrapolation of the pressure distribution $p_L \propto (-y)^{N+1}$ goes to zero. The unperturbed boundary is at the point $y = y|_\Gamma = -\varepsilon$. It is easy to see that the value of ε is determined by the pressure p_ε in the cavity:

$$\varepsilon \propto p_\varepsilon^{1/(N+1)}.$$

In the case of a ''cut-off'' power-law distribution, the problem is no longer self-similar. Its spectrum is shown in Fig. 3c. Here the straight line *SS* refers to the self-similar (power-law) spectrum; its value of σ_0 is $\sigma_0 = \Sigma_0(N,\gamma)\sqrt{gk}$. The asymptote *BV* corresponds to the Brunt-Väisälä frequency $\sigma_{BV} = \sqrt{g|s_{0y}'|/\gamma s_0}$, calculated at the edge of the profile at the point $y = -\varepsilon$. It limits the growth of σ. The square of the dimensionless ratio is $\Sigma_{BV}^2 = (\theta/\gamma)/k\varepsilon$, where θ is defined in Eq. (2.2). The resulting dispersion relation is given by curve *1*. For $k \to 0$ and $k \to \infty$ it goes to the limiting functions *SS* and *BV*, respectively. The transition region between these asymptotes occurs for $k \sim k_\varepsilon = 1/\varepsilon$.

In the section where the shells are matched the pressure in the cavity is low: $p_\varepsilon \ll p_A$. Therefore the shift of ε is small compared to the total thickness of the multilayer shell. Under these conditions there is some justification for calculating a correction to the self-similar growth rate due to the backpressure. We shall therefore be interested in the situation where the parameter $k\varepsilon$ is small: $k\varepsilon \ll 1$. Let us find the first-order (in $k\varepsilon$) correction to the growth rate $\sigma_0 = \Sigma_0\sqrt{gk}$.

The general dispersion relation, which is valid for an arbitrary value of $k\varepsilon$, is found from the condition $p_L(-k\varepsilon) = 0$. To find the solution $p_L(\eta)$ that satisfies the condition of vanishing at infinity, we use expressions (2.4) and (2.4)′ and substitute them into this boundary condition. As a result, we obtain the general relation

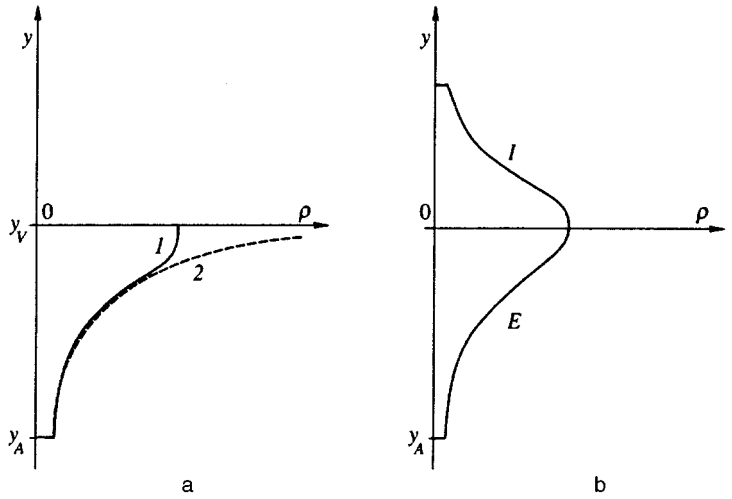

FIG. 4. a — The conclusion that partial stabilization occurs in the acceleration stage carries over to the case of any finite smooth profile *1*. b — The two-sided profile E–I inhibits the instability during both the acceleration stage and the deceleration stage.

$$(2k\varepsilon)^{N+1}\frac{\Gamma(-N)}{\Gamma(a)}$$

$$-\frac{\Gamma(N+2)}{\Gamma(a+N+1)}\frac{M(a,-N,2k\varepsilon)}{M(a+N+1,N+2,2k\varepsilon)}=0. \qquad (4.1)$$

As $\varepsilon\to 0$ the first term in Eq. (4.1) vanishes, since $N>-1$, and the function M tends to unity (see their definition, Eq. (2.4)). In this case, which was analyzed in Sec. 2, it is required that the argument of the gamma function coincide with a pole

$$a+N+1=-m, \quad a=a_M, \quad m=0,1,\ldots.$$

Let us see how the answer changes for $\varepsilon\neq 0$. For $k\varepsilon\ll 1$, as before we are in the neighborhood of a pole. In this neighborhood, as is well known (Ref. 23, p. 867) the asymptotic behavior is

$$\Gamma(z)\approx[(-1)^m/m!]/(z+m), \quad m=0,1,\ldots .$$

This gives us the desired expression for the correction

$$a_m=-m-N-1+\Delta a_m,$$

$$\Delta a_m=\frac{(-1)^m}{m!}\frac{\Gamma(-N)}{\Gamma(-m-N-1)\Gamma(N+2)}(2k\varepsilon)^{N+1}. \qquad (4.2)$$

It is interesting to note that the function $a_m(k)$ for acoustic (p) modes and gravitational ($g^{\pm}$) modes are the same. It thus follows, by the way, that even their eigenvalues coincide (for any value of ε). The reason is that relations (2.4) and (2.4′), which define p_L, do not contain their frequencies, which of course are different, but the quantity a. The rest of the eigenfunctions (v, ρ, etc) of the p modes and the g modes are different, so that the frequency enters into their definition.

Let us analyze these relations. The function $\Gamma(x)$ for real $x<1$ alternates between identical positive and negative parts.[23] It follows that for any $N>-1$ the corrections Δa_m (Eq. (4.2)) are negative. Substituting Eq. (4.2) into (2.3), which connects a and σ, we obtain a biquadratic equation for σ. One of the roots of this equation (taken with a minus sign) gives a negative frequency ($\sigma^2=-\omega^2$) for the acoustic modes, and the other to the gravitational modes. An analysis of the frequency dependence shows that because of the pressure p_ε the p modes are hard, and the stable and the unstable gravitational modes are soft. As a consequence, curve *1* in Fig. 3c is deflected below the straight line SS with increasing k.

Therefore the conclusion regarding partial stabilization still holds even for a profile with a finite density (see points α and β in Figs. 3c and 3a). The same conclusion also holds for profile *1* with any smooth cutoff *2* of the power-law dependence for $y\to 0$ (see Fig. 4a). We note that this carries over to a two-sided stabilizing profile (Fig. 4b), which weakens the instability and retards the onset of mixing, both in the stage in which the shell is pushed toward the center by the accelerating pulse, and in the stage where it is decelerated by the back-pressure of the fuel. In the acceleration stage the instability is stabilized by the outer part of the profile, E, and in the deceleration stage by the inner part, I.

5. STRUCTURE OF THE POLYTROPIC PERTURBATIONS

Let us address the analysis of the eigenfunctions and their structure. From formulas (2.4) and (2.4)′, and the dispersion relation $a_0=-N-1$, we find that for the g_0^- mode the field of the Lagrangian pressure is

$$p_L=(-\eta)^{1+N}e^{\eta}, \quad \eta=ky.$$

The maximum in the function $p_L(\eta)$ is at the point $\eta_{PL}^{\max}=-1-N$. The vertical component of the velocity v is calculated from p_L with the relation

$$v=i\frac{k}{\omega}\frac{\Sigma^2}{(\Sigma^2+1)(\Sigma^2-1)}\frac{\Sigma^2(p_L)'_\eta-p_L}{\rho_0}. \qquad (5.1)$$

This follows from the system

$$v'_\eta+\frac{v}{\Sigma^2}=\left(1+\frac{\sigma^2}{k^2c_0^2}\right)\frac{k}{\sigma\rho_0}p_L,$$

$$\frac{v'_\eta}{\Sigma^2}+v=\frac{k}{\sigma\rho_0}\left[(p_L)'_\eta+\frac{g}{kc_0^2}p_L\right], \qquad (5.1)'$$

which precedes Eq. (2.1). This component is equal to

$$v=(1-b^{-1}\eta)e^{\eta},\quad -\eta_V^{\max}=1-b,\quad b=\frac{1+N}{\Sigma_0^{-2}-1}>0. \tag{5.2}$$

The velocity Eq. (5.2) is normalized to unity on the surface $y=0$. The position of the maximum of v is also written down in Eq. (5.2). It is located within the polytrope. The proof of this fact will not be presented because it is very long and involved. For $N\to -1$ the pressure becomes a monotonic function. The localization of the g^- modes near the surface must be emphasized. They are localized in a layer adjacent to the boundary with the vacuum. The thickness of this layer, $|\Delta y|$, is of order $|\eta_V^{\max}|/k\sim 1/k$.

Let us see how the velocity distribution depends on N. The expansion in N of the expression that gives the position of the velocity maximum $(-\eta_V^{\max})$ near the point $N=-1$ is

$$-\eta_V^{\max}=1/2+(1/2-1\gamma)(\delta N)+(\delta N)^2/\gamma+O[(\delta N)^3].$$

In the incompressible case we have

$$(-\eta_V^{\max})=1+N/2.$$

Now we shall compare the position of the maxima of the perturbed Lagrangian pressure $(-\eta_{PL}^{\max})$ and velocity $(-\eta_V^{\max})$. One can show that there exists an $N_{\rm SEP}$ such that for $-1<N<N_{\rm SEP}$ the maximum in the velocity is above the maximum in the pressure, i.e., furthest from the vacuum interface, while for $N_{\rm SEP}<N<0$ on the other hand, the maximum pressure is found at the greatest depth.

We shall now analyze the behavior of the eigenfunctions in the neighborhood of the point $y=0$. An analysis of the linearized system gives

(i) $p_L\to 0$ for $y\to 0$;

(ii) the velocity v remains finite at the interface with the vacuum;

(iii) the ratio $v/c_0\propto 1/\sqrt{-y}$ for $|y|\ll 1$;

(iv) the perturbation of the Eulerian pressure satisfies $p_E=p_L+(g/i\omega)\rho_0 v\to\infty$ for $y\to 0$. This last condition, which includes the details of the density-inverted polytrope, is what distinguishes it from astrophysical situations, where $\rho_0\to 0$ as $y\to 0$, and consequently $p_E\to 0$ as $y\to 0$.

The degree of nonlinearity is governed by the function $a(\eta)=p_L/p_0$. It characterizes the relative amplitude of the perturbation. The vacuum amplitudes of the velocity $v(0)$ and of the relative pressure $a(0)$ are proportional to each other ($v\propto a, |y|\ll 1$). If this ratio is small, $p_L\ll p_0$, then the perturbation is linear. One can see that $p_L/p_0\propto e^{\eta}$ for $m=0$. The maximum of the function $a(\eta)$ occurs at the boundary with the vacuum, $y=0$. An important point is that this function remains finite at $y=0$. It therefore follows that if $a(0)\ll 1$ holds, then the perturbation is linear everywhere.

The singularities in v/c_0 and p_E (see iii and iv) are fictitious. Their existence does not mean that the perturbation goes over to a nonlinear regime, or that shock waves are formed in the vicinity of the vacuum boundary for any arbitrarily small perturbation at large distances from the boundary. They are, rather, associated with the displacement $\delta\eta$ of the perturbed boundary, because of which the physical boundary is located not at the point $y=0$, but at $y=0+\delta\eta$.

Although the Mach number v/c_0 is formally large, near the boundary the essential factor is the displacement of a thin near-boundary layer as a whole.[2)] In this layer the local velocity of sound c_0 is lower than $v(0)$. However, the buildup of this velocity occurs more slowly than the time required for sound to pass through the layer. This means that the acceleration is small $|(\delta\eta)_{tt}''|\ll g$. Accordingly, the accelerating pressure gradient is also small; that is to say, $p_L\ll p_0$ holds for the pressure itself, which means that the perturbation is linear.

6. CONTACT WITH A RIGID BOUNDARY

Let us consider yet another interesting modification. We shall examine what happens when the upper boundary condition is replaced. The condition we consider is in a sense opposite to the free-boundary condition. We assume that the upper wall is rigid. We take it to be located at the point $y=-\varepsilon$.

The solution $p_L(\eta)$ of Eq. (2.1), matched with the lower boundary condition, is given by the relations (2.4) and (2.4)′. To find the spectrum we must require that a condition imposed on the wall be satisfied. Namely, we set

$$v(-k\varepsilon)=0.$$

This condition along with relation (5.1) gives the dispersion equation

$$[(p_L)_\eta'\Sigma^2-p_L]|_{-k\varepsilon}=0.$$

Substituting Eqs. (2.4) and (2.4′) into this and differentiating, we obtain

$$\begin{aligned}&-\frac{1-\Sigma^{-2}}{2}\frac{M(a,-N,2k\varepsilon)}{\Gamma(a+N+1)\Gamma(-N)}\\&\quad+(2k\varepsilon)^{N+1}\frac{1-\Sigma^{-2}}{2}\frac{M(a+N+1,N+2,2k\varepsilon)}{\Gamma(a)\Gamma(N+2)}\\&\quad-\frac{a}{N}\frac{M(a+1,-N+1,2k\varepsilon)}{\Gamma(a+N+1)\Gamma(-N)}-(N+1)\\&\quad\times(2k\varepsilon)^N\frac{M(a+N+1,\,N+2,\,2k\varepsilon)}{\Gamma(a)\Gamma(N+2)}\\&\quad-(2k\varepsilon)^{N+1}\frac{a+N+1}{N+2}\frac{M(a+N+2,\,N+3,\,2k\varepsilon)}{\Gamma(a)\Gamma(N+2)}=0.\end{aligned} \tag{6.1}$$

We examine first the long-wavelength limit $k\varepsilon\ll 1$. If $N>0$ holds, then in the zeroth approximation in the small parameter $k\varepsilon$ the answer coincides with that calculated above. Specifically, the spectrum is given by the equation

$$a_m+N+1=-m.$$

The first order correction in $k\varepsilon$ is now different. It is equal to

$$\Delta a_m=\frac{(-1)^m}{m!}\left(-\frac{1-\Sigma_m^{-2}}{2}+\frac{N+1+m}{N}\right)^{-1}\frac{(N+1)\Gamma(-N)}{\Gamma(-N-1-m)\Gamma(N+2)}(2k\varepsilon)^N.$$

As can be seen, the power to which the parameter $k\varepsilon$ enters into the answer is smaller (compare Eq. (4.2)).

However, if we have $-1<N<0$, then the answer changes even in zero order, compared to the case of an isobaric boundary, as described in Sec. 2. In this case, it is necessary to eliminate the large fourth term in Eq. (6.1). To do so, it must be located at a pole of $\Gamma(a)$. It then follows that $a_m=-m$, $m=0, 1, \ldots$. To determine Σ from this equation we use formula (2.3) for a. Solving the equation, we find

$$\Sigma_m^2=-\beta+\sqrt{\beta^2+\theta/(N+1)},$$

where

$$\beta=(\gamma/2)(2m-N)/(N+1).$$

It is the largest growth rate that is of interest. Therefore we must consider the first values of m.

For $m=0$ the radicand is a perfect square. Therefore the instability for any values of $N(-1<N<0)$ and γ corresponds to the answer $\Sigma^2\equiv 1$. This should be taken as an indication that we should study the case $\Sigma^2=1$. In this case the system (5.1′) degenerates, and relation (5.1) loses any meaning (it contains the expression 0/0). We shall describe briefly the analysis. Setting $\Sigma^2=1$ in Eq. (2.1), we find the solution to this equation

$$p_L=e^{\eta}.$$

Now we must find a relation between p_L and v. Setting $\Sigma^2=1$ and $p_L=e^{\eta}$ in the system (5.1′), we see that the equations in the system become identical. They are equal to

$$v'_{\eta}+v=\left(1+\frac{\sigma^2}{k^2c_0^2}\right)\frac{p_L}{\rho_0}. \tag{6.2}$$

The unimportant factor is chosen to renormalize p_L. The field p_L is known, so we shall study v. Making the substitution $v=we^{-\eta}$ in Eq. (6.2), we obtain

$$w'_{\eta}=\left(1+\frac{\sigma^2}{k^2c_0^2}\right)\frac{e^{2\eta}}{\rho_0}. \tag{6.3}$$

The general solution of Eq. (6.2) is

$$v(\eta)=(w_p+C)e^{-\eta}.$$

where w_p is a particular solution of Eq. (6.3) and C is an arbitrary constant. The right-hand side, $R(\eta)$, of Eq. (6.3) is positive $(\sigma^2/k^2c_0^2=(N+1)/(-\eta)\gamma>0)$. Therefore, $w_p(\eta)$ is a monotonically increasing function of η. For $(-\eta)\gg 1$ the function R is small. Therefore at large distances, this function increases slowly. For $(-\eta)\to 0$ we have

$$R\propto 1/(-\eta)^{1-(-N)}.$$

Since $0<(-N)<1$, this singularity is integrable. This means that the integral

$$w_p(0)-w_p(-\infty)=\int_{-\infty}^{0}R\ d\eta \tag{6.4}$$

is finite. In order to satisfy the condition at the rigid upper boundary, we must set $C=-w_p(0)$. With this choice of C the function $v(\eta)$ vanishes at the boundary. It is equal to

$$v(\eta)=[w_p(\eta)-w_p(0)]e^{-\eta}.$$

By virtue of Eq. (6.4) this function diverges exponentially at the lower boundary. Consequently it does not satisfy one of the two boundary conditions, and must be discarded.

Thus it has been shown that the value $m=0$ must be discarded. Let us go now to the case $m=1$. The corresponding function $\Sigma_R(N,\gamma)$ is smaller than unity and considerably smaller than the function $\Sigma_F(N,\gamma)$, which was obtained with isobaric upper boundary conditions with $m=0$ (here we have used the subscripts R and F in order to distinguish the rigid and the free cases). This is as it should be, since the rigid boundary provides a stabilizing influence.

The above discussion has referred to the asymptotic behavior $k\varepsilon\ll 1$. For intermediate wavelengths, $k\sim 1/\varepsilon$ the increment $\Sigma_R(k)$ is adjusted to the Brunt–Väisälä asymptotic behavior. This means that the short-wavelength BV asymptotes in Fig. 3c of the growth rates $\Sigma_R(k)$ and $\Sigma_F(k)$ are the same.

7. ISOSPECTRAL DEFORMATIONS AND INVARIANT POINT IN THE ACOUSTIC DISPERSION RELATION

The special behavior of the mode with $m=0$ in the case with the rigid boundary (see Section 6) stimulates interest in analyzing modes that are invariant with respect to the structure of the hydrodynamic profile and obey the limiting dispersion law $\omega^2=\pm gk$. Let us consider the $f^{\pm}$ modes. The spectral problem of perturbations in incompressible fluids has a hidden symmetry. It is known that the isospectral deformation

$$\rho_0(y)\to I\{\rho_0(y)\}=\tilde{\rho}_0(y),$$

leaves the spectrum stationary. Unlike the eigenvalues, the eigenfunctions transform in a nontrivial manner. It is a matter of interest to apply the transformation I to the invariant $f^{\pm}$ modes, since they are not ''tied'' to any specific profile. It turns out that in this way we obtain a new type of invariant mode. The solution with $m=0$, obtained in Section 6 belongs to this type.

7.1 Isospectral Density Inversion

Inversion of the density

$$\rho(\eta)\to\tilde{\rho}(\eta)=\frac{1}{\rho(-\eta)} \tag{7.1}$$

does not change the eigenvalue spectrum. This property was observed in Ref. 13 in numerical calculations of the characteristic equations of step-function density profiles with several steps. Then in Ref. 24 it was derived for the particular case of a three-layer transition zone between two homogeneous half-spaces. A rigorous general proof of the symmetry of Eq. (7.1) was given in Ref. 25. The proof was based on algebraic transformations of arbitrary tridiagonal matrices. They describe the spectral problem in the case of an arbitrary step-function distribution with any number of steps. An arbitrary continuous profile can be approximated to any degree of accuracy using step functions. Another proof of Eq. (7.1) in the continuous case has been provided in Ref. 26, p. 297. In Ref. 26 relations also were obtained connecting the initial

velocity v with the transformed velocity $I\{v\}$. A new and short proof of the symmetry of Eq. (7.1), which reveals the property of duality between the eigenfunctions of the Lagrangian pressure and velocity, has been given in Ref. 21. A number of other recent papers have exploited the symmetry of Eq. (7.1) (Ref. 27, p. 526 and Ref. 28, Sec. IV).

The transformation (7.1) is not trivial. It changes the profile qualitatively, transforming one function $\rho(\eta)$ into another. It has been shown[21] that when Eq. (7.1) is inverted, the eigenfunctions are transformed according to the rule

$$\tilde{v}(\eta)=p_L(-\eta), \quad \tilde{p}_L(\eta)=v(-\eta). \tag{7.2}$$

It is clear that the functions v and p_L are transformed and the sign of the argument is changed.

For the $f^{\pm}$ modes there is no perturbation of the pressure: $p_L(\eta)\equiv 0$. We shall denote these modes that are invariant with respect to the profile as $f_G^{\pm}$ modes. The velocity is found from any of the equations of system (5.1′) for $p_L\equiv 0$. They are $v^{\pm}=e^{\pm\eta}$. If we transform them according to the rule (7.2), we find that

$$(p_L)^{\pm}(\eta)=e^{\mp\eta}, \quad \omega^2=\pm gk. \tag{7.3}$$

Here and in what follows the upper and lower signs correspond to the upper and lower signs in the dispersion law in the form it is written in Eq. (7.3) . We shall call the solutions of (7.3) the $f_P^{\pm}$ modes. These same solutions are found from Eq. (2.1) after substituting into it the dispersion laws (7.3). If we operate using Eq. (2.1), then it is easy to see that the modes $f_P^{\pm}$ are invariant. This means that the distributions p_L (7.3) do not depend on the equilibrium profiles of the density and pressure.

7.2 Velocity Profile and the Boundary Conditions

Let us determine the $v^{\pm}(\eta)$ functions of the $f_P^{\pm}$ modes. It is not possible to determine them from the rules (7.2). It is necessary to use Eq. (6.2). In the general case it is equal to

$$(v^{\pm})'_{\eta}\mp v^{\pm}=\left(1\mp\frac{g}{kc^2}\right)\frac{e^{\mp\eta}}{\rho}. \tag{7.4}$$

At this point we can find these functions.

It may be noted that the general solution to Eq. (7.4) depends on two constants, C_P and C_G. One of the constants is related to the function p_L, since $(p_L)^{\pm}=C_P e^{\mp\eta}$, and the other is related to a first-order differential equation (7.4). The general solution of the homogeneous equation corresponding to Eq. (7.4),

$$(v^{\pm})'_{\eta}\mp v^{\pm}=0,$$

is equal to $C_G e^{\pm\eta}$. This means that the general solution of Eq. (7.4) is a mixture of $f_P^{\pm}$ and $f_G^{\pm}$ modes, taken with the weights C_P and C_G, respectively.

Thus in the general case the form of the $f_P^{\pm}$ mode is described. Now it is a matter of interest to determine whether one of them matches the physical boundary conditions. The isobaric conditions must be discarded because the functions $p_L^{\pm}=e^{\mp\eta}$ do not have any zeros. Therefore we shall investigate a layer bounded by two walls. Here we require solutions of Eq. (7.4) with two zeros.

The case when the right-hand side R does not change sign was studied in Section 6. This includes the case of an incompressible fluid. In Section 6 it was shown that the solutions of Eq. (7.4), taken with the lower sign, are monotonic and cannot have more than one zero. Consequently it remains to examine the case of the upper sign. Let us make a substitution similar to that in Section 6. We obtain

$$w'_y=k\left(1-\frac{g}{kc^2}\right)\frac{e^{-2ky}}{\rho}=kR(y), \quad v^{+}(y)=e^{ky}w(y). \tag{7.5}$$

We consider an arbitrary monotonically increasing or decreasing function $c(y)$, defined on the interval $y_D<y<y_U$, where y_D and y_U are, respectively, the lower and the upper rigid boundaries. It is necessary to study those values of k for which the zeros of the function $R(y)$ (see Eq. (7.5)) lie within this segment. We set $R(y_z)=0$, and for definiteness assume that the function $c(y)$ falls off with increasing y. Then for $y<y_z$ we find that $R(y)>0$. We shall integrate Eq. (7.5) from the point $y=y_D$, and set $w(y_D)=0$. Then for the velocity we also have $v(y_D)=0$ The solution $w(y)$ of Eq. (7.5) increases for $y_D<y<y_z$, it reaches a maximum at $y=y_z$ and then begins to fall off.

Variations in k cause variations in y_z. With decreasing k the point y_z moves towards the lower boundary y_D. If the point y_z is close enough to the boundary y_D, then the function $w(y)$ after going over the maximum decreases again and becomes zero at the point y_{zz} inside the section $y_z<y_{zz}<y_U$. It is understood that if now k increases, then the point y_{zz} moves upwards towards the upper boundary y_U. There exists a value k_P for which the equality $y_{zz}=y_U$ is satisfied. This wave number corresponds to the frequency $\omega_P=\sqrt{gk_P}$. For these values of k and ω the velocity function $v(y)$ satisfies both the upper and the lower boundary conditions and is an eigenmode. It corresponds to the point $I(k_P,\omega_P)$ on the (k,ω) plane.

7.3 Inverse Transition Between Acoustic and Gravitational Branches

In Section 7.2 we considered a layer bounded by rigid walls. The related spectrum is shown in Fig. 5a. There exist p and g modes. For the acoustic modes p_m $(m=0,1,\ldots)$ the index m denotes the number of zeros of the function $v(y)$ inside the layer. For $k\to 0$ their frequencies tend to constant values $\simeq c/d$, and for $k\to\infty$ they approach the asymptote $\omega=ck$ (here $d=y_U-y_D$ is the distance between the boundaries and c is the characteristic speed of sound).

There exists yet another acoustic mode with a different asymptotic limit at zero. It is indicated by the letter L in Fig. 5a, and is called a Lamb mode. The origin of this mode can best be understood by considering the limit $g\to 0$. It is clear that in a gaseous layer a mode exists that propagates precisely horizontally. Let us consider a rectangular cell with hard walls. The acoustic modes are classified by a pair of numbers (m_y,m_x), where m_y denotes the number of half-waves that fit into the cell in the y direction, and m_x denotes

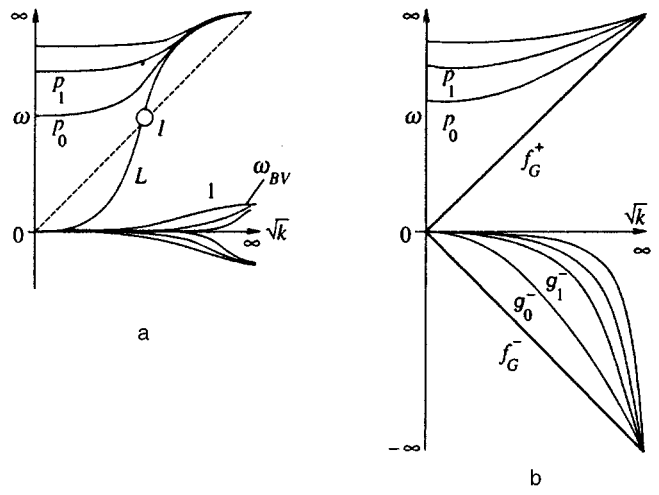

FIG. 5. a — The case of a gaseous layer between two rigid boundaries. The family of acoustic modes $(L,p_0,p_1,\ldots)$ and $(g^{\pm})$ gravitational modes (curve *1*). The dashed lines give the dispersion law $\omega=\sqrt{gk}$. The point *I* indicates the invariant acoustic mode f_p^+. b — The case of a layer between two isobaric boundaries (the pressure is $p_V=0$ on the upper boundary).

the same in the x direction. There are modes (0,1), (1,0), (1,1) ... . The modes $p_0,p_1,\ldots$ correspond to $m_y=1,2,\ldots$, and the L mode corresponds to $m_y=0$.

In a homogeneous layer with $g=0$ the vertical component of the velocity v vanishes identically, $v(y)\equiv 0$, in the case of the L mode. We note, incidentally, that in the inhomogeneous case the form of the function $v(y)$ is essentially defined by the inhomogeneity. It may not have zeros inside the layer in the case of a monotonic profile, or it may have one or several zeros depending on the nature of the nonmonotonicity. The L mode vanishes if even one of the boundaries is isobaric (see, e. g., Fig. 5b). For the present purposes it is important that for $g \neq 0$ the qualitative structure of the acoustic L and p_m modes is retained.

For $g \neq 0$ gravitational modes $g_m^{\pm}$ arise, where $m=0, 1, \ldots$ denotes the number of zeros of the function $v(y)$. In the limit $k\to 0$ they have the asymptotic behavior of a shallow-water wave

$$\omega_m=\xi_m\sqrt{\frac{d\ \ln s}{dy}}\ d\ ck=\alpha\xi_m\sqrt{\frac{d\ \ln s}{dy}}\ d\sqrt{gh}\ k,$$

where $h=c^2/g$, the numbers ξ_m and α depend on the specific profile, and $\xi\to 0$ as $m\to\infty$.

The point I, which refers to the invariant mode f_P^+, is located on the intersection of the dispersion curve of the L mode and the curve $\omega=\sqrt{gk}$. Its coordinates are k_P and ω_P (see Sec. 7.2).

For the sake of comparison, Fig. 5b also shows the complete spectrum of polytropes investigated here, bounded by two isobaric boundaries. There are p_m, f_G^+, g_m^-, and f_G^- modes. The index m denotes the number of zeros of the function $p_L(y)$. The asymptotes for $k\to 0$ remain in the same qualitative relation as for the case of two rigid boundaries. The asymptotes for $k\to\infty$ are $\sigma_m=\sqrt{gk}\Sigma_m$ for the g modes and $\omega_m=\sqrt{gk}\Omega_m$ for the p modes. The f_G^+ and g_m^- modes are related to the upper boundary y_V, and the f_G^- mode to the lower boundary y_A. Therefore, they are spatially separated.

As mentioned above, different eigenfunctions correspond to the invariant fundamental $f_P^{\pm}$ and $f_G^{\pm}$ modes for the former the pressure p_L is exponential, whereas for the latter the velocity v is exponential. The $f_G^{\pm}$ modes are gravitational. This follows from the fact that in the incompressible limit $c\to\infty$ their frequencies remain finite, whereas the frequencies of the acoustic waves go to infinity. Moreover, they have a gravitational asymptote. For example, in the case of a layer bounded below by a rigid boundary and above by an isobaric boundary, the mode with the asymptote of the fundamental mode, $\omega\to\sqrt{gk}$ as $k\to\infty$, joins the family of gravitational modes as $k\to 0$. For $k\to 0$ they approach the asymptote of the shallow-water mode. This mode is the hardest in this family. Conversely, the mode f_P^+ should be grouped with the acoustic family. In fact, it is situated on the acoustic L mode and coincides with it as $c\to\infty$, along with all the rest of the acoustic modes. Therefore the complete gravitational-acoustic classification scheme has the form

$$p,\ f_P^+,\ f_G^{\pm},\ g^{\pm}.$$

8. TURBULENT MIXING OF POLYTROPIC PROFILES

The analysis in the previous sections has given some exact results pertaining to the linear spectral theory. This theory is applicable as long as the amplitudes of the perturbation are small:

$$t\ll\frac{1}{\sigma(k)}\ \ln\frac{1}{ka_0},$$

where a_0 is the initial amplitude. Let us consider the motion in the nonlinear stage

$$t\sim\frac{1}{\sigma(k)}\ \ln\frac{1}{ka_0}$$

and the highly nonlinear (turbulent) stage

$$t\gg\frac{1}{\sigma(k)}\ \ln\frac{1}{ka_0}.$$

It is found that definite conclusions can be drawn about the turbulence from the spectral theory. The most important of these are the following.

A. In the case of polytropes bounded by the vacuum $(p_V=0)$ the growth rate σ increases without bound as $k\to\infty$ (see Sec. 2, and Fig. 5b, the g_0^- mode). This indicates preferential mixing of small-scale perturbations over long-wavelength perturbations in the early stages of turbulent mixing

B. The perturbations are localized in a near-surface layer of thickness $\sim 1/k$ at the vacuum boundary (see Section 5, which deals with an analysis of the fields of the eigenfunctions). This means first, that the mixing starts at this boundary, and second, the mixed layer is adjacent to it.

C. A periodic perturbation in the layer develops in two stages. The first is a stage of exponential growth, and the second is the saturation stage. Let us estimate the typical velocity in the second stage by the formula

$$v \sim \sigma(k)/k. \tag{8.1}$$

The development of a periodic perturbation causes ''smearing'' of the power-law singularities. The smearing causes averaging of the density profile over the horizontal coordinate, as for profile *1* in Fig. 4a. The thickness h_P of the smeared layer increases with time (here the letter P stands for ''periodic''). The maximum rate of expansion

$$\frac{dh_P}{dt} \sim \Sigma_0 \sqrt{\frac{g}{k}} \tag{8.2}$$

is attained when the system makes the transition from the first to the second stage. Then the velocity Eq. (8.1) decreases. This decrease occurs because the parameter kh_P becomes larger than unity, and the growth rate in the estimate (8.1) begins to decrease, since it is no longer calculated from the formula $\sigma=\Sigma_0\sqrt{gk}$ but from the formula $\sigma=\sigma_{BV} \sim \sqrt{g/h_P}$. The reason is the transition from the SS asymptote to the BV asymptote (see Fig. 3c in Sec. 4).

D. Let us compare the development of small-scale and large-scale periodic perturbations. According to point (A) enumerated above, small-scale perturbations are dominant in the early stage. However, the expansion velocity related to the large-scale phenomena, Eq. (8.2), is greater. Therefore in the later stages the large-scale perturbations dominate.

E. Self-similar growth rates (Sec. 2) and self-similar eigenfunctions (Sec. 5) are associated with a uniform (power-law) distribution. The self-similarity or uniformity means in this context that all the space and time scales are of equal weight. This then implies equal weight of all the pairs of ''small-large'' scales (these are denoted by the letters λ_S and λ_L, respectively) that differ by a fixed factor $(\lambda_L/\lambda_S=Q)$.

From the discussion in point (D) regarding the change in the dominant scales that occurs in time, and the discussion in point (E) regarding the self-similarity, it follows that during the development of turbulence there is a kind of relay sequence of instantaneous dominant horizontal scales $\langle\lambda\rangle$. It may be called a cascade of increasing instantaneous scales. This is an inverse cascade, since in this process the characteristic wave number k decreases. Correspondingly, the Fourier transforms of the gasdynamic functions such as ρ are self-similar. Let us designate by $\rho(x,y,t)$ and $\rho(k,y,t)$ the density distribution and its Fourier transform. The self-similarity of the Fourier transform means that it can be written as a function of two variables, $\rho(q,\eta)$, where $q=kgt^2$ and $\eta=y/gt^2$. The dependence on q for fixed η has a maximum corresponding to the dominant scale. We use the notation $q=q_{\max}$ to indicate the value of q at the maximum, $2\pi/\alpha_H$.

Then the dominant horizontal scale is $\langle\lambda\rangle=2\pi/k_{\max}$, where $q_{\max}=k_{\max}gt^2$, is equal to $\langle\lambda\rangle=\alpha_H gt^2$.

In a similar way, we can infer from dimensional arguments (the acceleration g is the only dimensional parameter) that the thickness of the turbulent mixing layer is self-similar:

$$h_T=\alpha_V(N,\gamma)gt^2. \tag{8.3}$$

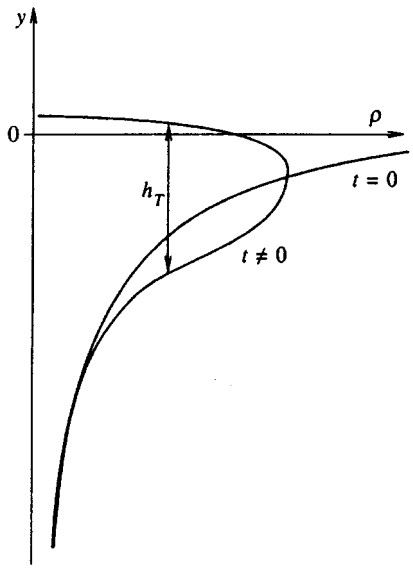

FIG. 6. The development of polytropic compressible turbulence. $h_T=\alpha_V gt^2$ is the width of the mixing layer.

The development of turbulence is shown in Fig. 6. During mixing, the power-law singularity is smoothed out by the turbulence. This is due to the elimination of the low-entropy Lagrangian particles and the increase of the high-entropy ones. It should be emphasized that the thickness h_T increases according to the self-similarity rule (8.3) in the compressible case. The turbulent mixing coefficient α_V depends on the exponents N and γ characterizing the compressible material. It decreases with increasing N and fixed γ.

This work was carried out with the financial support of the Russian Fund for Fundamental Research (Grant No. 95-02-06381-a) and INTAS (Grant No. 94-1105).

[1)]In the theory of atmospheric oscillations the situations in which an isobaric boundary is below a gaseous layer simply does not arise, and therefore there are no f^- modes.

[2)]Here an example is pertinent. The velocity of sound in the Earth's atmosphere is low compred to the earth's orbital motion. Accordingly, the Mach number calculated for the orbital velocity is very large, ≈ 100. This does not interfere with the atmosphere being in subsonic equilibrium.

[1] J. H. Nuckolls, L. Wood, X. Thiessen, and G. B. Zimmermann, Nature **239**, 139 (1972).
[2] K. A. Brueckner and S. Jorna, Rev. Mod. Phys. **46**, 325 (1974).
[3] A. M. Prokhorov, S. I. Aniskimov, and P. P. Pashinin, Usp. Fiz. Nauk **119**, 401 (1976) [Sov. Phys. Usp. **19**, 547 (1976)].
[4] S. I. Anisimov, A. M. Prokhorov, and V. E. Fortov, Usp. Fiz. Nauk **142**, 395 (1984) [Sov. Phys. Usp. **27**, 181 (1984).
[5] J. D. Lindl, R. L. McCrory, and M. Campbell, Phys. Today **45**(9), 32 (1992).
[6] S. Yu. Gus'kov, N. V. Zmitrenko, and V. B. Rozanov, Zh. Éksp. Teor. Fiz. **108**, 548 (1995) [JETP **81**, 296 (1995)].
[7] S. W. Haan, S. M. Pollaine, J. D. Lindl *et al.*, Phys. Plasmas **2**, 2480 (1995).
[8] R. McEachern, C. Moore, G. E. Overturf *et al.*, ''Inertial Confinement Fusion'' ICF Quarterly Report. LLNL **4**, 25 (1993).
[9] S. A. Letts, G. W. Collins, E. M. Fearon *et al.*, ''Inertial Confinement Fusion'' ICF Quarterly Report. LLNL **4**, 54 (1994).
[10] N. A. Inogamov, Dokl. Akad. Nauk SSSR **278**, 57 (1984) [Sov. Phys. Dokl. **29**, 714 (1984)].
[11] W. Unno, Y. Osaki, H. Ando, and H. Shibahashi, *Nonradial Oscillations of Stars* (University of Tokyo Press, Tokio, 1979).

[12] J. P. Cox, *Theory of Stellar Pulsation* (Princeton University Press, Princeton New Jersey, 1980).
[13] K. O. Mikaelian, Phys. Rev. Lett. **48**, 1365 (1982).
[14] R. N. Ardashova, S. I. Balabin, N. P. Voloshin, Yu. A. Kucherenko *et al.*, in Vopr. Atom. Nauki Tekh. Ser. Teor. Prikl. Fiz. No. 1, 20 (1988).
[15] H. Takabe, K. Mima, L. Montierth, and R. L. Morse, Phys. Fluids **28**, 3676 (1985).
[16] H. Lamb, *Hydrodynamics* (Dover Publications, New York, 1945), Ch. X.
[17] C. L. Pekeris, Phys. Rev. **73**, 145 (1948).
[18] A. Skumanich, Astrophys. J. **121**, 408 (1955).
[19] J. Christensen-Dalsgaard, Monthly Notes Royal Astronomical Soc. **190** 765 (1980).
[20] J. Christensen-Dalsgaard, D. O. Cough, and J. Toomre, Science **229**, 923 (1985).
[21] N. A. Inogamov, Zh. Éksp. Teor. Fiz. **110**, 559 (1996) [JETP **83**, 299 (1996)].
[22] L. J. Slater, in *Handbook of Mathematical Functions*, NBS Appl Math. Ser. No 55, edited by M. Abramowitz and I. A. Stegun, Washington DC, GPO (1964).
[23] *Mathematical Encyclopedia* [in Russian] edited by I. M. Vinogradov *et al.*, (Sov. Entsikopediya, Moscow, 1977), Vol. 1.
[24] K. O. Mikaelian, Phys. Rev. A **26**, 2140 (1982).
[25] N. A. Inogamov, in *Problems of Dynamics and Stable Plasmas* [in Russian] (Moscow Physicotechnical Institute (MFTI), Moscow, 1990), p. 100.
[26] H. J. Kull, Phys. Rep. **206**, 197 (1991).
[27] C. Cherfils and K. O. Mikaelian Phys. Fluids **8**, 522 (1996).
[28] K. O. Mikaelian, Phys. Rev. E **53**, 3551 (1996).

Translated by J. R. Anderson

Effect of an electric field on the surface tension of a liquid at low temperatures

L. A. Mel'nikovskiĭ and S. A. Kriminskiĭ

P. L. Kapitsa Institute of Physical Problems, Russian Academy of Sciences, 117334 Moscow, Russia
(Submitted 1 November 1996)
Zh. Éksp. Teor. Fiz. **111**, 1369–1372 (April 1997)

An external electric field changes the dispersion law of waves on the surface of a liquid. Besides the usual capillary term ($\propto k^3$, k is the wave number) and gravitational term ($\propto k$), a term quadratic in the wave vector appears in the expression for the square of the frequency in a homogeneous field. These excitations are associated with the variation of the coefficient of surface tension of the liquid at low temperatures. In the case of a large field tangent to the surface, the correction is proportional to $T^{8/3}$, unlike the $T^{7/3}$ correction in the absence of a field.
 [S1063-7761(97)01804-0]

The temperature dependence of the coefficient of surface tension close to absolute zero is associated with low-energy excitations of the liquid–gas boundary (surface waves).[1] The usual classification separates capillary and gravitational waves. The properties of the former are mainly determined by surface-tension forces, which are substantial in the limit of small wavelengths. The gravitational forces, on the other hand, correspond to long wavelengths. This paper discusses the spectral variation effect associated with an external electric field constant in time. It is substantial in the region intermediate (in wavelength) between the capillary and gravitational regimes.

Let us find the dispersion relation for a surface wave propagating in the presence of external field $\mathbf{E}_0$ (for definiteness, let $\mathbf{E}_0$ be the field outside the liquid). The x axis is in the wave-propagation direction, while the z axis is upward, perpendicular to the surface of the unperturbed liquid. The surface displacement from the equilibrium position in this wave is described by a function of the form

$$z=\zeta(x,t)=\zeta e^{ikx-i\omega t}.$$

Assuming that the liquid is incompressible, the continuity condition of Ref. 2 is imposed on the velocity $\mathbf{v}$ and takes the form

$$\operatorname{div}\mathbf{v}=0. \tag{1}$$

The Euler equation in the linear approximation (deviations from equilibrium are considered small) gives a second condition for the velocity:

$$\frac{\partial v_i}{\partial t}=\frac{1}{\rho}\frac{\partial \sigma_{ik}}{\partial x_k}, \tag{2}$$

where

$$\sigma_{ik}=-(\rho g z+P)\delta_{ik}-\frac{E^2}{8\pi}\left[\varepsilon-\rho\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T\right]\delta_{ik}+\frac{\varepsilon E_i E_k}{4\pi} \tag{3}$$

is the stress tensor.[3] Here P is the pressure corresponding to the same density ρ in the absence of a field, and ε is the permittivity of the liquid. The electric field itself obeys the following equations (all the velocities are much less than the velocity of light, and the liquid is assumed not to have any free charges):

$$\operatorname{div}\mathbf{E}=0, \quad \operatorname{curl}\mathbf{E}=0. \tag{4}$$

The corresponding field distribution, periodic along the x axis, has the form

$$\mathbf{E}^g=\mathbf{E}_0+\mathbf{A}e^{ikx-kz}, \quad \mathbf{E}^l=\mathbf{E}_0^l+\mathbf{B}e^{ikx+kz},$$

where the superscripts l and g refer to the liquid and gas, respectively, with

$$A_x=-iA_z=-iA, \quad B_x=iB_z=iB, \quad A_y=B_y=0.$$

The possibility of making the following transformation of the right-hand side of Eq. (2) by using Eq. (4) is essential for the subsequent treatment:

$$\frac{\partial v_i}{\partial t}=\frac{1}{\rho}\frac{\partial\sigma_{ik}}{\partial x_k}=\frac{1}{\rho}\frac{\partial}{\partial x_i}\left[-(\rho g z+P)+\frac{E^2}{8\pi}\rho\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T\right].$$

As a consequence, the motion of the liquid is irrotational; i.e., $\mathbf{v}=\nabla\psi$. Using Eq. (1), we get

$$\psi=\dot{z}\,\frac{e^{kz}}{k}=-\zeta\,\frac{i\omega}{k}\,e^{ikx+kz-i\omega t}.$$

Euler's equation itself in terms of the potential ψ takes the form

$$\frac{\partial\psi}{\partial t}=-\left(gz+\frac{P}{\rho}\right)+\frac{E^2}{8\pi}\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T. \tag{5}$$

At the interface, the electric field satisfies[3]

$$\mathbf{E}_\parallel^g=\mathbf{E}_\parallel^l, \quad \mathbf{E}_\perp^g=\varepsilon\mathbf{E}_\perp^l.$$

We can use the surface $z=0$ for the interface (we assume $k\zeta\ll 1$). After elementary transformations, we get

$$E_{0x}=E_{0x}^l, \quad E_{0y}=E_{0y}^l, \quad E_{0z}=\varepsilon E_{0z}^l,$$

$$A=k\zeta\,\frac{\varepsilon-1}{\varepsilon+1}(E_{0z}-iE_{0x}), \quad B=k\zeta\,\frac{\varepsilon-1}{\varepsilon+1}\left(\frac{E_{0z}}{\varepsilon}+iE_{0x}\right). \tag{6}$$

Using explicit expressions for σ_{ij}^g and σ_{ij}^l in the boundary conditions for the stress tensor,

$$\sigma_{ij}^g n_j=\sigma_{ij}^l n_j-\alpha\,\frac{\partial^2\zeta}{\partial x^2}\,n_i$$

(here **n** is the unit vector normal to the interface, while α is the coefficient of surface tension), we get the equation for the pressure at the boundary:

$$P=\text{const}=-\alpha\frac{\partial^2\zeta}{\partial x^2}+\frac{E^{l^2}\rho}{8\pi}\left(\frac{\partial\varepsilon}{\partial\rho}\right)_T-\frac{\varepsilon-1}{8\pi}(\varepsilon E_\perp^{l^2}+E_\parallel^{l^2}). \quad (7)$$

Finally, substituting Eq. (7) into Eq. (5) and using Eq. (6), we find

$$\omega^2=gk+\frac{\alpha}{\rho}k^3+\frac{(\varepsilon-1)^2}{4\pi\rho\varepsilon(\varepsilon+1)}(\varepsilon E_\parallel^2\cos^2\theta-E_\perp^2)k^2. \quad (8)$$

Above, for brevity, we have omitted the subscript 0 in the field and have introduced into the treatment the angle θ between the wave vector **k** and the projection of field **E** onto the horizontal plane. It is clear that the dispersion relation for surface waves in a liquid in a magnetic field is obtained from Eq. (8) by replacing the permittivity by the permeability. In fact, in the particular cases of tangential and normal fields, our result coincides with that of Ref. 4.

The resulting formula is inapplicable for large vertical fields. The unperturbed surface of the liquid in this case cannot be regarded as a horizontal plane. For such a surface to be stable, the quantity ω^2 must be positive for all k, since otherwise the amplitudes of the waves with the corresponding wave vectors will grow without limit. It is easy to obtain the required limitation on the field:

$$(\varepsilon-1)^4E_\perp^4<64\pi^2\rho\alpha g\varepsilon^2(\varepsilon+1)^2.$$

For example, in water, the density is $\rho=1$ g/cm^3, $\varepsilon=81$, $\alpha=73$ dyne/cm, and the critical field is $E=2.5$ kV/mm. In the limit $\varepsilon\to\infty$, corresponding to a conductor, the stability condition goes over to the inequality $E_\perp^4<64\pi^2\rho\alpha g$ (Ya. I. Frenkel', 1935; see, for example, Ref. 3, section 5).

The generalization of the dispersion relationship to a weakly inhomogeneous external field is obvious. If the characteristic scale l over which E_0^2 varies is much larger than the corresponding scale of the wave (ζ or $1/k$), the only thing that changes in Eq. (8) is the coefficient g (which, of course, will no longer mean the acceleration of gravity, but will characterize the force acting on unit mass of the material in the combined electric and gravitational fields). For example, in the geometry of a charged jet surrounded by a cylindrical layer of liquid with radius r,

$$g=\frac{\varepsilon-1}{4\pi\varepsilon}\frac{E_0^2}{\rho r}.$$

Here E_0 is the field at the surface, gravitation is considered small, and the applicability condition is $\zeta\ll r$.

In a strong horizontal field, when the capillary and gravitational terms can be neglected, $\omega\propto k|\cos\theta|$, and the group velocity is independent of **k** and is always parallel to the field. The surface thus becomes effectively rigid in the direction of the field.

The correction $\delta\alpha$ to the coefficient of surface tension (relative to the value at $T=0$) equals the derivative of the ''quasi-partial'' (associated with the surface waves) potential Ω with respect to the surface area:

$$\delta\alpha=\left(\frac{\partial\Omega}{\partial S}\right)_T,$$

for an ideal two-dimensional Bose gas, it is determined by[5]

$$\delta\alpha=T\int\ln\left[1-\exp\left(-\frac{\hbar\omega}{T}\right)\right]\frac{d^2k}{4\pi^2}.$$

Assuming that the variation of the spectrum associated with the field is small, we get

$$\begin{aligned}\delta\alpha&=-\frac{T}{8\pi^2}\int_0^{2\pi}d\theta\int_0^\infty\frac{k^2}{e^q-1}\,dq\\&\approx-\Gamma\left(\frac{7}{3}\right)\zeta\left(\frac{7}{3}\right)\frac{\rho^{2/3}T^{7/3}}{4\pi\alpha^{2/3}\hbar^{4/3}}\\&\quad+\Gamma\left(\frac{5}{3}\right)\zeta\left(\frac{5}{3}\right)\frac{(\varepsilon-1)^2\rho^{1/3}(\varepsilon E_\parallel^2-2E_\perp^2)T^{5/3}}{48\pi^2\varepsilon(\varepsilon+1)\alpha^{4/3}\hbar^{2/3}}\\&\approx-0.1341\,\frac{\rho^{2/3}T^{7/3}}{\alpha^{2/3}\hbar^{4/3}}\\&\quad+0.004\,\frac{(\varepsilon-1)^2\rho^{1/3}(\varepsilon E_\parallel^2-2E_\perp^2)T^{5/3}}{\varepsilon(\varepsilon+1)\alpha^{4/3}\hbar^{2/3}},\end{aligned}$$

where $q=\hbar\omega/T$. In the opposite case of a strong tangential field, however,

$$\delta\alpha\approx\sqrt[3]{\frac{T^8}{\alpha\hbar^5}}\sqrt{\frac{\varepsilon+1}{\pi^3}}\frac{2\rho^{5/6}}{(\varepsilon-1)E}\int_0^\infty dr\int_0^\infty ds\,\ln[1-\exp(-\sqrt{r^2+s^3})]\approx-0.59\,\frac{\rho^{5/6}T^{8/3}\sqrt{\varepsilon+1}}{\alpha^{1/3}\hbar^{5/3}(\varepsilon-1)E}.$$

We should point out that it is convenient to use waves propagating perpendicular to the field direction to experimentally measure the surface tension. Their dispersion relationship depends on the field only via the renormalized surface tension. Atkins's theory is actually applicable only to liquid helium and hydrogen. Since the former is only very weakly polarizable ($\varepsilon=1.047$), the effect under consideration apparently cannot be measured with present-day experimental technique. The susceptibility is much larger in hydrogen ($\varepsilon=1.231$), while a fairly low temperature can be attained. Thus, it is preferable to choose hydrogen for measuring the correction to the surface tension.

A stimulating discussion with A. F. Andreev, K. O. Keshishev, and A. Ya. Parshin was essential to the writing of this article. We would also like to thank Yu. A. Kosevich for pointing out Ref. 4.

[1] K. R. Atkins, Can. J. Phys. **31**, 1165 (1953).
[2] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Nauka, Moscow, 1988; Pergamon Press, Oxford, 1987).
[3] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Nauka, Moscow, 1992; Pergamon Press, Oxford, 1960).
[4] J. R. Melcher, *Fluid Coupled Surface Waves* (MIT Press, Cambridge, Mass., 1963).
[5] L. D. Landau and E. M. Lifshitz, *Statistical Physics* (Nauka, Moscow, 1995; Pergamon Press, Oxford, 1980), part 1.

Translated by W. J. Manthey

The effect of pulsed magnetic fields on Cz-Si crystals

M. N. Levin and B. A. Zon[a)]

Voronezh State University, 394693 Voronezh, Russia
(Submitted 30 December 1995; resubmitted 23 October 1996)
Zh. Éksp. Teor. Fiz. **111**, 1373–1397 (April 1997)

It has been established that short-term effects of pulsed magnetic fields initiate long-term low-temperature decay of a supersaturated solid solution of oxygen in silicon crystals grown by the Czochralski method (Cz-Si), which results in the generation of oxygen-containing defects in the form of O–V centers and more complicated $Si_xO_yV_z$ complexes, where V is a vacancy. The process of defect formation after the action of the pulsed magnetic fields culminates in the formation of spatially ordered oxygen–vacancy clusters and/or the precipitation of oxide phases, depending on the original defects present in the crystal. The action of such fields also initiates crystallization of an amorphized layer when it is present on the surface of the original crystal. The detected effects are characterized by a threshold field strength, are accumulate with successive pulses, and reach saturation in terms of these parameters of the action. The effects induced in Cz-Si crystals by pulsed magnetic fields are analyzed in terms of a possible cause consisting of the excitation of the Si–O bond of an interstitial oxygen by nonequilibrium population of the vibrational levels of a metastable electronic term of the bond.
 [S1063-7761(97)01904-5]

1. INTRODUCTION

It is well known that relatively weak magnetic fields (<1 T) can substantially affect the actual structure and physical properties of solids. Weak magnetic fields have been observed to have an effect on the mechanical properties of solids,[1–4] the luminescence of crystals,[5] the resonance absorption of ultrasound by dislocations,[6] and the optical absorption of impurity centers.[7] Detailed studies have been made of the magnetoplastic effect (the increase of the mobility of dislocations in a magnetic field), which is detected in the alkali halide crystals NaCl, LiF, and CsI and the nonmagnetic metals Al and Zn under the action of weak magnetic fields, both constant[8,9] and variable.[10,11] It has been observed that dislocations in NaCl crystals continue to have increased mobility after the magnetic action is removed (the magnetic memory effect).[12,13]

The influence of weak dc and ac magnetic fields on chemical reactions in condensed media[14–18] has been detected and actively studied.

It is assumed in most papers that weak magnetic fields affect materials because they remove the forbiddenness for any transitions in which the electron spin changes. For example, theoretical models of the effect of weak magnetic fields on radical chemical reactions are based on the assumption that the magnetic field removes the spin forbiddenness on intercombination transitions between states with different multiplicity.[5,14–16] Ideas about the removal of spin forbiddenness in intercombination transitions have made it possible to explain the main features of the magnetoplastic effect,[19–21] the effect of a magnetic field on dislocation friction, and the crystal-strengthening effect in a dc magnetic field.[22,23] In particular, the increased mobility of dislocations in a magnetic field was explained by the effect of the field on the probability of intercombination transitions between singlet and triplet states of radical pairs, which arises when unpaired nuclear spins of the dislocations interact with paramagnetic impurities (dislocation stoppers), and by an increase in the occupation of antibonding triplet states of radical pairs that possess reduced breaking energies.

Interest has recently arisen in studying the effect of pulsed magnetic fields (PMFs) on condensed media.

It has been experimentally established that short-term actions of weak PMFs initiate long-term changes of the structure and physical properties of a wide class of nonmagnetic materials. Long-term structural changes after the action of PMFs have been observed in alkali halide crystals,[24] chalcogenide semiconductors,[25] II–VI semiconductor crystals,[26] III–V semiconductor crystals,[27] Ge,[27] Si,[28,29] and $Si–SiO_2$ structures.[30–33]

The distinctive features of PMF-induced effects are that there is a delay in their appearance after the magnetic action ends and that the kinetics has a long-term nonmonotonic character. Thus, after a PMF acts on silicon crystals for several seconds, the processes of structural changes in these crystals last hundreds of hours at room temperature.[29,32]

Despite the large amount of accumulated experimental data, the phenomenological picture of PMF-induced effects is not yet complete, and existing model concepts have not adequately explained them.

The goal of this paper is to study PMF-induced effects in silicon crystals grown by the Czochralski method (Cz-Si) in order to establish the mechanism by which a PMF effects long-term structural changes and to construct a qualitative model of the phenomenon.

2. EXPERIMENTAL RESULTS

The PMF-induced effects were studied on crystals of dislocation-free semiconductor silicon with substantially different concentrations of dissolved oxygen. Silicon crystals with a low concentration of dissolved oxygen ($[O]<10^{16}$

cm^{-3}) were obtained by crucibleless zone melting (Zm-Si), and crystals with a high oxygen concentration ($[O] > 10^{18}$ cm^{-3}) were obtained by pulling from the melt by the Czochralski method (Cz-Si).[34] Samples with *n*-type conductivity and crystallographic orientation of the (111) planes were studied. The dopant (phosphorous) concentration was less than 5×10^{15} cm^{-3}. The crystals for the study were cut from the central part of the Cz-Si or Zm-Si bars. The sample-fabrication methods are specifically indicated below for each of the investigation methods used here.

PMFs were imposed as series of from one to 10^4 unipolar triangular pulses with similar widths of the leading and the trailing edges. The pulse amplitude H was varied from 10^2 to 10^6 A/m, while the pulsewidths τ were varied from 10^{-2} to 10^{-6} sec, with a fixed repetition frequency of 2×10^{-2} sec.

The magnetic processing was carried out at room temperature, except for the specifically mentioned cases in which it was done at liquid-nitrogen temperature. In the intervals between measurements, the PMF-processed samples were stored in an inert gas (a mixture of argon and nitrogen) at room temperature.

2.1. Optical microscopy

When a cleavage face of a silicon plate was visually observed in an MII-4 optical microscope before and after magnetic action, the formation of precipitates was detected in Cz-Si crystals subjected to PMF processing. Typical results are shown in Fig. 1. The size and concentration of the observed precipitates increased for hundreds of hours after the magnetic action ended. The precipitates that formed after the action of the PMF possessed a very nonequilibrium spatial distribution. They were mainly localized in the regions of the cleavage face with the most significant original structural damage. No precipitates formed on the control cleavage faces of Cz-Si crystals not subjected to PMF action during the entire storage time. Images of the corresponding control cleavage faces are shown in the insets in Fig. 1.

No formation of precipitates as a result of PMF action was detected on the cleavage faces of Zm-Si plates.

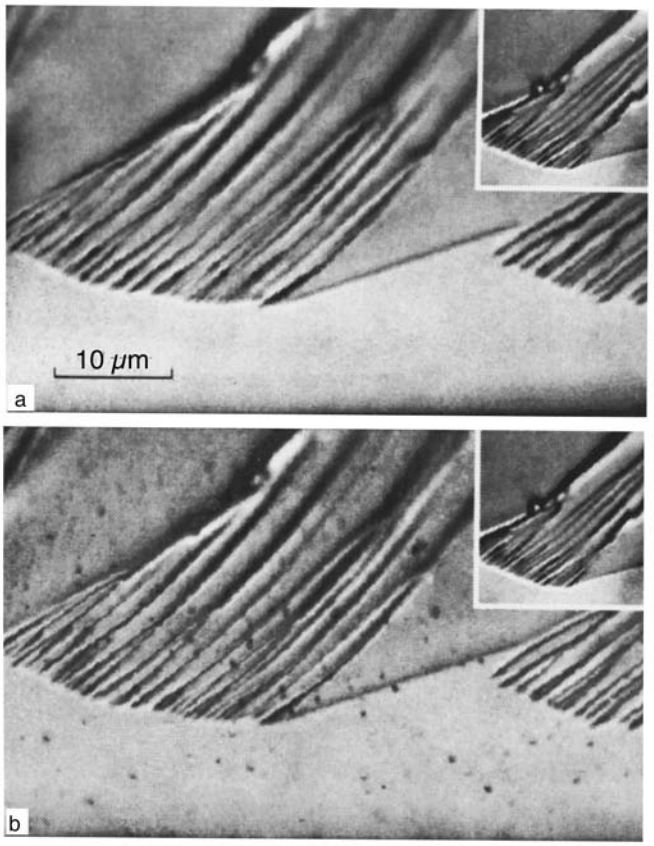

FIG. 1. Image of a cleavage face of a Cz-Si crystal in an optical microscope: (a) before the action of PMF; (b) 250 h after the action of PMF. The PMF regime is $H = 10^5$ A/m, $\tau = 2 \times 10^{-5}$ sec, $N = 10^3$ pulses. The insets show images of a cleavage face of a control sample that was not subjected to the action of PMF.

2.2. Transmission electron microscopy

PMF-induced structural changes in silicon crystals were studied in an ÉMV-100AK transmission electron microscope.

To prepare a sample, a disk 3 mm in diameter was cut from the central part of a silicon plate about 250 μm thick. An indentation was ground out at the center of the disk to a residual thickness of about 30 μm, after which the surface of the crystal was sputtered with a beam of Ar^+ ions with an energy of about 6 keV until a hole opened up. The thickness of the sample thus prepared decreased radially from the edge to the center.

Typical images of the precipitates that appeared in the Cz-Si crystals after the action of PMFs are shown in Fig. 2. No such defects were observed on the original samples and the control samples not subjected to the action of PMFs. A feature of the transmission images of PMF-induced precipitates on thin sections of the sample close to the perforation is the complex picture of the extinction contours accompanying them (Fig. 2a). As the storage time of the samples after magnetic action increased, the extinction contours around the precipitates became less pronounced. Extinction contours did not appear on the transmission electron micrographs of the precipitates made close to the outer edge of the sample, because of the large thickness of the crystal (Fig. 2b).

The observed extinction contours can be caused by elastic-stress fields that arise because of the wide disparity of the molar volumes of the oxide precipitates and the crystal matrix, and they can be made to disappear by removing these elastic stresses as a result of the gradual accumulation of vacancies in the regions where the crystal is compressed around the precipitates. As is well known, elastic stresses can be compensated by forming vacancy-type dislocation loops around the precipitates that create compressive stresses in the crystal.[34]

It should be pointed out that the ion sputtering of the crystal was nonequilibrium and created island-like relief of the sample surface on the side subjected to the sputtering. This was manifested on images in the direct beam of the transmission microscope by the presence of dark regions separated by lighter boundaries where the crystal was more deeply sputtered. Moreover, bombardment of the surface of the crystal with heavy ions caused its surface layer to be

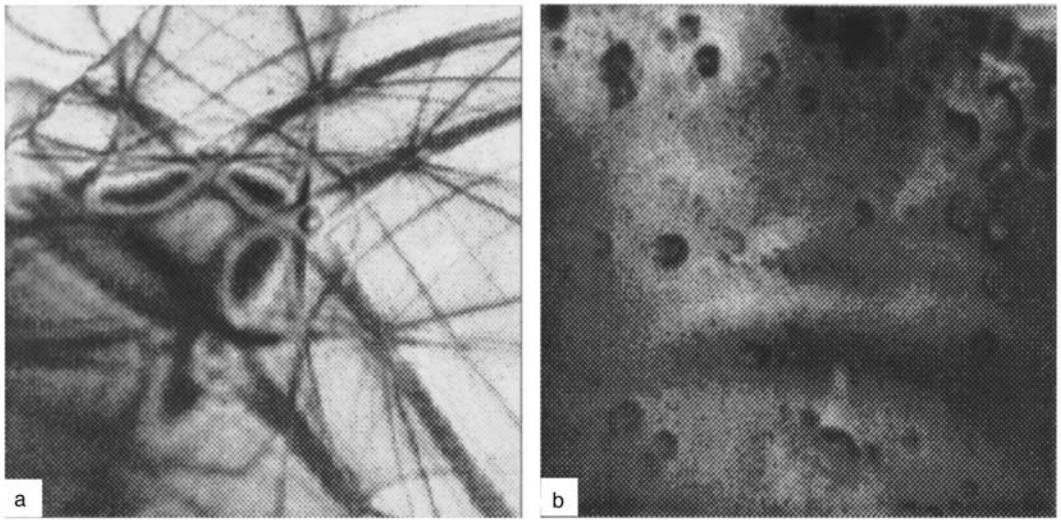

FIG. 2. Transmission electron micrographs of precipitates formed in a Cz-Si crystal after the action of a PMF: (a) on a thin section of the sample (75 h after PMF); (b) on a thick section of the sample (250 h after PMF). Magnification 12 000.

amorphized, and this resulted in a corresponding halo in the electron-diffraction pattern of the sample.

Complete images of a Cz-Si sample in the direct beam of the transmission electron microscope and its electron-diffraction patterns, obtained before and after the action of a PMF, are shown in Fig. 3.

Short-term action (a few seconds) of PMFs caused the halo on the electron-diffraction patterns of the Cz-Si crystals to be completely eliminated, showing that the amorphized surface of the sample was crystallized. The halo disappeared within tens of hours after the completion of the magnetic action. In control samples the original halo remained as long as the samples were kept. When the control samples were specially exposed in the electron microscope for several minutes under a 170-keV electron beam, the initial halo did not disappear.

These results show that the detected effect of crystallization of the amorphized phase of silicon results from the action of the PMF.

Another indication of PMF-induced crystallization of the amorphized surface was that the electron micrographs became more distinct because the diffuse background caused by electron scattering in the amorphized surface layer of the original sample was eliminated.

The action of PMFs on the Zm-Si samples did not sig-

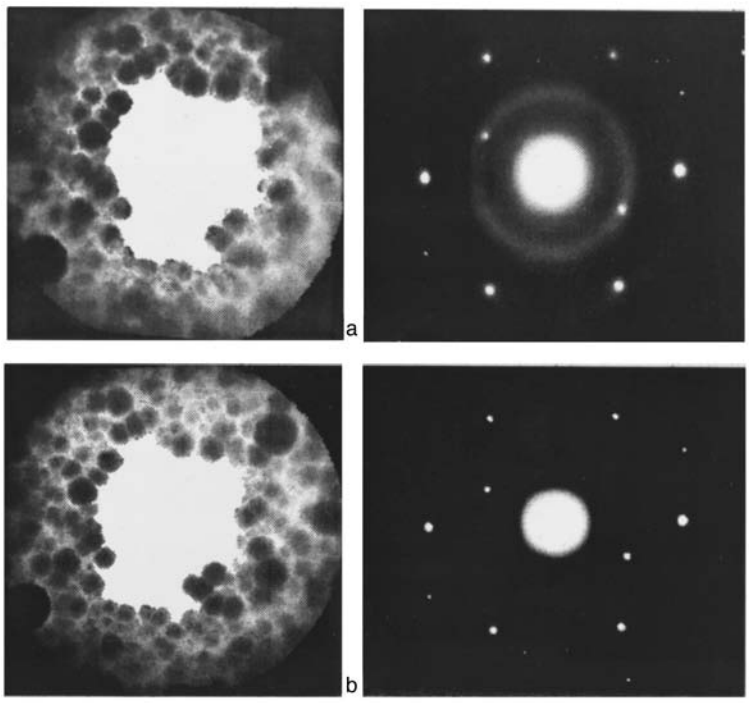

FIG. 3. Images of a Cz-Si sample in the direct beam of a transmission electron microscope and its electron diffraction pattern: (a) before the action of a PMF, (b) 150 h after the action of a PMF. Sample diameter 3 mm.

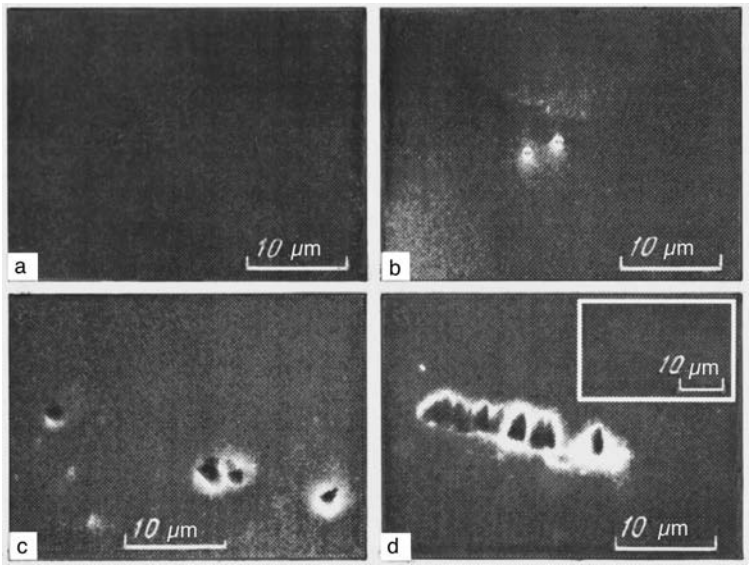

FIG. 4. Images in a scanning electron microscope of a selectively etched surface of a Cz-Si crystal: (a) before the action of a PMF; (b), (c), and (d) 20, 75, and 150 h after the action of a PMF, respectively. The inset to (d) shows the selectively etched surface of a control sample.

nificantly change the brightness of the halo of the original electron-diffraction pattern.

2.3. Scanning electron microscopy

The effect of PMFs on silicon crystals was studied by scanning electron microscopy with a CamScan microscope as follows: A Cz-Si plate was cleaved into several samples, one of which was subjected to the action of PMFs, while the others served as control samples. The surface topology of the samples was studied in the secondary-electron recording regime. To reveal structural defects, the surface of the samples was subjected to selective chemical etching. All the samples were processed by PMFs together, and each sample was etched individually but in identical regimes for different time intervals after the magnetic action, immediately before monitoring the surface topology.

The main results obtained by this method for the action of PMFs on Cz-Si crystals are shown in Fig. 4. Selective etching did not reveal any surface defects on the original samples (Fig. 4a). However, etch pits were detected after magnetic processing, and their depth increased with the time interval that passed after the end of the action of the PMFs (Fig. 4b–4d). These pits, detected when the crystal surface was selectively etched, had no crystallographic faceting inside them characteristic of the emergence of linear dislocations onto the surface. The absence of faceting is evidence that the detected etch pits are caused by spatially limited defects that are etched more rapidly than is the defect-free crystal. Such defects may be the precipitates observed in the optical and transmission electron microscopes.

No etch pits appeared on the control samples stored along with the samples processed by PMFs (inset in Fig. 4d).

2.4. Scanning tunnelling microscopy

Studies in a scanning tunnelling microscope of the surface topology of Cz-Si crystals with layer-by-layer etching showed that PMF-induced effects appeared very different in the near-surface layers and in the depth of the crystal.

The action of a PMF caused isolated etch pits to appear in the near-surface layers of the crystal (Fig. 5a). As with the scanning-electron studies, the absence of faceting of the detected etch pits showed that their appearance is caused by spatially limited defects. Such defects can be oxide precipitates visually similar to those observed in an optical microscope on cleavage faces of a Cz-Si crystal. In the near-surface layers, the process of PMF-induced precipitation is also facilitated by an increase in the concentration of the original structural disturbances that can serve as primary nucleation centers.

Note that, along with oxygen-containing defects, other point defects can be formed in the surface layers of the Cz-Si crystals—for example, clusters of interstitial atoms that form during the growth of precipitates of oxide phases.

In the bulk of the Cz-Si crystal, with higher structural perfection, the action of PMFs resulted in the formation of regular structures of nanometer scale (Fig. 5b).

The formation of such structures was detected only in Cz-Si crystals and was not observed in Zm-Si crystals, which is evidence that oxygen plays a key role in the appearance of the effect.

The regular structure may result from PMF-induced formation of clusters of the type $Si_xO_yV_z$, where V is a vacancy.

2.5. IR transmission spectroscopy

The IR transmission spectra in the region of wavenumbers ν from 400 to 1200 cm^{-1} were recorded before and after the action of PMFs on a UR-brand IR spectrometer at room temperature. The samples were of a Cz-Si plate polished on two sides.

The long-term variations of the IR transmission spectra of the Cz-Si crystals after magnetic action directly showed that dissolved oxygen participated in the PMF-induced processes. It is known that the main band ($\nu = 1106$ cm^{-1}) in the spectral interval in Fig. 6a is caused by IR absorption at Si–O bonds of interstitial oxygen.[35] The permanent distor-

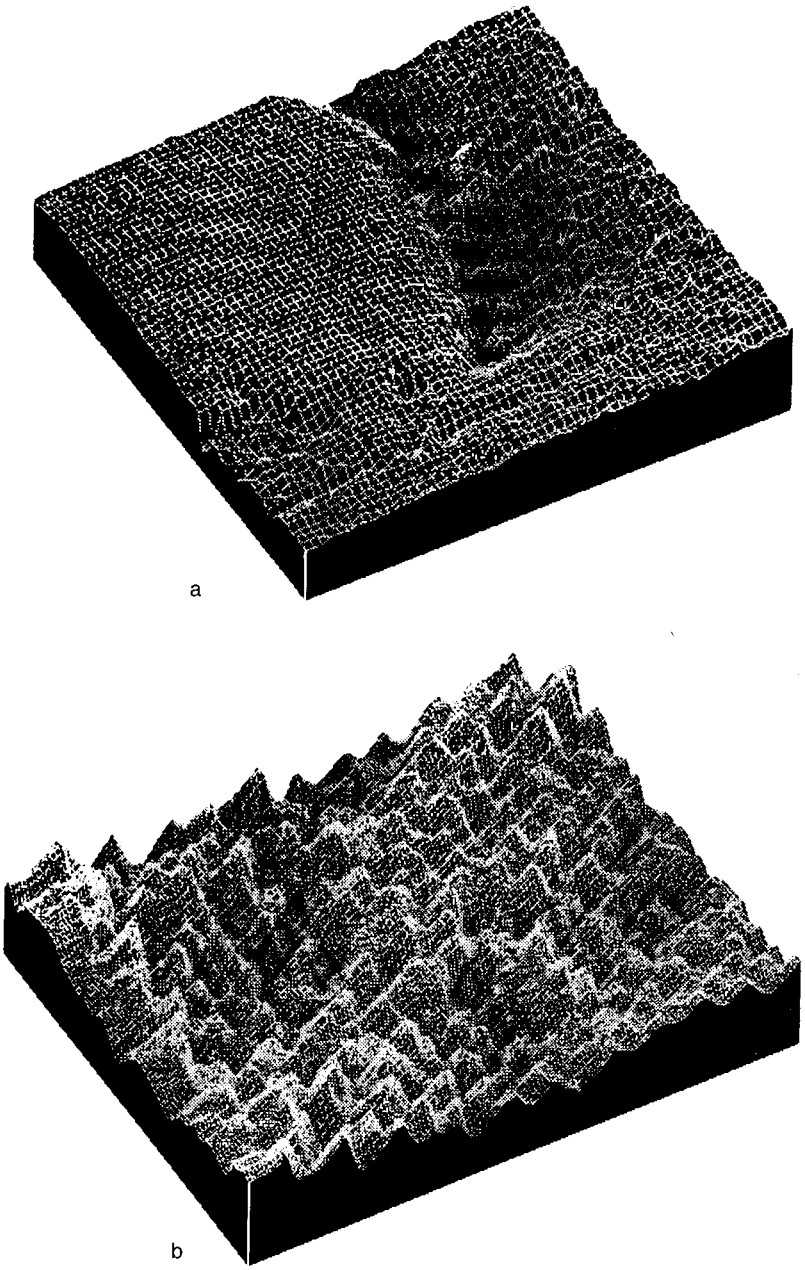

FIG. 5. Topology of a selectively etched surface of a Cz-Si crystal 250 h after the action of a PMF. The thickness of the etched layer is about 2 μm (a) and about 20 μm (b). The image size is 300 nm along the x axis, 300 nm along the y axis, and 45 nm along the z axis.

tions of this band as a result of the action of PMFs are evidence that the silicon–oxygen bond is excited by the magnetic field.

Another sign that oxygen participates in PMF-induced effects was the appearance of a narrow absorption band ($\nu = 830$ cm^{-1}). The presence of this band in the IR spectrum is associated with absorption at an oxygen–vacancy (O–V) complex, well known as the radiation point defect called the A center.[36–38]

The IR absorption also increases in the wavenumber interval 400–650 cm^{-1} (Fig. 6b). The appearance of absorption bands in this part of the spectrum was observed earlier after long (tens and hundreds of hours) heat treatments in the temperature interval 600–800 K and was associated with the formation of low-temperature thermodonors.[39] The nature of the thermodonors is still being discussed in the literature, but it seems certain that they are oxygen-containing SiO_4 complexes or have a more complicated composition.[34,35,40,41]

2.6. Deep-level transient spectroscopy

The possibility of PMF-induced formation of point defects with deep levels in the band gap of a semiconductor has been tested by deep-level transient spectroscopy.[42] The studies were carried out in the constant-capacitance regime[43] on Au–Si Schottky barriers formed on Cz-Si crystals. Polished plates of KÉF-7.5 silicon were used for the samples. Rectifying metal–semiconductor contacts were obtained by thermally depositing gold electrodes 0.8 mm in diameter and about 500 nm thick through a molybdenum mask in a vacuum better than 10^{-3} Pa with no special heat treatment. An ohmic contact was created by mechanically depositing an indium–gallium eutectic mixture on the back side of the sample.

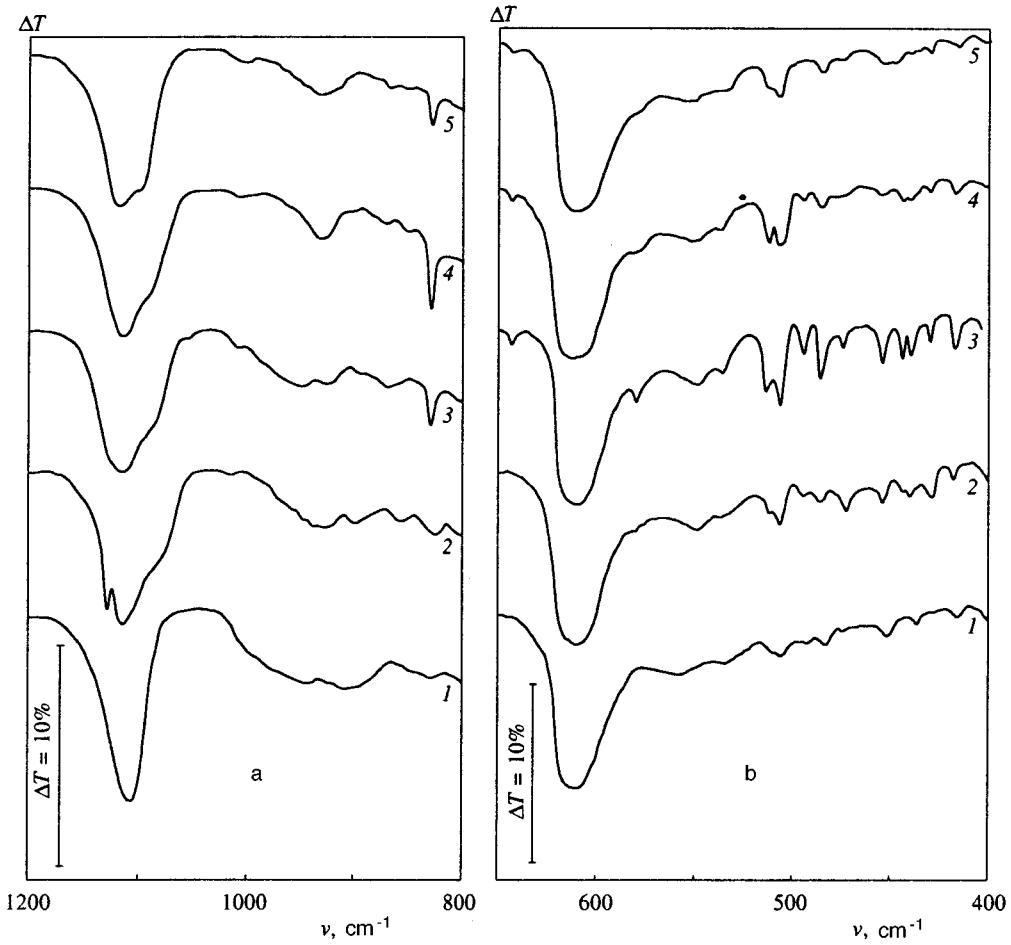

FIG. 6. IR transmission spectra ΔT of a Cz-Si crystal in the wavenumber intervals 800–1200 cm^{-1} (a) and 400–650 cm^{-1} (b); *1*—before the action of a PMF; *2*, *3*, *4*, and *5*—25, 150, 250, and 600 h after the action of a PMF, respectively.

On the original samples (before magnetic action) as well as on companion samples periodically monitored in the course of hundreds of hours after fabrication, the concentration of any deep levels that contribute to the spectrum was less than 10^{13} cm^{-3}. The action of PMFs caused the appearance of a signal whose amplitude substantially increased for several days after the completion of the magnetic action and then decreased (Fig. 7).

A characteristic of the detected peak was that its shape was distorted on the high-temperature side. Such distortion is evidence that the observed signal could be caused by a deep-level band.[44] In fact, the relaxation of the nonequilibrium charge of the band of deep levels forming the overall signal peak results in a nonexponential variation of the recorded parameter (the voltage on the electrode in the method used here, in which a constant capacitance is maintained on the Schottky barrier). In this case, it becomes incorrect to use Lang's classical technique[42] for determining the energy position of the levels and their majority-carrier capture cross sections.[44]

It is well known[45] that *A* centers in silicon give a peak in the signal spectrum obtained by deep-level transient spectroscopy, with a maximum at a temperature $T_m \sim 80$ K for a relaxation time of $\tau_m \sim 10$ μs, where $\tau_m \sim (t_2 - t_1)/\ln(t_1/t_2)$, while t_1 and t_2 are the variable times of recording the relaxation parameter. It is also known[36] that *A* centers belong to

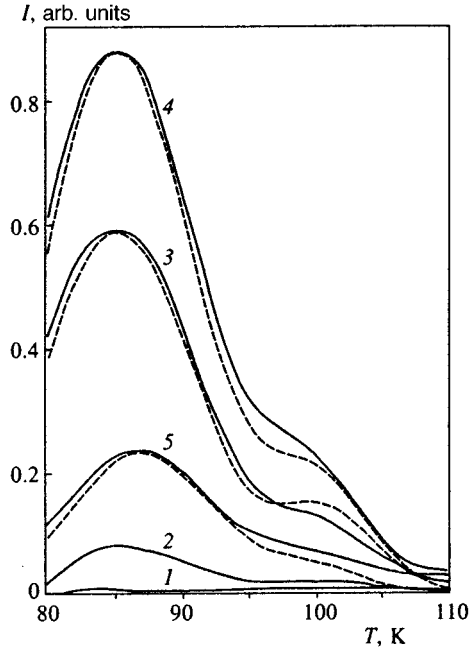

FIG. 7. Signal obtained by deep-level transient spectroscopy of a Cz-Si crystal: *1*—before the action of a PMF; *2*, *3*, *4*, and *5*—25, 150, 250, and 600 h, respectively, after the magnetic action. The solid curves show the experimental dependences, and the dashed curves show the results of a numerical calculation with the parameters $E_0 = E_c - E_t = 0.17$ eV, $\sigma_n = 5 \times 10^{-15}$ cm^2, $t_1 = 2 \times 10^{-3}$ sec, and $t_2 = 10^{-2}$ sec and with a compressive hydrostatic pressure of $P_0 = 0.32$, 0.34, and 0.27 GPa for curves, *2*, *3*, *4*, respectively.

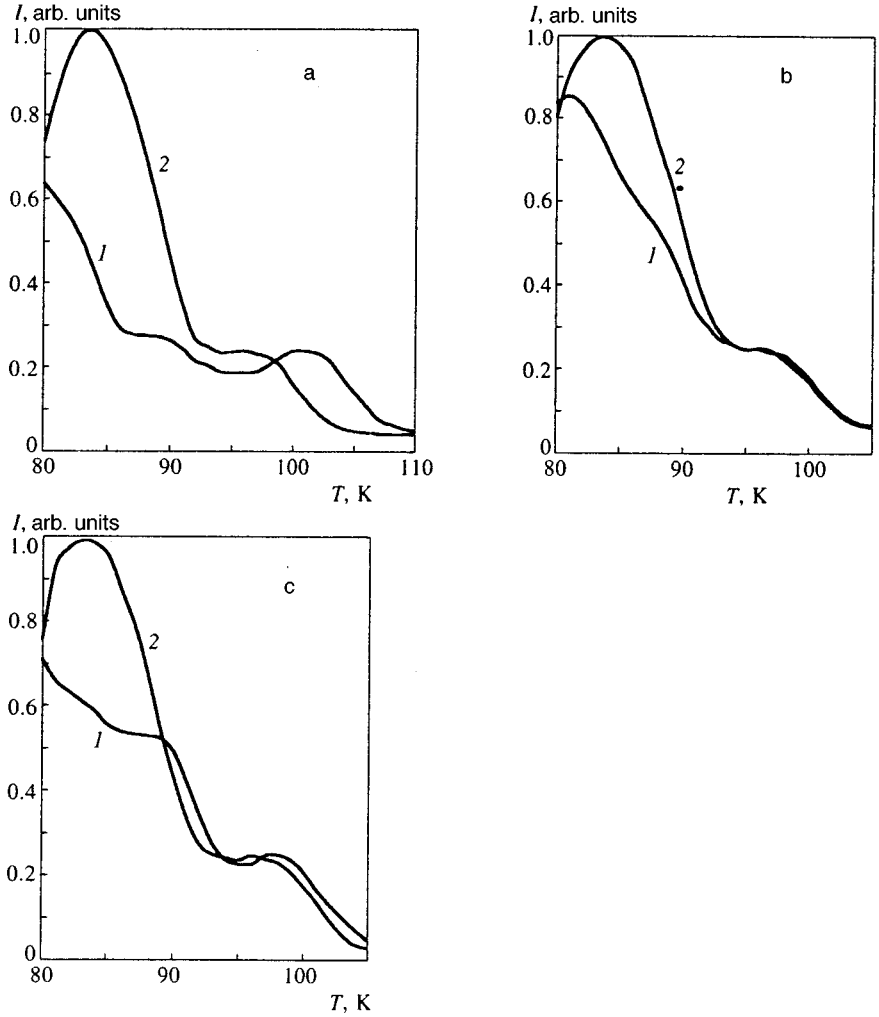

FIG. 8. Effect of uniaxial compression (curves *1*) on the signal obtained by deep-level transient spectroscopy of a PMF-induced center in a Cz-Si crystal: (a) $P_{[100]}=0.3$ GPa, (b) $P_{[110]}=0.3$ GPa, and (c) $P_{[111]}=0.3$ GPa. The signal-recording regime is $t_1=8\times10^{-3}$ sec, $t_2=10^{-2}$ sec; curves *2*—initial signal with no external compression.

the C_{2v} symmetry group and, in the absence of external effects, are sixfold degenerate.

Using the parameters of the A center (an energy position in the band gap of $E_c-E_t=0.17$ eV, where E_c is the energy level in the conduction band and E_t is the energy of the deep level, and an electron-capture cross section of $\sigma_n=5\times10^{-15}$ cm^2, Refs. 34 and 35) and taking into account elastic hydrostatic-compression fields, a numerical calculation of the signal makes it possible to accurately reproduce the experimental signals of the PMF-induced defect. Figure 7 shows the experimental and calculated signals. The signal observed after the action of PMFs shows features characteristic of a center of C_{2v} symmetry in a cubic crystal.

To conclusively determine the symmetry group of the PMF-induced defect, the effect of uniaxial compression of the crystal was studied by the method proposed in Ref. 46. The samples specially fabricated for this method were $5\times1.5\times1.5$ mm^3 parallelepipeds cut from the central part of a Cz-Si sample. To optimize the signal resolution, the measurements were made with close-lying values of the parameters t_1 and t_2.

Structures with Schottky barriers and ohmic contacts were formed on the opposite side faces of the samples. The long sides of the samples were oriented along one of the crystallographic axes [100], [110], and [111]. Uniaxial compression along these directions was created by means of a device similar to that described in Ref. 36 and an additional lever whose free end was loaded with the test sample.

The results of the action of uniaxial compression are shown in Fig. 8. When the sample was compressed along the [100] axis, the recorded signal invariably showed a second maximum (Fig. 8a). The signal was shifted toward lower temperatures relative to the initial value, its amplitude was reduced, and an inflection point appeared on the high-temperature side at a height close to the amplitude of the second maximum. Substantially smaller distortions of the original signal and reduction of its height were observed when the sample was compressed along the [110] axis (Fig. 8b). The application of external pressure in the [111] direction significantly reduced the amplitude of the initial signal and caused the appearance of a second maximum and an inflection point at a height twice that of the second maximum (Fig. 8c).

These effects of uniaxial compression on the recorded signal confirms that the PMF-induced point defect belongs to the C_{2v} symmetry group.

Since there are no data on the electronic–vibrational characteristics of the detected center caused by the Jahn–Teller effect nor on the character of the reconstruction of the partially filled broken silicon bonds, the PMF-induced defect

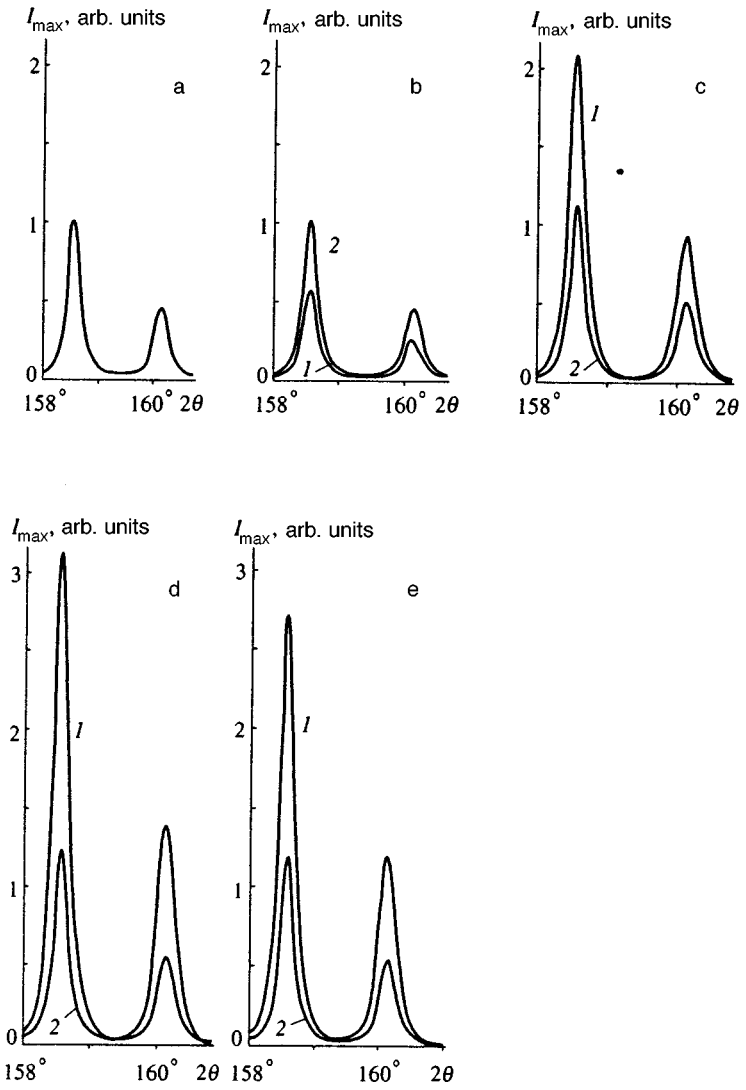

FIG. 9. Diffraction patterns of a Cz-Si crystal (111): (a) before the action of a PMF; (b), (c), (d), (e) 25, 150, 250, and 600 h after the action of a PMF, respectively. Diffraction patterns *1* and *2* correspond to a PMF-processed sample and a control sample.

cannot be unconditionally identified as an *A* center. However, the fact that the detected defect coincides in energy position, electron-capture cross section, and symmetry type with the corresponding characteristics of an *A* center, in our opinion, justifies the use of the term ''*A*-like center'' to designate the detected PMF-induced point defect.

2.7. X-ray diffractometry

X-ray structural studies of the effect of PMFs on silicon crystals were carried out on a DRON4-07 diffractometer. The diffraction patterns were measured with CuK_α x-rays ($\lambda = 1.54051$ Å), in the regime in which the sample is automatically angularly displaced with a step of 0.01° and an exposure time of 3 sec at each point. The doublet character of the K_α x-ray line caused the x-ray diffraction lines to be split into two peaks.

The samples were polished Cz-Si and Zm-Si plates with the crystallographic orientation of the (111) planes.

The brief action of PMFs produced long-term nonmonotonic variations of the interplanar distances d in the Cz-Si crystals and also substantially changed the total and maximum intensities of the x-ray diffraction lines of these crystals. Typical changes of the diffraction patterns of the Cz-Si crystals after the action of PMFs are shown in Fig. 9. For comparison, the same figure shows the results of measurements on control samples that were not subjected to magnetic processing.

The detected intensity changes of the x-ray diffraction lines of Cz-Si samples after the action of PMFs can be caused by the formation of the quasiregular spatial structures observed in the scanning tunnelling microscope. Deformation of the atomic planes of the original Cz-Si crystal as a result of PMF-induced formation of oxygen-containing $Si_xO_yV_z$ clusters should cause the primary extinction during dynamic x-ray scattering to be suppressed and indeed results in a substantial increase of the total and maximum intensity of the x-ray diffraction lines (making them several times as strong), while their ratio, which characterizes the half-width of the line, is maintained.

No appreciable changes of the x-ray diffraction line intensities were observed in the Zm-Si crystals after the action of PMFs.

The relative changes $\Delta d/d$ of the interplanar distances of Cz-Si crystals after PMFs act on them reached values of

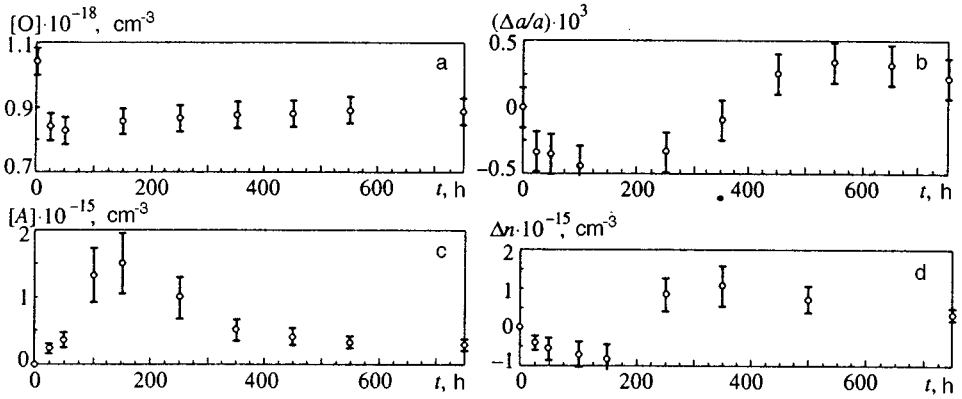

FIG. 10. Variations of the parameters of a Cz-Si crystal vs the time after the action of a PMF: (a) interstitial oxygen concentration (from the IR transmission at about 9 μm); (b) relative variation of the lattice constant (from the shift of the x-ray diffraction line); (c) concentration of A centers (from the area of the signal obtained by deep-level transient spectroscopy); (d) free-electron concentration.

about 5×10^{-5} for the (444) planes, while this parameter did not vary by more than 1.5×10^{-5} on the control samples. The corresponding temporal variations of the lattice constant a of the crystal ($a=d\sqrt{h^2+k^2+l^2}$, where h, k, and l are the crystallographic indices) are shown in Fig. 10.

We should also point out that, in $Si-SiO_2$ structures obtained by thermal oxidation of Cz-Si plates, PMF-induced changes of the x-ray diffraction lines and of the interplanar silicon spacings were significantly weaker than on unoxidized plates.

The temporal variations of the interplanar distances in Zm-Si crystals after the action of PMFs was less than the variation of this parameter on the control samples.

Thus, the results of x-ray structural studies confirm that oxygen dissolved in the silicon lattice plays a key role in the appearance of PMF-induced effects and indicate that the surface state of the crystal is very significant in these effects.

2.8. Time dependences of the variation of the parameters

The main features of the kinetics of the long-term variations of the parameters of the Cz-Si crystals, observed after the short-term action of PMFs, are shown in Fig. 10. Besides the time dependence of the PMF-induced change of the lattice constant of the Cz-Si crystal mentioned above, Fig. 10 shows the time dependence of the concentration of interstitial oxygen, PMF-induced A-like centers, and conduction electrons.

The concentration of interstitial oxygen dissolved in the silicon lattice was determined by comparing the IR transmission coefficients of the test crystal at a wavelength of 9 μm ($\nu=1106$ cm^{-1}) with a reference crystal whose oxygen concentration was known. The concentration of A-like centers was estimated from the area of the signal obtained by deep-level transient spectroscopy. The conduction-electron concentration was determined from the ratio of the minimum and maximum values of the steady-state high-frequency capacitance of poly-$Si-SiO_2-Si$ structures specially formed on part of the test plates of silicon. The errors noted in Fig. 10 and the subsequent figures are caused by statistical scatter of the measured results on batches of at least twenty samples. Note that the time variations of the parameters of control samples not subjected to the action of PMFs were less than the statistical scatter of the corresponding parameters.

A comparison of the time dependence of the PMF-induced changes of the parameters of a Cz-Si crystal is evidence that the concentration of interstitial oxygen and the lattice constant of the crystal both decrease at the initial stage of the evolution. In other words, magnetic action causes part of the oxygen atoms to leave the interstitial state and causes internal elastic compressive stresses to appear in the crystal.

The subsequent increase of the lattice constant can be caused by the arrival of additional vacancies from the surface of the crystal, compensating the compressive stresses that appeared earlier.

The entry into the crystal of vacancies from the surface can increase the concentration of A-like centers formed by the interaction of these vacancies with interstitial oxygen. The gradual decrease of the concentration of PMF-induced A-like centers at the later stages of the evolution is determined by two factors: the consumption of the A-like centers in cluster-formation processes and the reduction of the flux of vacancies from the surface as the elastic stresses in the crystal are compensated.

The nonmonotonic change of the concentration of free carriers reflects the multistage process of the conversion of electrically neutral interstitial oxygen into electrically active defects, including A-like acceptor centers, thermal donors, and other oxygen-containing centers. However, this cause for the variation of the free-carrier concentration is not the only one possible. This parameter can also vary as a result of PMF-induced redistribution of the dopants. Such a redistribution effect has been observed in II–VI semiconductor crystals[26] and is not ruled out in Cz-Si crystals.

2.9. The effect of the regimes of action of PMFs

The experimental data presented above show that the long-term changes of the structural and electrophysical parameters of the Cz-Si crystals that occur after the short-term action of PMFs are associated with the change of state of part of the interstitial oxygen atoms.

To explain the mechanism of PMF-induced excitation of the silicon–oxygen bond of interstitial oxygen in the silicon lattice, studies were carried out to establish how the variation of the magnetically sensitive parameters of the Cz-Si crystals depends on the regimes of the magnetic action. The relative variation of the IR transmission at $\lambda=9\,\mu$m and the concen-

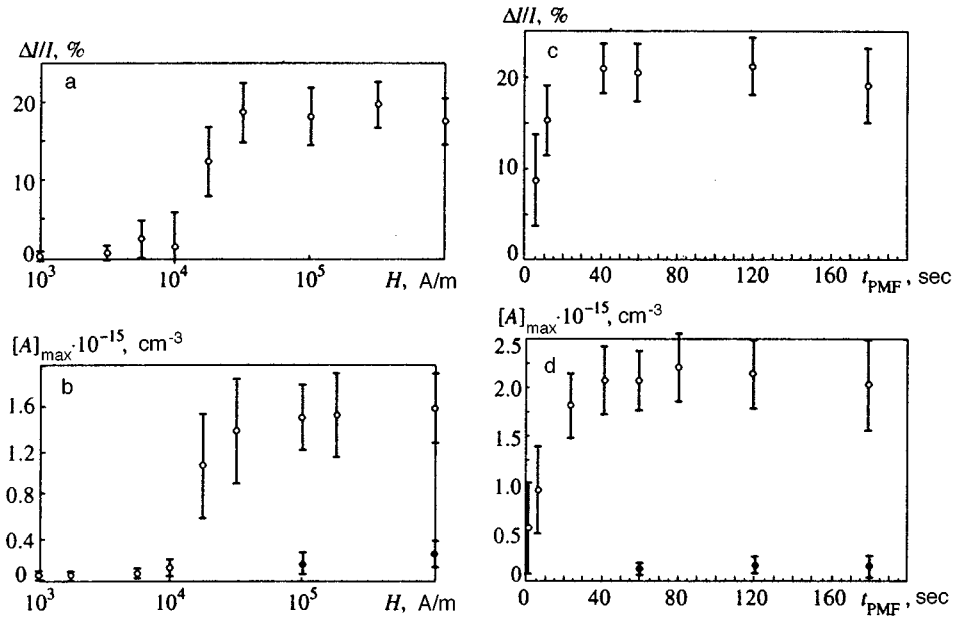

FIG. 11. Maximum variation of the parameters of a Cz-Si crystal vs the amplitude H of the magnetic-field pulses (a) and (b) and vs time t_{PMF} of the action of a PMF with frequency 50 Hz (c), (d); (a) and (c) are the relative change of the IR transmission at about 9 μm ($\nu=1106$ cm^{-1}); (b) and (d) are the concentrations of A centers. The open dots are for processing by a PMF at $T=300$ K; the closed dots are for $T=77$ K.

tration of A-like centers were chosen as the test parameters, reflecting the state of the oxygen in the crystal.

The variable parameters of the action were the amplitude, the pulsewidth, and the number of pulses of the magnetic field. The magnetic action was implemented at room temperature and liquid-nitrogen temperature.

Figures 11a and 11b show the maximum values of the changes of the test parameters (in the entire period of their evolution after the action of a PMF) as a function of the magnetic-field pulse amplitude. These dependences are evidence that the appearance of PMF-induced effects has a threshold character, with saturation in the amplitude of the magnetic-field pulses. The orientation of the samples in the magnetic field had no appreciable effect on the results of the action of the PMFs.

Figures 11c and 11d show the dependences of the variation of the test parameters on the time of action of the PMFs (at a pulse-repetition rate of 50 Hz). The effectiveness of the magnetic action increased with the number of pulses and reached saturation.

The variation of the IR absorption coefficient at the Si–O–Si bond of interstitial oxygen vs the magnetic-field pulsewidth is shown in Fig. 12. No appreciable effect of this parameter on the results of the magnetic action was observed over a wide pulsewidth variation. However, for pulsewidths comparable to their repetition interval, the efficiency of the magnetic action began to decrease.

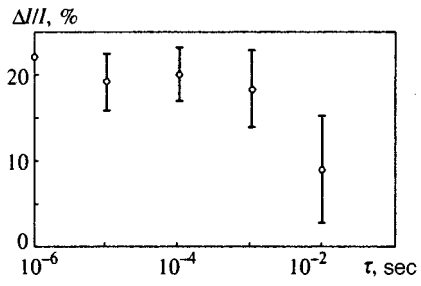

FIG. 12. Relative variation of the IR transmission of a Cz-Si crystal at about 9 μm ($\nu=1106$ cm^{-1}) vs the magnetic-field pulsewidth τ.

Unlike a PMF, a constant magnetic field with a value of 10^6 A/m and a duration of about 15 min caused no changes of the test parameters.

PMF-induced effects were also substantially suppressed if the magnetic processing was carried out at reduced temperature ($T=77$ K). Experimental data confirming this are shown in Fig. 11.

3. QUALITATIVE THEORY OF THE PHENOMENON

Let us evaluate the possible mechanisms for the action of PMFs on Cz-Si crystals. First, it should be pointed out that the energy of the carriers associated with the Zeeman effect changes by less than $10^{-3}kT$ under our conditions. Second, magnetostriction effects are also extremely small and moreover depend on the magnetic field ($\propto H^2$), which does not agree with the experimental results.

Finally, as pointed out in Sec. 2.9, the absence of a dependence of the observed effects on the pulsewidth of the magnetic pulses in fairly wide limits makes it impossible to explain these effects by induction fields.

The induction electric field E_{ind} can be estimated from

$$E_{\mathrm{ind}} \simeq \mu\mu_0 r H \tau^{-1},$$

where μ is the permeability of silicon, μ_0 is the permeability of free space, r is the maximum radius of the contour at which an induction emf appears (which is smaller than the linear dimensions of the sample; i.e., $r<1$ cm in our experiments), H is the amplitude of the magnetic pulses, and τ is the width of the triangular pulses, with identical rise and fall times of the field.

For the regime typically used for the action of PMFs ($H=10^5$ A/m; $\tau=2\times10^{-5}$ sec), the induction fields are less than 50 V/cm, which is clearly inadequate, for example, to

appreciably heat the carriers or to lower any internal potential barriers associated, in particular, with the Stark effect.

Since the detected effects cannot be explained in terms of the mechanisms enumerated above, we propose a new qualitative model for PMF-induced phenomena.

A characteristic of PMF-induced effects is that there is a significant delay for the manifestation of the magnetic action on silicon crystals, reaching tens and hundreds of hours. It should be pointed out that such a delay is not unique and is encountered in many areas of solid-state physics. The phenomenon of phosphorescence is a well-known example.[47] The generally accepted model of phosphorescence associates this phenomenon with the de-excitation of metastable excitations of a crystal that are formed by cascade transitions from optically allowed states populated by radiation incident on the crystal.

Another phenomenon related to the model under consideration is magnetic quenching of the fluorescence of gases and molecular crystals, which was studied comparatively recently in a series of papers (see, for example, the reviews in Refs. 5 and 16). These papers reported that the delayed luminescence of materials disappeared when they were placed in a constant magnetic field. The explanation of the effect is based on concepts concerning the removal by the magnetic field of the forbiddenness of the decay of metastable states.

After these preliminary remarks, let us pass on to an explanation of the qualitative theory of the phenomena that we observed.

As follows from the experimental data, the effect of a magnetic field on Cz-Si crystals is largely determined by the presence of dissolved oxygen in the silicon. It is well known that oxygen in silicon mainly occupies a bridge position, in which its two valence bonds are distributed between two silicon atoms of the lattice.[34] The equilibrium position of the oxygen atom in this case is shifted relative to the straight line connecting the silicon atoms, so that the angle α between the oxygen–silicon valence bond and this straight line is about 20°.[48]

We assume that the electronic excitation of this valence bond is metastable, with the dependence of the electronic terms on a certain coordinate, for example, the angle α mentioned above, being qualitatively shown by Fig. 13a.[1)] Such excited states, not optically coupled with the ground state, are well known in free molecules. A classical example is the triplet state in the hydrogen molecule. In this case, however, since α varies in a limited interval, several vibrational levels can exist in the metastable state.

A fundamental point, shown in Fig. 13a, is the assumption that there is a point where the ground and metastable terms intersect. When a magnetic field is present, the degeneracy of the levels at the point of intersection is removed, and anti-intersection of the terms appears, as shown in Fig. 13b. Such a transformation of the terms is completely analogous to the transition from diabatic to adiabatic terms in the theory of vapor-phase chemical reactions, which occurs when the finite collision rate of the atoms or molecules[49] is taken into account.[2)]

It is evident that, under the conditions of such a transformation, it becomes impossible to speak of vibrational levels belonging to the ground or the metastable terms if the energies of these levels lie in the region of the ''window'' between the terms that appears because of the magnetic field. The vibrational levels inside this window are a superposition of the vibrational levels of the ground and the metastable states. Therefore, if, before the magnetic field was turned on, the system was on a vibrational level of the ground term lying in the region of the window (level E_g in Fig. 13a), then, when the field is turned on, this level is smeared over the doublet levels of the superposition (E_{gm} in Fig. 13b). After the field is switched off, the system has an appreciable probability of being in a vibrational state of the metastable term (E_m in Fig. 13a). Thus, the metastable level is populated by a pulsed magnetic field.

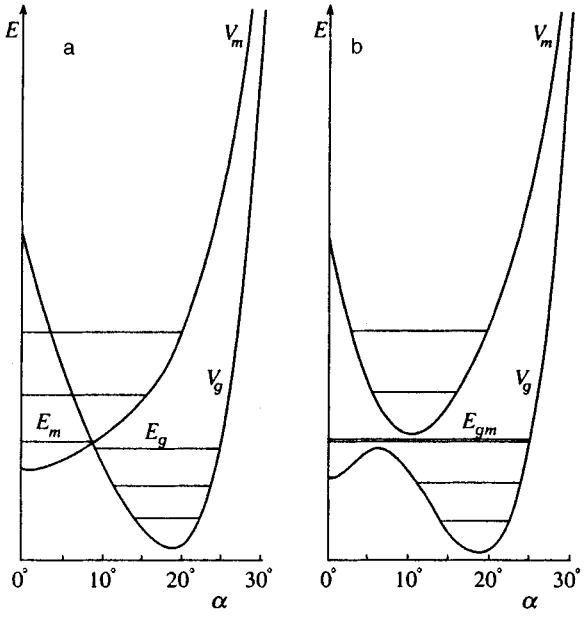

FIG. 13. Qualitative form of the electronic ground term (V_g) and metastable term (V_m) of the Si–O bond of interstitial oxygen in a silicon lattice: (a) with no magnetic field, (b) in a magnetic field. The horizontal lines conventionally show the positions of the vibrational levels in each term. The interacting levels are designated E_g and E_m. When the ''window'' is opened, these levels form the doublet E_{gm}.

For the subsequent evolution of the system, it is essential for there to be a long time interval in which there is no magnetic field. During this time, the metastability of the excited electronic term is restored, and its upper vibrational levels are populated according to Boltzmann's law. It should be taken into account in this case that the lower vibrational levels of the ground and metastable terms have different energies, and therefore the populated vibrational levels in the metastable state have on the average an absolute energy that is greater than the mean absolute energy of the vibrational levels in the ground state. The difference of these energies is determined by just the energy difference of the lower vibrational levels of the two terms, which means, in other words, by the activation of the Si–O bond.

The picture described here by no means contradicts the principles of thermodynamics, as it might seem to at first glance. The very fact that the mean energies of the ground and metastable degrees of freedom of the Si–O bond are

different reflects the property of metastability of one of them. Of course, after some time all the degrees of freedom will have equal mean energies; however, this time can be extremely long for the metastable state.

If the nonequilibrium-populated vibrational levels of the metastable term are able to partially relax during the pause after the magnetic-field pulse, for example, by tunnelling from an upper vibrational level through the potential barrier along any generalized coordinate, the Si–O bond will become excited, and this excitation will not be destroyed by the next magnetic-field pulse. In this case, periodic repetition of the pulses causes the excitation of the silicon–oxygen bond to accumulate.

The proposed model thus reduces to the following: The magnetic field opens a window between the ground and metastable states for a short time, as a result of which the metastable state is populated. After the window is closed, thermal processes cause the upper vibrational levels of the metastable state to be excited. As a result, the energy for the excitation and consequently also for the chemical activation of the Si–O bond is drawn not from the magnetic field but from the phonon thermostat (the crystal lattice). In this sense, the activation of the Si–O chemical bond can be regarded as endothermic. The emptying of the upper vibrational levels of the metastable term, for example, by tunnelling (for brevity, we shall call this the nonthermal channel) ensures that the excited state is maintained until the next magnetic-field pulse and that this excitation will grow (be ''pumped'') as the pulses are periodically repeated.

This description explains in a natural way why the observed effects are substantially suppressed if the crystals are processed by PMFs at reduced temperature, and also why the effect is absent in a constant magnetic field.

The excited-state pumping effect of the silicon–oxygen bond by a pulsed magnetic field can be achieved only if the nonthermal emptying of the upper vibrational levels of the metastable term during the action of the PMF occurs in a time that is shorter than the interval between the magnetic pulses, which, we recall, was 2×10^{-2} sec in our experiment. The population time can be estimated from the halfwidth of a spectral line of the vibrational level (1–100 cm^{-1}) and is about 10^{-10}–10^{-12} sec.

The proposed model makes it possible to explain the threshold dependence of the appearance of PMF-induced effects on the magnetic-field amplitude. The metastable state actually becomes populated as soon as the energies of the vibrational levels from both terms fall into the region of the magnetic window. Increasing the amplitude further does not strengthen the magnetic action, so that the field dependence of PMF-induced effects should have the characteristic form of curves with saturation, which is observed in experiment.

The threshold field H_0 is determined by the condition

$$\mu_B H_0 \simeq |E_g - E_m|,$$

where μ_B is the Bohr magneton.

When the magnetic field is $H\sim3\times10^4$ A/m, the window is about 0.02 cm^{-1} wide. This value can be compared with the characteristic spacing between the vibrational levels in the ground electronic term, which is several tens of inverse centimeters.[35] Making the natural assumption that the characteristic spacing between the vibrational levels of the metastable term is no greater than this value, the resulting estimate can be regarded as not contradicting the experimental measurement of the thermal processes. Moreover, it should be taken into account that the vibrational levels are strongly broadened by thermal processes. And, although the fact of broadening does not in itself fundamentally change the picture of the phenomenon,[3)] nonetheless the quantitative estimates given here can give a close approximation to the experimental data.

Based on the proposed model, the PMF-induced effects in Cz-Si crystals presented above can be interpreted as follows: The excitation of the Si–O bond by a pulsed magnetic field increases the chemical activity of the interstitial oxygen. As a result, part of the activated oxygen atoms can enter into interaction with the vacancies that exist in the crystal before the magnetic action. The fact that the original vacancies are present in the crystal not in isolated but in clustered form, for example, in the form of thermally stable point defects such as pentavacancies or the well-known *D* defects,[35] does not rule out the possibility of such interaction. In other words, part of the oxygen atoms excited by the action of PMFs pass from an interstitial position to the position of nearby vacant lattice sites (possibly with some shift from the position of the regular site similar to what takes place in the *A* center).

The decrease of the initial concentration of vacancies in the Cz-Si crystal because their positions are occupied by oxygen atoms that were at interstitial sites before the action of the PMFs causes elastic compressive stresses to appear in the crystal. The resulting deficiency of vacancies in the crystal can be compensated by the arrival of additional vacancies from its surface.

We should point out that the PMF-induced crystallization of the surface of the samples described in Sec. 2.2 and the lower sensitivity to the action of PMFs of Cz-Si crystals with thermal oxide on the surface noted in Sec. 2.7 also agrees with the assumption that nonequilibrium vacancies from the surface enter into a Cz-Si crystal processed by PMFs.

The vacancies that arrive from the surface and enter the volume of the crystal interact with the interstitial oxygen atoms, which also manifests itself in the long-term generation of *A*-like centers observed after the completion of the action of PMFs. A feature of the PMF-induced generation of *A*-like centers is that, unlike the radiational generation of ordinary *A* centers, it accompanies not an excess but a deficiency of vacancies. This feature causes PMF-induced *A*-like centers to form a band of energetically closely spaced deep levels because their degeneracy is removed by the internal elastic stress field, and this is shown by the splitting of the corresponding signal obtained by deep-level transient spectroscopy.

The transformation of interstitial oxygen atoms into the more mobile *A*-like centers, caused by the arrival of vacancies from the crystal surface, causes a multistage process of complex formation, which manifests itself in the formation of clusters of the type of $Si_xO_yV_z$ in the volume of the crystal and/or of oxide precipitates in the more defective surface

layers. The appearance and growth of clusters and precipitates, which are point sources of elastic stresses, increases the internal stress level in the crystal. This increase of the elastic stresses in the crystal in turn stimulates the arrival of vacancies from the surface. That is, positive feedback occurs between the generation of vacancies and elastic fields in the crystal and is maintained by the conversion of nonmobile interstitial oxygen into the more mobile *A*-like center when the oxygen captures a vacancy. This produces further growth of the clusters that creates local elastic stresses.

The process continues until the concentration of interstitial oxygen is decreased to levels such that the vacancies can reach the regions of elastic compression close to the precipitates without being captured at the interstitial oxygen atoms.

The subsequent arrival of vacancies is not accompanied by the formation of *A*-like centers and continues until the elastic stresses in the crystal are compensated and their thermodynamic equilibrium concentration is reestablished.

The indicated positive feedback should become stronger the larger the size of the PMF-induced precipitates and clusters. This makes it possible to understand why there is a time delay in the appearance of PMF-induced effects after the magnetic action ends. The latency of the PMF-induced defect formation at the initial stage is caused by the small size of the clusters that are formed (the sources of elastic stresses), for which the positive feedback that strengthens the appearance of the PMF-induced effects is still weak.

The proposed interpretation of the PMF-induced defect formation assumes that the mobility of the oxygen-containing complexes in silicon is rather high at room temperature. We know of no data on the mobility of such complexes when the elastic-stress fields are inhomogeneous, when nonequilibrium vacancies arrive from the surface, and when there is a high original concentration of interstitial oxygen. It is significant that each of the indicated factors is capable of stimulating the migration of oxygen in the crystal at low temperatures. For example, Ref. 52 found that, in Cz-Si crystals with an inhomogeneous distribution of elastic stresses created by carbon implantation (about 0.2 GPa, Ref. 53), the oxygen diffusion coefficient reaches anomalously high values (about 10^{-10} cm^2/sec at 720 K), where the normal value of this parameter is 10^{-20} cm^2/sec.

From an analysis of the splitting of the signal obtained by deep-level transient spectroscopy of the PMF-induced center (Fig. 7), the local compressive stresses in the crystal reach values of about 0.3 GPa, while the observed variations of the lattice constant of silicon (Fig. 10) correspond to the appearance of elastic stresses of about 0.1 GPa. These estimates coincide in order of magnitude with the data of Refs. 52 and 53 and consequently can result in equally significant changes of the diffusion coefficient.

Another factor that can be considered responsible for accelerated low-temperature diffusion of oxygen in silicon is that nonequilibrium intrinsic defects of the crystal can participate in its transport.[35] In our model, the PMF-induced migration of oxygen is counter to the flux of nonequilibrium vacancies from the surface, which agrees with this concept.

Finally, the ability of oxygen to migrate in silicon crystals increases as the original concentration of dissolved oxygen increases. Thus, when the oxygen concentration is about 10^{18} cm^{-3}, annealing of the radiation-induced *A* centers in a Cz-Si crystal occurs not at 300–400°C, as it is assumed to do, but in a much lower temperature interval ($\simeq$100 °C).[54] We recall that the Cz-Si crystals that we studied had a similarly high concentration of dissolved oxygen.

The combination of all three factors acting together apparently ensures that the capability of oxygen to migrate is anomalously high and produces the effects that we observed.

It should be pointed out that other mechanisms for the migration of oxygen-containing complexes are not ruled out—for example, exchange or cooperative mechanisms. However, the establishment of all possible mechanisms of PMF-induced oxygen transport requires further study and goes beyond the limits of this article. We should also point out that PMF-induced anomalous migration of an impurity in crystals at room temperature has been observed earlier. We recall in this connection the indium-redistribution effect, observed at room temperature in CdClTe crystals after the brief action of PMFs.[26]

The appearance of a flux of vacancies from the surface into the bulk of the crystal makes it possible to understand the effect of PMF-induced crystallization of an amorphized surface layer. It is well known that a flux of point defects into an amorphized region of a crystal can cause it to crystallize. Thus, in the phenomenon of low-temperature ion-induced crystallization, the crystallization of an amorphous layer is associated with diffusion to the phase interface of intrinsic point defects of the crystal, created by ion beams.[55]

There is also fundamental significance in the fact that PMF-induced excitation of Si–O bonds occurs coherently in the entire crystal. For a sufficiently high concentration of oxygen dissolved in the lattice of a Cz-Si crystal, this can result in the appearance of long-range order and can cause the observed formation of spatially ordered structures of nanometer scale.

4. CONCLUSIONS

1. The action of PMFs on Cz-Si crystals initiates the decay of a supersaturated solid solution of oxygen and manifests itself in the long-term generation of oxygen-containing defects, which culminate in the formation of spatially ordered clusters of the type $Si_xO_yV_z$ (V is a vacancy) in defect-free regions of the crystal and/or cause the precipitation of oxide phases in the near-surface layers having initial structural disturbances.

2. As the result of the action of PMFs, amorphized layers are crystallized when they are present on the surface of Cz-Si crystals.

3. PMF-induced effects have a threshold dependence on the PMF amplitude, accumulate as the number of pulses increases, and reach saturation in these parameters of the action.

4. A possible start-up mechanism for the appearance of PMF-induced effects in Cz-Si crystals is the excitation of the Si–O bond of interstitial oxygen by the population of the vibrational levels of its metastable term.

The authors are deeply grateful to M. I. Molotskiĭ, V. M. Maslovskiĭ, and V. S. Pershenkov for numerous fruitful dis-

cussions of the model concepts explained in this article, as well as to S. A. Ozol' for making the measurements with a scanning tunnelling microscope.

This work was carried out with the support of the Competitive Center of Fundamental Research in the area of the natural sciences, State Committee on Higher Education of the Russian Federation (Grant 94-7.17-346) and the Russian scientific program ''Noosfer.''

[a)]e mail: zon@niif.vucnit.voronezh.su

[1)]Of course, the configuration coordinates that determine the vibrational spectrum and the spin state of the Si–O bond can differ from each other. Moreover, the energetics of the bond are determined by the surface shape in multidimensional configuration–spin space, by its orientation relative to the crystallographic axes, etc. Therefore, the one-dimensional model presented here has a rough, qualitative character.

[2)]We should point out that the existing theory of vapor-phase reactions remains largely one-dimensional, even though the real potential surfaces for colliding molecules are multidimensional (see the preceding note).

[3)]In connection with this assertion, we can indicate Refs. 50 and 51, in which the role of broadening was analyzed in a Landau–Zener process similar to the one considered here.

[1] S. Hayashi, J. Phys. Soc. Jpn. **32**, 349 (1972).
[2] J. M. Galligan and C. S. Pang, J. Appl. Phys. **50**, 6353 (1979).
[3] M. S. C. Bose, Phys. Status Solidi A **86** (649) (1984).
[4] Y. Y. Su, R. F. Hochman, and J. R. Schaffer, J. Phys. C **2**, 3629 (1990).
[5] Ya. B. Zel'dovich A. L. Buchachenko, and E. L. Frankevich, Usp. Fiz. Nauk **155**, 3 (1988) [Sov. Phys. Usp. **31**, 385 (1988)].
[6] R. Jemielniak and J. Krolikowski, J. Phys. (Paris) Colloq. **46**, C10 (1985).
[7] U. Nadel and G. Liidja, Z. Phys. B **91**, 21 (1993).
[8] V. I. Al'shits, E. V. Darinskaya, and T. M. Perekalina, Fiz. Tverd. Tela (Leningrad) **29**, 467 (1987) [Sov. Phys. Solid State **29**, 265 (1987)].
[9] V. I. Al'shits, R. Voska, E. V. Darinskaya *et al.*, Fiz. Tverd. Tela **35**, 70 (1993) [Phys. Solid State **35**, 37 (1993)].
[10] V. I. Al'shits, E. V. Darinskaya, and E. A. Petrazhik, Fiz. Tverd. Tela (Leningrad) **34**, 81 (1992) [Sov. Phys. Solid State **34**, 43 (1992)].
[11] V. I. Alshits, E. V. Darinskaya, O. L. Kazarinova *et al.*, J. Alloys Compounds **211**, 548 (1994).
[12] Yu. I. Golovin and R. B. Morgunov, JETP Lett. **58**, 191 (1993).
[13] Yu. I. Golovin and R. B. Morgunov, Fiz. Tverd. Tela **37**, 1352 (1995) [Phys. Solid State **37**, 734 (1995)].
[14] Yu. N. Molin, K. M. Salikhov, A. L. Buchachenko, and R. Z. Sagdeev, *Spin Polarization and Magnetic Effects in Radical Reactions* (Elsevier, Amsterdam, 1984).
[15] K. M. Salikhov, Yu. N. Molin, R. Z. Sagdeev, and A. L. Buchachenko, *Spin Polarization and Magnetic Effects in Radical Reactions* (Elsevier, Amsterdam, 1984).
[16] U. E. Steiner and T. Ulrich, Chem. Rev. **89**, 51 (1989).
[17] V. O. Saik, O. A. Anisimov, A. V. Koptyng *et al.*, Chem. Phys. Lett. **165**, 142 (1990).
[18] K. M. Salikhov and Yu. N. Molin, J. Phys. Chem. **97**, 13 259 (1993).
[19] M. I. Molotskiĭ, Fiz. Tverd. Tela (Leningrad) **33**, 3112 (1991) [Sov. Phys. Solid State **33**, 1760 (1991)].
[20] M. I. Molotskiĭ, Fiz. Tverd. Tela **35**, 5 (1993) [Phys. Solid State **35**, 1 (1993)].
[21] M. I. Molotskiĭ and V. Fleurov, Phys. Rev. B **52**, 15 829 (1995).
[22] M. I. Molotskiĭ, R. E. Kris, and V. Fleurov, Phys. Rev. B **51**, 12 531 (1995).
[23] M. I. Molotskiĭ and V. Fleurov, Phil. Mag. Lett. **73**, 11 (1996).
[24] G. L. Distler, V. M. Kanevskiĭ, V. V. Moskvin *et al.*, Dokl. Akad. Nauk SSSR **268**, 591 (1983) [Sov. Phys. Dokl. **28**, 43 (1983)].
[25] S. A. Dembovskiĭ, E. A. Chechetkina, and A. L. Kozyunin, JETP Lett. **41**, 88 (1985).
[26] V. P. Vlasov, F. A. Zaitov, V. M. Kanevskiĭ *et al.*, Fiz. Tverd. Tela (Leningrad) **34**, 3264 (1992) [Sov. Phys. Solid State **34**, 1747 (1992)].
[27] V. N. Davydov, E. A. Loskutova, and E. P. Naĭden, Fiz. Tekh. Poluprovodn. **23**, 1596 (1989) [Sov. Phys. Semicond. **23**, 989 (1989)].
[28] N. V. Kukushkin, S. N. Postnikov, Yu. A. Terman *et al.*, Zh. Tekh. Fiz. **55**, 2083 (1985) [Sov. Phys. Tech. Phys. **30**, 1227 (1985)].
[29] M. N. Levin and V. M. Maslovsky, Mat. Res. Soc. Symp. Proc. Ser. **319**, 429 (1994).
[30] A. D. Kadmenskiĭ, S. G. Kadmenskiĭ, M. N. Levin *et al.*, Pis'ma Zh. Tekh. Fiz. **19(3)**, 41 (1993) [Tech. Phys. Lett. **19**, 86 (1993)].
[31] M. N. Levin, Yu. O. Lichmanov, and V. M. Maslovskiĭ, Pis'ma Zh. Tekh. Fiz. **20(4)**, 27 (1994) [Tech. Phys. Lett. **20**, 145 (1994)].
[32] M. N. Levin and V. M. Maslovsky, Solid State Commun. **90**, 813 (1994).
[33] V. M. Maslovskiĭ, Yu. A. Klimov, N. S. Samsonov *et al.*, Fiz. Tekh. Poluprovodn. **28**, 772 (1994) [Semiconductors **28**, 453 (1994)].
[34] K. V. Ravi, *Imperfections and Impurities in Semiconducting Silicon* (Wiley, New York, 1981; Mir, Moscow, 1984).
[35] V. S. Vasilov, V. F. Kiselev, and B. N. Mukashev, *Defects in Silicon and on Its Surface* (Nauka, Moscow, 1990).
[36] G. D. Watkins and J. W. Corbett, Phys. Rev. **121**, 1001 (1961).
[37] J. W. Corbett, G. D. Watkins, R. M. Chrenko *et al.*, Phys. Rev. **121**, 1015 (1961).
[38] G. G. De Leo, W. B. Fowler, and G. D. Watkins, Phys. Rev. B **29**, 3193 (1984).
[39] Ya. A. Latushko, A. V. Mudryĭ, and V. D. Tkachev, Zh. Prikl. Spektrosk. **40**, 983 (1984).
[40] F. Warger and J. Hage, Appl. Phys. A **49**, 123 (1989).
[41] J. L. Lindström and T. Halberg, J. Appl. Phys. **77**, 2684 (1995).
[42] D. V. Lang, J. Appl. Phys. **45**, 3023 (1974).
[43] N. M. Johnson, D. J. Bartelink, R. B. Gold *et al.*, J. Appl. Phys. **50**, 4828 (1979).
[44] L. S. Berman and A. A. Lebedev, *Capacitance Spectroscopy of Deep Centers in Semiconductors* (Nauka, Leningrad, 1981).
[45] L. C. Kimerling, in *Radiation Effects in Semiconductors*, vol. 31, ed. N. B. Urli and J. W. Corbett (Inst. of Physics, London, 1977), p. 2211.
[46] J. M. Meese, J. W. Farmer, and C. D. Lamp, Phys. Rev. Lett. **51**, 1286 (1983).
[47] P. Pringsheim, *Fluorescence and Phosphorescence* (Interscience, New York, 1949; Gostekhizdat., Moscow, 1951).
[48] J. Plans, G. Diaz, E. Martinez *et al.*, Phys. Rev. B **35**, 788 (1987).
[49] E. E. Nikitin and S. Ya. Umanskiĭ, *Nonadiabatic Transitions for Slow Atomic Collisions* (Moscow, Atomizdat, 1979).
[50] L. M. Kishinevskiĭ and É. M. Parilis, Zh. Éksp. Teor. Fiz. **55**, 1932 (1968) [Sov. Phys. JETP **28**, 1020 (1968)].
[51] B. A. Zon, Zh. Éksp. Teor. Fiz. **75**, 834 (1978) [Sov. Phys. JETP **48**, 422 (1978)].
[52] B. N. Romanyuk, V. G. Popov, V. G. Litovchenko *et al.*, Fiz. Tekh. Poluprovodn. **29**, 166 (1995) [Semiconductors **29**, 87 (1995)].
[53] V. V. Artamonov, M. Ya. Valakh, F. Kirsht *et al.*, Fiz. Tekh. Poluprovodn. **25**, 1704 (1991) [Sov. Phys. Semicond. **25**, 1025 (1991)].
[54] V. V. Emtsev, G. A. Oganesyan, and K. Shmal'ts Fiz. Tekh. Poluprovodn. **27**, 1549 (1993) [Semiconductors **27**, 856 (1993)].
[55] J. Linnros, B. Svensson, and G. Holmen, Phys. Rev. B **30**, 3629 (1984).

Translated by W. J. Manthey

The Peierls system in a light field

A. L. Semenov

Ul'yanovsk Branch of the M. V. Lomonosov Moscow State University, 432700 Ul'yanovsk, Russia
(Submitted 4 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 1398–1409 (April 1997)

This paper studies the behavior of the low-temperature phase of the Peierls system in a quasimonochromatic time-independent random light field whose frequency is much lower than the frequency of the lower edge of band-to-band transitions. The density matrix method in the dipole approximation is used to derive equations for the band gap. A dependence between the band gap and the light-field intensity is established in an approximation in which the concentration of the nonequilibrium electrons in the conduction band is low. The possibility of, and the conditions for, the existence of a light-induced semiconductor–semiconductor phase transition and cavityless optical bistability with increasing absorption are established.
 [S1063-7761(97)02004-0]

1. INTRODUCTION

The Peierls system is a one-dimensional chain of atoms each of which contains one outer electron. In the high-temperature metallic phase, the atoms in the chains are positioned equidistantly and the conduction band is half-filled. When the system is cooled and the temperature becomes lower than a critical value T_0, the metallic phase becomes unstable, and a metal–semiconductor phase transition occurs. As a result the chain atoms move closer together in pairwise fashion.[1]

The theoretical results obtained from the Peierls model are used to describe the experimentally observed properties of many quasi-one-dimensional materials: complex vanadium compounds,[1] *m*-TaS_3compounds,[2] organic conductors,[3] oxide vanadium bronzes,[4,5] and others.[6,7] In particular, this model makes it possible to explain the effect on the metal–semiconductor transition in vanadium dioxide films of uniaxial pressure,[8] alloying,[9–11] cohesion of a vanadium dioxide film and the substrate,[12] adsorption, and other factors.[13]

The behavior of the low-temperature phase of the Peierls system in a light field whose frequency exceeds the frequency of the lower edge of band-to-band transitions has been studied both theoretically[14] and experimentally (in vanadium dioxide).[15,16] It was found that because of strong electron–phonon coupling under irradiation, the forbidden band narrows, and in strong fields there occurs a semiconductor–metal phase transition. This result is based on the fact that an increase in the concentration of conduction electrons narrows the forbidden band of the Peierls system.

This paper studies the behavior of the low-temperature phase of the Peierls system in a quasimonochromatic light field whose frequency lies somewhat below the lower edge of band-to-band transitions. Nevertheless, the finite width of the optical spectrum of the light field allows processes in which electrons ''hop'' from the valence band to the conduction band. In such conditions, when the irradiation intensity increases, there may be a sudden drop in the band gap in the system (a light-induced semiconductor–semiconductor phase transition) by virtue of positive feedback. Initially a nonresonance light field causes a slight decrease in the band gap. As a result, electron–phonon coupling becomes more resonant and hence more intensive, which in turn leads to a further decrease in the band gap. Thus, positive feedback lowers the stability of the system. If the positive feedback is strong, stability can be lost, and there is a sudden transition to a new state of equilibrium.

Thus, thanks to the strong coupling between the electron subsystem and the static phonon mode in the Peierls model,[8] one of the mechanism causing cavityless optical bistability with increasing absorption comes into play.[17] Generally, the given type of bistability is due to the nonlinear dependence of the absorption coefficient on the incident radiation intensity and may be brought on by different causes:[17] atomic correlations within a small volume, the temperature dependence of the band gap, renormalization of the band gap due to electron–phonon coupling,[17,18] and variations in the frequency of transitions in an ensemble of two-level atoms due to local-field effects or electron–phonon coupling.[19] However, the author knows of no mechanism similar to that of cavityless optical bistability in the Peierls system considered in this paper.

2. ELECTRON SPECTRUM OF THE SYSTEM

Let us examine a chain of atoms each of which has one outer electron. The Hamiltonian of the electron subsystem in the tight-binding approximation can be written as[1]

$$H=\sum_n B_{n,n+1}(a_n^+ a_{n+1}+a_{n+1}^+ a_n), \tag{1}$$

where n is the number of the atom in the chain, $B_{n,n+1}$ is the overlap integral of the wave functions of neighboring electrons, and a_n^+ and a_n are the electron creation and annihilation operators at the nth atom.

For narrow-gap systems, e.g., for the Peierls model, the separation of adjacent atoms, $r_{n,n+1}$, exceeds the effective radius R of the atomic wave function of an electron several-fold. In this case $B_{n,n+1} \propto \exp(-r_{n,n+1}/R)$ (Ref. 20). The separation $r_{n,n+1}$ can be written

 1063-7761/97/040774-06$10.00

$$r_{n,n+1}=r_0+(-1)^{n+1}R\xi, \tag{2}$$

where r_0 is the atomic separation in the metallic phase, and ξ is the period-doubling parameter for a one-dimensional crystal. If we allow for (2), the overlap integral $B_{n,n+1}$ assumes the form

$$B_{n,n+1}=b\ \exp((-1)^n\xi), \tag{3}$$

where b is the overlap integral in the metallic phase ($\xi=0$). With Hamiltonian (1), the phases of the atomic wave functions are selected in such a way that b in (3) is real.

To diagonalize the Hamiltonian (1) we employ Bogolyubov's method of canonical transformations.[21] We introduce collective second-quantization Fermi operators c_k and c_k^+ as follows:

$$a_n=\frac{1}{\sqrt{N}}\sum_k\ c_k e^{ikn}, \tag{4}$$

where N is the number of atoms in the chain, $k=0$, $\pm 2\pi/N$, ..., $\pm\pi$, and $c_{k+2\pi}=c_k$.

In the new operator representation the Hamiltonian (1) becomes

$$H=\sum_k\ 2b(c_k^+c_k\ \cosh\ \xi\ \cos\ k+ic_k^+c_{k-\pi}\ \sinh\ \xi\ \sin\ k). \tag{5}$$

Now in (5) we apply the canonical transformation to the operators α_k and α_k^+:

$$c_k=\frac{\alpha_k+i\varphi_k\alpha_{k-\pi}}{\sqrt{1+\varphi_k^2}}. \tag{6}$$

The function φ_k is selected in such a way that the resulting Hamiltonian is diagonal in the new variables α_k and α_k^+:

$$H=\sum_k\ \varepsilon_k\alpha_k^+\alpha_k. \tag{7}$$

Plugging (6) into (5) and zeroing out the off-diagonal elements, we arrive at an expression for φ_k and a dispersion relation for ε_k:

$$\varphi_k=\frac{\cosh\ \xi\ \cos\ k-\mathrm{sgn}(\cos\ k)\sqrt{\cos^2k+\sinh^2\xi}}{\sinh\ \xi\ \sin\ k}, \tag{8}$$

$$\varepsilon_k=2b\ \mathrm{sgn}(\cos\ k)\sqrt{\cos^2k+\sinh^2\xi}. \tag{9}$$

Equation (9) shows that for $\xi\neq 0$ the spectrum of ε_k has two bands, with the lower band in the ground state completely occupied and the upper band vacant (the semiconducting phase). At $\xi=0$ the spectrum (9) consists of one half-filled band (the metallic phase).

3. DIPOLE MOMENT OPERATOR

In the tight-binding approximation, the dipole moment operator of the Peierls system is

$$\mathbf{d}=\sum_n\ (\mathbf{d}_{n,n+1}a_n^+a_{n+1}+\mathbf{d}_{n,n+1}^*a_{n+1}^+a_n), \tag{10}$$

where the dependence of $\mathbf{d}_{n,n+1}$ on ξ is similar to (3):

$$\mathbf{d}_{n,n+1}=(\mathbf{d}_1+i\mathbf{d}_2)\exp[(-1)^n\xi]$$
$$=-e\int_V\psi_n^*(\mathbf{r})\mathbf{r}\psi_{n+1}(\mathbf{r})dV. \tag{11}$$

Here $\psi_n(\mathbf{r})$ is the atomic wave function of the electron at the nth site, and e is the electron charge. By selecting the phases of the wave functions $\psi_n(\mathbf{r})$ in such a way that the overlap integral (3) is real we ensure that both $\mathbf{d}_1$ and $\mathbf{d}_2$ in (11) are uniquely defined.

Plugging (4) into (10) and allowing for (11) and the identity $\exp[(-1)^n\xi]=\cosh\ \xi+(-1)^n\sinh\ \xi$, we obtain

$$\mathbf{d}=2\sum_k\ [\cosh\ \xi(\mathbf{d}_1\ \cos\ k-\mathbf{d}_2\ \sin\ k)c_k^+c_k$$
$$+i\ \sinh\ \xi(\mathbf{d}_1\ \sin\ k+\mathbf{d}_2\ \cos\ k)c_k^+c_{k-\pi}]. \tag{12}$$

Introducing the Fermi operators α_k and α_k^+ into (12) and allowing for (6) and (8) we finally obtain

$$\mathbf{d}=\sum_k\ \left\{\left[\frac{\mathbf{d}_1\varepsilon_k}{b}-\frac{2\mathbf{d}_2}{1+\varphi_k^2}[(1-\varphi_k^2)\cosh\ \xi\ \sin\ k+2\varphi_k\ \sinh\ \xi\right.\right.$$
$$\left.\times\cos\ k]\right]\alpha_k^+\alpha_k+i\frac{2\mathbf{d}_2}{1+\varphi_k^2}[(1-\varphi_k^2)\sinh\ \xi$$
$$\left.\times\cos\ k-2\varphi_k\ \cosh\ \xi\ \sin\ k]\alpha_k^+\alpha_{k-\pi}\right\}. \tag{13}$$

Note the formal similarity of the operators (1) and (10) at $\mathbf{d}_2=0$. The operators (7) and (13) are also formally similar.

Suppose that in the absence of an external electric field the total dipole moment of the system is zero. Then Eqs. (13) and (9) imply that $\mathbf{d}_1=0$. Thus, the fact that we have chosen the phases of the wave functions $\psi_n(\mathbf{r})$ in such a way that the overlap integral $B_{n,n+1}$ in (1) is real ensures, in the present case, that the interstitial dipole-moment matrix element $\mathbf{d}_{n,n+1}$ in (10) is imaginary. The case of $\mathbf{d}_1\neq 0$ can probably be observed in systems exhibiting ferroelectric properties, but we will not consider such systems here. As $\xi\to 0$, as Eq. (8) indicates, $\varphi_k\to 0$ for all $k\neq\pm\pi/2$, with the result that $\mathbf{d}_{k,k-\pi}\to 0$ in (13) and all dipole transitions are forbidden. If $\xi\neq 0$, then $\mathbf{d}_{k,k-\pi}\neq 0$ in (13), and the corresponding dipole transitions are allowed. Since in this case the interval $k\in[-\pi/2,\pi/2]$ is the first Brillouin zone, the given transitions in the spectrum (9) are vertical band-to-band transitions.

4. EQUILIBRIUM EQUATION

In describing the behavior of the Peierls system in a light field we assume that the electron intraband relaxation time is much shorter than the band-to-band relaxation time τ (Ref. 22). In this case, following the approximation adopted in Refs. 14 and 23, we can assume that the electron statistics in each band is described by a Fermi distribution with a specific Fermi quasilevel. Then, allowing for (9), we arrive at the following expression for the population difference Δ between the valence and conduction bands in a light field:

$$\Delta = 2\sum_{|k|\leqslant \pi/2} \Delta_k = 2\sum_{|k|\leqslant \pi/2} \tanh\left(\frac{\varepsilon_k - \mu}{2k_BT}\right). \tag{14}$$

Here μ is the Fermi quasilevel of the conduction band (in view of the symmetry of the dispersion relation (9) the Fermi quasilevel of the valence band has the opposite value), k_B is Boltzmann's constant, and $\Delta_k = \rho_{k-\pi,k-\pi} - \rho_{kk}$, with ρ_{kk} an element of the density matrix ρ of the electron subsystem. The factor of 2 in front of the sum in (14) is present because of the spin degeneracy of each level.

Treating the parameter ξ as a generalized coordinate, we can write the equilibrium equation for the Peierls system:

$$\left(\frac{\partial F_1}{\partial \xi}\right)_{T,N_1} + \left(\frac{\partial F_2}{\partial \xi}\right)_{T,N_2} + \left(\frac{\partial F_{\rm ph}}{\partial \xi}\right)_{T} = 0, \tag{15}$$

where

$$F_j = \mu_j N_j - k_BT\sum_k \ln\left[1 + \exp\left(\frac{\mu_j - \varepsilon_k}{k_BT}\right)\right], \tag{16}$$

$$F_{\rm ph} = \frac{A}{2}\xi^2,$$

with F_j, μ_j, and N_j the electron free energy, the Fermi quasilevel, and the number of electrons belonging to the jth band ($j=1,2$), and A is the coefficient in the expansion of the free energy of the phonon subsystem, $F_{\rm ph}$, in powers of ξ (see Ref. 1). Summing over k in (16) must be done within the jth band of the spectrum (9).

Plugging (16) into (15) and performing the necessary transformations, we finally get

$$-2\sum_{|k|\leqslant \pi/2} \frac{\partial \varepsilon_k}{\partial \xi} \tanh\left(\frac{\varepsilon_k - \mu}{2k_BT}\right) + A\xi = 0. \tag{17}$$

The expansion coefficient A can be expressed in terms of the critical temperature T_0 of a thermodynamic equilibrium metal–semiconductor transition (in the absence of a light field) and other characteristics of the system.[8,11] Such a transition occurs when the metallic phase loses stability ($\partial^2 F(T_0, \xi=0)/\partial \xi^2 = 0$, where F is the free energy of the system). In the absence of a light field, the Fermi quasilevels of the valence and conduction bands of the spectrum (9) are zero: $\mu_{1,2} = \mp \mu = 0$. Combining this with Eqs. (17) and (9), we get

$$A = 2\sum_{|k|\leqslant \pi/2} \left(\frac{\partial^2 \varepsilon_k}{\partial \xi^2} \tanh \frac{\varepsilon_k}{2k_BT_0}\right)_{\xi=0}. \tag{18}$$

Calculating the sum in (18), we arrive at the following approximate expression for A:

$$A = \frac{4bN}{\pi}\left[\ln\left(\frac{\pi b}{2k_BT_0}\right) + 1\right]. \tag{19}$$

Thus we have obtained an equation (Eq. (14)) expressing the dependence of the population difference Δ on the Fermi quasilevel μ and an equilibrium equation (Eq. (17)) describing the relationship between the parameter ξ of the Peierls system and μ. The Fermi quasilevel μ is determined in turn by the degree to which light acts on the system. An equation describing such action must depend on the specific mechanism of interaction of the radiation and the electron subsystem.

5. INTERACTION WITH RADIATION

The interaction between the system and the light field is described by an operator V, which in the dipole approximation can be written as

$$V = -\mathbf{d}\cdot\mathbf{E}(t) = -\mathbf{d}\cdot\int \mathbf{E}_\omega e^{-i\omega t} d\omega, \tag{20}$$

where $\mathbf{E}_\omega$ and ω are the amplitude and frequency of a spectral component of the light field.

We assume the incident radiation $\mathbf{E}(t)$ to be a quasimonochromatic time-independent random process linearly polarized along the crystal's axis.[24] As a result, all the spectral components $\mathbf{E}_\omega$ are statistically independent:[24,25]

$$\langle \mathbf{E}_\omega \cdot \mathbf{E}_{\omega_1} \rangle = G(\omega)\,\delta(\omega + \omega_1). \tag{21}$$

Here $G(\omega)$ is the spectral density of the light field, which for a quasimonochromatic signal can be written as[24]

$$G(\omega) = Ig(|\omega| - \omega_0), \tag{22}$$

where ω_0 is the carrier frequency, and $g(x)$ is a bell-shaped function, with its maximum at $x=0$, satisfying the normalization condition $\int g(x)dx = 1$. The width $\Delta\omega$ of the spectrum $G(\omega)$ satisfies the inequality $\Delta\omega \ll \omega_0$. The quantity $I = \int G(\omega)d\omega/2$ is the intensity of the light field (in the Gaussian system of units, to within the factor $cn/2\pi$, where c is the speed of light, and n is the medium's refractive index).

Using Liouville's equation[26]

$$i\hbar \frac{\partial \rho}{\partial t} = [H + V, \rho] \tag{23}$$

and allowing for Eqs. (20) and (21), we arrive at an equation for the diagonal elements ρ_{kk} of the density matrix ρ of the electron subsystem in second-order perturbation theory:

$$\frac{\partial \rho_{kk}}{\partial t} = \frac{2\pi}{\hbar^2}\sum_s |\mathbf{d}_{ks}|^2 G\left(\frac{\varepsilon_s - \varepsilon_k}{\hbar}\right)(\rho_{ss} - \rho_{kk}), \tag{24}$$

where $\mathbf{d}_{ks}$ is the matrix element of the dipole moment operator (13). In the special case of a monochromatic field $E(t) = E_0\cos(\omega_0 t + \varphi)$ with a uniformly distributed random phase φ, the spectral density is

$$G(\omega) = \tfrac{1}{4}E_0^2[\,\delta(\omega - \omega_0) + \delta(\omega + \omega_0)\,].$$

Then Eq. (24) becomes Fermi's well-known Golden Rule for the probability of stimulated transitions.[25]

If we combine (13) and (9) with (24), we get

$$\frac{\partial \rho_{kk}}{\partial t} = \frac{2\pi}{\hbar^2} d_k^2 G\left(\frac{2\varepsilon_k}{\hbar}\right)\Delta_k, \tag{25}$$

where

$$d_k = \frac{2d_2}{1+\varphi_k^2}\left|(1-\varphi_k^2)\sinh\xi\cos k - 2\varphi_k\cosh\xi\sin k\right|. \tag{26}$$

Employing Eq. (14) and the fact that $\rho_{kk}+\rho_{k-\pi,k-\pi}=1$, which is true due to the symmetry of the electron spectrum (9), we can use (25) to derive the following kinetic equation for the population difference Δ:

$$\frac{\partial\Delta}{\partial t}=-\frac{8\pi}{\hbar^2}\sum_{|k|\leqslant\pi/2} d_k^2\,\tanh\left(\frac{\varepsilon_k-\mu}{2k_BT}\right)G\left(\frac{2\varepsilon_k}{\hbar}\right)-\frac{\Delta-\Delta_0}{\tau}, \tag{27}$$

where Δ_0 is the band population difference in the absence of a light field: $\Delta_0=\Delta(\mu=0)$. The last term on the right-hand side of Eq. (27) allows for band-to-band relaxation with a relaxation time τ. Equation (27) shows that the variation in Δ is due to the interaction of the electron subsystem and the spectral components of the light field with frequencies $\omega_k=2\varepsilon_k/\hbar$.

Combining (22) with Eq. (27) in the steady-state ($\partial\Delta/\partial t=0$), we finally arrive at the following expression for the intensity:

$$I=\frac{(\Delta_0-\Delta)\hbar^2}{8\pi\tau}\left(\sum_{|k|\leqslant\pi/2} d_k^2\,\tanh\left(\frac{\varepsilon_k-\mu}{2k_BT}\right)g\left(\frac{2\varepsilon_k}{\hbar}-\omega_0\right)\right)^{-1}. \tag{28}$$

Equations (14), (17), and (28) constitute a closed system of equations with respect to the internal parameters μ, Δ, and ξ of the Peierls system with given external parameters T, I, ω_0, etc.

6. RELATIONSHIP BETWEEN THE BAND GAP AND THE INTENSITY OF THE LIGHT FIELD

Let us analyze Eqs. (14), (17), and (28) for the case where the light frequency is smaller than the lower edge of band-to-band transitions: $\varepsilon_0-\hbar\omega_0>0$, where ε_0 is the band gap of the system in the absence of a light field.

The light is assumed weak, so that

$$\frac{\varepsilon/2-\mu}{k_BT}>2, \tag{29}$$

where ε is the band gap. Physically, Eq. (29) means that 1) $4k_BT<\varepsilon$, and 2) the concentration of excited electrons lies so low that the Fermi quasilevel is in the forbidden band. Then, taking $\exp[-(\varepsilon/2-\mu)/k_BT]$ as a small parameter, we can approximately transform Eq. (14) for Δ into

$$\Delta=N\left[1-\sqrt{\frac{\varepsilon k_BT}{\pi b^2}}\exp\left(-\frac{\varepsilon/2-\mu}{k_BT}\right)\right]. \tag{30}$$

We write the equilibrium equation (17) in the following form:

$$\exp\left(-\frac{\varepsilon/2-\mu}{k_BT}\right)=\sqrt{\frac{\varepsilon}{4\pi k_BT}}\left[\arcsin\frac{\pi}{4}+\ln\left(\frac{2\pi b}{\varepsilon}\right)-\frac{\pi A}{4bN}\right]. \tag{31}$$

Here in addition to assuming that condition (29) is met we assume that $\xi\leqslant 0.5$. This is justified because such values of ξ are realized in the majority of materials described by the Peierls model.[1,6,8]

Having (31) in mind, we can use (19) to calculate the band gap ε_0 in the absence of light ($\mu=0$) for $k_BT\ll\varepsilon$:

$$\varepsilon_0=4k_BT_0\,\exp\left[\arcsin\frac{\pi}{4}-1\right]\simeq 3.63k_BT_0. \tag{32}$$

The condition $k_BT\ll\varepsilon$, for which (32) is valid, then becomes $T<T_0/2$. Note that a similar result $\varepsilon\simeq 3.53k_BT_0$ was obtained by Bulaevskiĭ.[1]

Plugging (31) into (30) and allowing for (32), we arrive at a relationship linking the nonequilibrium population difference Δ to the band gap ε:

$$\Delta=N\left(1-\frac{\varepsilon}{2\pi b}\,\ln\frac{\varepsilon_0}{\varepsilon}\right). \tag{33}$$

This is true when $\Delta\varepsilon\equiv\varepsilon_0-\varepsilon$ is small, and the reason for this lies in the adopted approximation (29), which with allowance for (31) and (32) can be written in the more convenient form

$$\Delta\varepsilon\equiv\varepsilon_0-\varepsilon<\frac{\sqrt{\varepsilon_0k_BT}}{2}, \tag{34}$$

where, as in (32), it is assumed that $T<T_0/2$.

To solve Eq. (28) we must know the function $g(\omega)$ characterizing the spectral density of the light field. To this end we assume that $g(\omega)$ has the Lorentzian form[25]

$$g(\omega)=\frac{\tau_0}{\pi[1+(\omega\tau_0)^2]}, \tag{35}$$

where τ_0 is the reciprocal of the spectral halfwidth.

Bearing in mind that near the lower edge of an optical transition $d_k\simeq 2d_2\cosh\xi$, as Eq. (26) implies, and combining Eq. (28) with (33), (35), and (9), we arrive at the following expression for the light field intensity:

$$I=\frac{\hbar^2x_0z(x)}{16\pi d_2^2\tau\sqrt{\tau_0^3\omega_0}}, \tag{36}$$

where

$$z(x)=\left(1-\frac{x}{x_0}\right)\sqrt{(1+x^2)(x+\sqrt{1+x^2})}, \tag{37}$$

and $x=\tau_0(\varepsilon/\hbar-\omega_0)$ is a dimensionless parameter characterizing the detuning of the lower-edge optical-transition frequency from the carrier frequency of the light field, with $x_0=\tau_0(\varepsilon_0/\hbar-\omega_0)$ the initial detuning.

7. SEMICONDUCTOR–SEMICONDUCTOR PHASE TRANSITION AND CAVITYLESS OPTICAL BISTABILITY

Equations (36) and (37) describe the relationship between the light-field intensity I and the band gap $\varepsilon=\hbar(\omega_0+x/\tau_0)$ of the Peierls system in a stationary state of equilibrium. An analysis of the kinetic equation (27) shows that for the state to be stable $\partial I/\partial\varepsilon$ must be negative, which in a form more convenient for analysis can be expressed as

$$\frac{\partial z}{\partial x}<0. \tag{38}$$

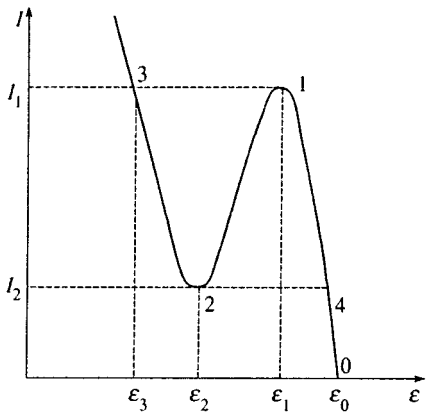

FIG. 1. The light-field intensity I as a function of the band gap ε.

Let us study the z vs. x dependence (37) in the range $|x|\leqslant 1$. To this end we expand (37) in Taylor series up to terms cubic in x:

$$z(x)=1+a_1x+a_2x^2+a_3x^3, \tag{39}$$

where

$$a_1=\frac{x_0-2}{2x_0}, \quad a_2=\frac{5x_0-4}{8x_0}, \quad a_3=\frac{3x_0-10}{16x_0}. \tag{40}$$

The points x_1 and x_2 corresponding to the loss of stability in the system can be found by solving the equation $\partial z/\partial x=0$. Combining this with (39) yields

$$x_{1,2}=-\frac{a_2\pm\sqrt{a_2^2-3a_1a_3}}{3a_3}. \tag{41}$$

We see that the behavior of the system is critical if $a_2^2-3a_1a_3>0$. Combining this fact with (40), we arrive at the necessary condition for a phase transition, a condition that imposes a contraint on the initial detuning of the light-field carrier frequency from the lower-edge optical-transition frequency, $x_0>1.55$, or, in the initial notation,

$$\varepsilon_0-\hbar\omega_0>\frac{1.55\hbar}{\tau_0}. \tag{42}$$

If condition (42) is met, an analysis of Eqs. (36) and (37) shows that I has the behavior depicted in Fig. 1 as a function of ε. As the light intensity increases, a point on the graph moves along the curve $0\to4\to1\to3$, while as the intensity decreases the point moves along the curve $3\to2\to4\to0$. The jumps in the value of ε in the $1\to3$ and $2\to4$ sections correspond to direct and inverse phase transitions. The interval $I_2<I<I_1$ constitutes the bistability region.

Analyzing Eq. (39), we can find the points x_3 and x_4 corresponding to the new positions of equilibrium for the direct and inverse transitions,

$$x_{3,4}=-\frac{a_2\mp2\sqrt{a_2^2-3a_1a_3}}{3a_3}, \tag{43}$$

and the dimensionless value Δx of the jump in the band gap, $\Delta\varepsilon=\hbar\Delta x/\tau_0$ in a phase transition:

$$\Delta x=-\frac{\sqrt{a_2^2-3a_1a_3}}{a_3}. \tag{44}$$

Here it turns out that Δx is the same for direct and inverse transitions.

Using (39), we can also derive an expression for the width of the hysteresis loop in a semiconductor–semiconductor phase transition (the size of the region of cavityless optical bistability in terms of the dimensionless light-field intensity):

$$\Delta z=z(x_1)-z(x_2)=\frac{8(a_2^2-3a_1a_3)^{3/2}}{27a_3^2}. \tag{45}$$

Equations (39)–(45) were derived for $|x|\leqslant1$, which contrains x_0. Using (40), we find that Eq. (41) for x_1 is valid if $x_0\leqslant1.77$ and for x_2 if $x_0\leqslant3.32$, while Eq. (43) for x_3 is valid if $x_0\leqslant2$ and for x_4 if $x_0\leqslant1.60$. Thus, within the present approximation we are able to describe the behavior of the system near the threshold value $x_0=1.55$, above which critical behavior is observed.

For the sake of numerical estimates, let us assume that $x_0=1.6$. Then, in accordance with Eqs. (39)–(45), we have

$$x_1\simeq0.757, \quad x_2\simeq0.271, \quad x_3\simeq0.028, \quad x_4\simeq1,$$

$$\Delta x\simeq0.729, \quad \Delta z\simeq0.409, \quad z(x_1)\simeq1.$$

If we put $d_2\approx10^{-18}$ esu (Refs. 19 and 25), $\tau\approx10^{-10}$ s (Ref. 22), $\tau_0\approx10^{-13}$ s, and $\omega_0\approx10^{15}$ s^{-1}, we find the detunings $\varepsilon_j-\hbar\omega_0=\hbar x_j/\tau_0\approx x_j\cdot10^{-2}$ eV for $j=0,1,2,3,4$ and the jump in the band gap in a phase transition, $\Delta\varepsilon=\hbar\Delta x/\tau_0\approx0.7\times10^{-2}$ eV. Employing (36), we find that $I_1\approx10^3$ esu and $\Delta I=I_1-I_2\approx0.4\times10^3$ esu, which corresponds to values of the light-field intensity of about 10^5 W/cm^2.

8. CONCLUSION

Our theory suggests that when the Peierls system is irradiated with light whose frequency is much lower than the frequency of the lower edge of band-to-band transitions, a semiconductor–semiconductor phase transition and cavityless optical bistability can be observed in such a system. These phenomena occur when condition (42) is met, i.e., when the light-field frequency ω_0 is bounded above. The fact that such a condition exists can be explained by competition of two feedback mechanisms. Positive feedback is reflected by the first term on the right-hand side of Eq. (27), and the mechanism has been described in detail in the Introduction. Negative feedback consists in the following: as the concentration n of the nonequilibrium electrons in the conduction band grows, the recombination rate increases, which drives n down. This process is represented by the second term on the right-hand side of Eq. (27). If negative feedback is stronger than positive feedback, the equilibrium state of the system is stable, and slow variations in the external parameters lead to smooth variations in the internal parameters. But if positive feedback is dominant, then, in accordance with (27), there is a sharp increase in fluctuations, which leads to a sudden variation in the internal parameters of the system.

When the initial detunings of the frequency ω_0 are much larger than the width $\Delta\omega$ of the optical spectrum, i.e., $\varepsilon_0-\hbar\omega_0\gg\hbar/\tau_0$, the intensity I needed for observing a possible semiconductor–semiconductor phase transition increases sharply. The present theory is unable to examine this frequency range, since the approximation (34) of a weak optical excitation is invalid in this range. At temperatures $T\approx 100$ K ($k_BT\approx 8.6\times 10^{-3}$ eV) and a band gap $\varepsilon_0\approx 0.5$ eV this condition imposes the constraint $0<\varepsilon_0-\varepsilon<0.033$ eV.

In conclusion it must be noted that the phenomenon of cavityless optical bistability with increasing absorption similar to that studied in the present paper (but caused by other mechanisms) was observed in experiments involving zinc selenide,[27] the amorphous semiconductor GeS_2 (Ref. 28), amorphous $GeSe_2$ films,[17] CdS plates, and other materials.[17,18] However, the author knows of no such experiments involving materials whose electronic properties can be described by the Peierls model.

[1] L. N. Bulaevskiĭ, Usp. Fiz. Nauk **115**, 263 (1975).
[2] V. E. Minakova, Yu. I. Latyshev, and V. A. Volkov, JETP Lett. **62**, 455 (1995).
[3] L. P. Gor'kov, Usp. Fiz. Nauk **144**, 381 (1984).
[4] A. A. Fotiev, V. L. Volkov, and V. K. Kapustkin, *Oxide Vanadium Bronzes* [in Russian], Nauka, Moscow (1978), p. 158.
[5] A. V. Dmitriev, N. A. Zhuravlev, and V. L. Volkov, Fiz. Tverd. Tela (Leningrad) **32**, 3420 (1990) [Sov. Phys. Solid State **32**, 1980 (1990)].
[6] A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, *The Metal–Insulator Phase Transition and Its Application* [in Russian], Nauka, Leningrad (1979), p. 28.
[7] R. O. Zaĭtsev, E. V. Kuz'min, and S. G. Ovchinnikov, Usp. Fiz. Nauk **148**, 603 (1986) [Sov. Phys. Usp. **29**, 322 (1986)].
[8] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Vestnik Moskov. Univ. Ser. III Fiz. Astronom. **30**, 52 (1989).
[9] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Fiz. Tverd. Tela (Leningrad) **31**, No. 10, 261 (1989) [Sov. Phys. Solid State **31**, 1803 (1989)].
[10] V. I. Emel'yanov, N. L. Levshin, and A. L. Semenov, Vestnik Moskov. Univ. Ser. III Fiz. Astronom. **31**, 99 (1990).
[11] A. L. Semenov, Fiz. Tverd. Tela (St. Petersburg) **36**, 1974 (1994) [Phys. Solid State **36**, 1079 (1994)].
[12] V. I. Emel'yanov and A. L. Semenov, Fiz. Tverd. Tela (Leningrad) **32**, 3083 (1990) [Sov. Phys. Solid State **32**, 1790 (1990)].
[13] V. I. Emel'yanov, N. L. Levshin, S. Yu. Poroĭkov, and A. L. Semenov, Vestnik Moskov. Univ. Ser. III Fiz. Astronom. **32**, 63 (1991).
[14] K. F. Berggren and V. A. Huberman, Phys. Rev. B **18**, 3369 (1978).
[15] A. A. Bugaev, B. P. Zakharchenya, and F. A. Chudnovskiĭ, JETP Lett. **33**, 629 (1981).
[16] A. A. Bugaev, V. V. Gudyalis, B. P. Zakharchenya *et al.*, JETP Lett. **34**, 430 (1981).
[17] H. Gibbs, *Optical Bistability: Controlling Light with Light*, Academic Press, New York (1985).
[18] P. I. Khadzhi, G. D. Shibarshina, and A. Kh. Rotaru, *Optical Bistability in a System of Excitons and Biexcitons in Semiconductors* [in Russian], Shtiintsa, Kishinev (1988), p. 120.
[19] A. V. Andreev, V. I. Emel'yanov, and Yu. A. Il'inskiĭ, *Cooperative Phenomena in Optics: Superradiation, Bistability, and Phase Transitions* [in Russian], Nauka, Moscow (1988), p. 256.
[20] O. Madelung, *Introduction to Solid State Physics*, Springer, Berlin (1978).
[21] N. N. Bogolyubov and N. N. Bogolyubov, Jr., *Introduction to Quantum Statistical Mechanics* [in Russian], Nauka, Moscow (1984), p. 282.
[22] N. R. Belashenkov, V. B. Karasev, A. A. Solunin, I. A. Khakhaev, K. Sh. Tsibadze, and F. A. Chudnovskiĭ, Fiz. Tverd. Tela (St. Petersburg) **36**, 2475 (1994) [Phys. Solid State **36**, 1347 (1994)].
[23] V. L. Bonch-Bruevich and S. G. Kalashnikov, *Semiconductor Physics* [in Russian], Nauka, Moscow (1977), p. 255.
[24] S. A. Akhmanov, Yu. E. D'yakov, and A. S. Chirkin, *Introduction to Statistical Radiophysics and Optics* [in Russian], Nauka, Moscow (1981), p. 42.
[25] D. N. Klyshko, *Physical Bases of Quantum Electronics* [in Russian], Nauka, Moscow (1986), p. 22.
[26] A. S. Davydov, *Solid-State Theory* [in Russian], Nauka, Moscow (1976), p. 296.
[27] V. A. Stadnik, Fiz. Tverd. Tela (Leningrad) **29**, 3594 (1987) [Sov. Phys. Solid State **29**, 2059 (1987)].
[28] V. M. Lyubin and V. K. Tikhomirov, JETP Lett. **55**, 23 (1992).

Translated by Eugene Yankovsky

Model of two systems with random *P*-spin interaction and a fixed correlation between their spin configurations

A. É. Allakhverdyan and D. B. Saakyan

Erevan Physical Institute, 375036 Erevan, Armenia; Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

(Submitted 17 April 1996)
Zh. Éksp. Teor. Fiz. **111**, 1410–1425 (April 1997)

The binary spins of each subsystem interact within the confines of its subsystem. The scalar product (coupling) of the spin configurations of the two subsystems is fixed. Nontrivial interference develops between the subsystems when (for certain values of the parameters) they can only slide into the spin-glass phase. © *1997 American Institute of Physics.* [S1063-7761(97)02104-5]

1. INTRODUCTION

The Derrida model[1] for *P*-spin glass is the simplest system that can be solved exactly for a single breakdown of replication symmetry.[2] It can be also used for optimum encoding.[3–6] In this paper we consider a system consisting of two subsystems with a *P*-spin interaction, having the ± 1 spin sets σ_i^1, $i=1,\ldots,N$ and σ_i^2, $i=1,\ldots,N$. Here the restriction

$$\frac{1}{N}\sum_i \sigma_i^1\sigma_i^2 = C \tag{1}$$

is imposed (where C is a parameter). The general hamiltonian has the form

$$H = -\sum_{1\leqslant i_1<i_2\ldots<i_P\leqslant N} (\sigma_{i_1}^1\ldots\sigma_{i_P}^1 j_{i_1\ldots i_P}^1 + \sigma_{i_1}^2\ldots\sigma_{i_P}^2 j_{i_1\ldots i_P}^2)\sqrt{\frac{N}{C_N^P}}, \tag{2}$$

where the C_N^P are binomial coefficients and j^1 and j^2 are gaussian random quantities with zero mean and dispersions of 1/2 and $J^2/2$, respectively. We assume $J\leqslant 1$. The parameter C in Eq. (1) varies over the range $0\leqslant C\leqslant 1$. For $C=0$ we have two independent systems. At high temperatures both subsystems will be in a paramagnetic phase. Then, as the temperature is lowered, the first system undergoes a transition into a spin-glass phase, while the second remains in the paramagnetic phase. Finally, a second phase transition takes place and the second subsystem also enters the spin-glass state. This sort of qualitative picture is retained in our model for nonzero (subcritical) values of C.

Here we encounter a situation in which the system can be solved exactly in the second order of the breakdown of replication symmetry. For $C=1$ we return to the usual Derrida model, in which

$$\langle (j_{i_1\ldots i_P}^1)^2\rangle = \frac{1+J^2}{2}. \tag{3}$$

In this system there is only one phase transition and a single breakdown of the replication symmetry. Then one of the subsystems is in a ferromagnetic or spin-glass phase and the thermodynamic properties of the other subsystem will correspond to the usual Derrida model in which there are only

$$\exp\left[-N\left(\frac{1+C}{2}\ln\frac{1+C}{2}+\frac{1-C}{2}\ln\frac{1-C}{2}\right)\right] \tag{4}$$

configurations, rather than 2^N. The thermodynamic properties of the Derrida model can be calculated in accordance with the following principle: perform the high-temperature expansion at high temperatures and calculate the free energy for the paramagnetic phase. Then continue the expression for the free energy from the temperature at which the entropy goes to zero to zero temperature. In the following we calculate the free energy rigorously by the average field method. All our results based on the above simple principles are in agreement with what we expected.

In the first five sections we investigate the model using the average field method and in the last two we derive these results by a method that is applicable to the random-energy model.

2. PARAMAGNETIC PHASE

Solving the model (1)–(3) using replicas, we obtain the following expression for the partition function Z^n:

$$\langle Z^n\rangle = \int_{-i\infty}^{i\infty}\prod_{\alpha<\beta} dQ_{\alpha\beta}\,\delta\left(C-\sum_{i=1}^N \sigma_i^1\sigma_i^2\right)\delta\left(Q_{\alpha\beta}^1 - \frac{1}{N}\sum_{i=1}^N \sigma_{\alpha i}^1\sigma_{\beta i}^1\right)\delta\left(Q_{\alpha\beta}^2-\frac{1}{N}\sum_{i=1}^N \sigma_{\alpha i}^2\sigma_{\beta i}^2\right)\times\exp\left\{Nn\,\frac{B^2}{4}\left[\sum_{\alpha,\beta}\sum_{s=1}^2 (J_s)^2 (Q_{\alpha\beta}^s)^P\right]\right\}, \tag{5}$$

where B is the reciprocal temperature, $Q_{\alpha,\beta}$ are the ordering parameters, and

$$J_1=1, \quad J_2=J. \tag{6}$$

On introducing the integral representation for the δ-function, by analogy with Ref. 2 we obtain

$$\langle Z^n\rangle = \int_{-i\infty}^{i\infty}\frac{dk}{2\pi}\prod_{\alpha<\beta}\int_{-i\infty}^{i\infty}\frac{N dQ_{\alpha\beta}d\lambda_{\alpha\beta}}{2\pi}\times\exp\left\{N\left\{\frac{B^2}{4}\left[(1+J^2)n+\sum_{\alpha\neq\beta}(Q_{\alpha\beta}^1)^P\right.\right.\right.$$

 1063-7761/97/040780-08$10.00

$$+J^2(Q^2_{\alpha\beta})^P\Big]-\sum_{s=1}^{2}\sum_{\alpha<\beta}Q^s_{\alpha\beta}\lambda^s_{\alpha\beta}\sum_{\alpha}CK_\alpha$$

$$+\ln\sum_{\sigma^1_\alpha\sigma^2_\alpha}\exp\Bigg[\sum_{s=1}^{2}\sum_{\alpha\neq\beta}\lambda^s_{\alpha\beta}\sigma^s_\alpha\sigma^s_\beta$$

$$-\sum_{\alpha}K_\alpha\sigma^1_\alpha\sigma^2_\beta\Bigg]\Bigg\}\Bigg\}. \tag{7}$$

In the paramagnetic phase we take

$$\lambda^s_{\alpha\beta}=\lambda^s,\quad Q^s_{\alpha\beta}=q^s,\quad K_\alpha=K. \tag{8}$$

Here the $\lambda_{\alpha\beta}$ are the Lagrangian multipliers in the expression for the δ-function. For $F=\langle\ln Z\rangle/N$ we obtain the expression

$$\frac{B^2(1+J^2)}{4}+\frac{B^2}{4}[q_1^P+J^2q_2^P]-KC+\frac{\lambda^1q^1+\lambda^2q^2}{2}$$

$$-\frac{\lambda^1+\lambda^2}{2}+\int_{-\infty}^{\infty}Dr\int_{-\infty}^{\infty}Ds\ \ln[2e^K\cosh(r\sqrt{\lambda^1}+s\sqrt{\lambda^2})$$

$$+2e^{-K}\cosh(r\sqrt{\lambda^1}-s\sqrt{\lambda^2})], \tag{9}$$

where

$$Dr=\frac{1}{\sqrt{2\pi}}\exp(-r^2/2),\quad Ds=\frac{1}{\sqrt{2\pi}}\exp(-s^2/2), \tag{10}$$

and r and s are the auxiliary variables in the Stratonovich transform. On differentiating Eq. (9) with respect to q_1 and q_2, we obtain the following equation for λ_1 and λ_2:

$$PB^2q_1^{P-1}/2=\lambda_1,\quad PJ^2B^2q_2^{P-1}/2=\lambda_2. \tag{11}$$

In the paramagnetic phase we must have $q_1<1$ and $q_2<1$. This implies that

$$\lambda_1=0,\quad \lambda_2=0. \tag{12}$$

The integrals are easily evaluated after substituting Eq. (12) in Eq. (9) and we obtain (setting $q_1=0$ and $q_2=0$)

$$F=B^2(1+J^2)/4-KC+\ln 4\cosh K, \tag{13}$$

where K satisfies the equation

$$C=\tanh K.$$

Finally, for the free energy we obtain

$$F=\frac{1}{N}\langle\ln Z\rangle=\frac{4(1+J^2)}{B^2}+[1+h(C)]\ln 2, \tag{14}$$

where

$$h(C)=\Bigg[-\frac{1+C}{2}\ln\frac{1+C}{2}-\frac{1-C}{2}\ln\frac{1-C}{2}\Bigg]/\ln 2. \tag{15}$$

When $C=0$ the free energy (14) is given by the sum of the free energies for two independent models. For $C=1$ we obtain the free energy for a single model with an effective constant.

3. SPIN-GLASS PHASE IN TWO SUBSYSTEMS

Let us examine the case where the replication symmetry fails once and both systems slide into a spin glass phase. As will become clear below, this happens when

$$J^2>h. \tag{16}$$

By analogy with Ref. 2, we introduce the variables λ_0^1, λ_1^1, λ_0^2, λ_1^2, q_0^1, q_0^2, q_1^1, and q_1^2, where the superscripts denote the number of the subsystem, and then for the free energy we obtain

$$F=\frac{B^2}{4}(1+J^2)+\frac{B^2}{4}[-m(q_0^1)^P+(m-1)(q_1^1)^P]$$

$$+\frac{B^2J^2}{4}[-m(q_0^2)^P+(m-1)(q_1^2)^P]+\frac{1}{2}[\lambda_0^1q_0^1m$$

$$-(m-1)\lambda_1^1q_1^1]+\frac{1}{2}[\lambda_0^2q_0^2m-(m-1)\lambda_1^2q_1^2]$$

$$-\frac{\lambda_1^1+\lambda_1^2}{2}+\frac{1}{m}\int Dr_0Ds_0\ln\int Dr_1Ds_1[2e^K$$

$$\times\cosh(\sqrt{\lambda_0^1}r_0+\sqrt{\lambda_1^1-\lambda_0^1}r_1+\sqrt{\lambda_0^2}s_0+\sqrt{\lambda_1^2-\lambda_0^2}s_1)$$

$$+2e^{-K}\cosh(\sqrt{\lambda_0^1}r_0-\sqrt{\lambda_1^1-\lambda_0^1}r_1+\sqrt{\lambda_0^2}s_0$$

$$-\sqrt{\lambda_1^2-\lambda_0^2}s_1)]^m, \tag{17}$$

where m is the replication symmetry breakdown parameter. Differentiating Eq. (17) with respect to q_0^1, q_1^1, q_0^2, and q_1^2 yields

$$\lambda_i^1=\frac{PB^2(q_i^1)^{P-1}}{2},\quad \lambda_i^2=\frac{PB^2(q_i^2)^{P-1}}{2}. \tag{18}$$

We again find that

$$\lambda_0^1\to 0,\quad \lambda_0^2\to 0, \tag{19}$$

and

$$\lambda_1^1\to\infty,\quad \lambda_1^2\to\infty. \tag{20}$$

Condition (20) implies that

$$F=\frac{B^2}{4}(1+J^2)m+\frac{1}{m}\ln 4\cosh Km-KC. \tag{21}$$

On differentiating Eq. (21) with respect to K and m, respectively, we obtain

$$\tanh Km=C, \tag{22}$$

and

$$\frac{B^2m^2}{4}(1+J^2)+\cosh(Km)\ln 4+Km\tanh Km=0. \tag{23}$$

Equations (22) and (23) have the following solutions:

$$\frac{B_c^2}{4}(1+J^2)=\ln 2[1+h(C)], \tag{24}$$

and

$$m=B_c/B, \tag{25}$$

where B_c is the critical value of the reciprocal temperature at which the phase transition takes place. Now we obtain the following expression for the free energy $F = \langle \ln Z \rangle / N$:

$$F = \frac{B}{B_c}\left[\frac{B_c^2}{4}(1+J^2) + \ln 2(1+h(C))\right]. \tag{26}$$

The spin-glass phase which we have found exists for

$$B > B_c, \quad J^2 > h(C). \tag{27}$$

The second of conditions (27) will be examined in the next section.

4. SPIN-GLASS PHASE WITH A DOUBLE FAILURE OF REPLICATION SYMMETRY AND THE SPIN-GLASS+PARAMAGNETIC MIXTURE

Let us examine the case in which the replication group for the first subsystem is broken down to an m_1 subgroup and that for the second, to m_2. Calculations similar to those in the previous section yield

$$\begin{aligned} F = {} & \frac{B^2}{4}(1+J^2) + \frac{B^2}{4} + (m_1-1)(q_1^1)^P + \frac{B^2J^2}{4}(m_2-1) \\ & \times (q_1^2)^P - \frac{\lambda_1^1+\lambda_1^2}{2} + \frac{1}{2}[(1-m_1)\lambda_1^1 q_1^1 + (1 \\ & - m_2)\lambda_1^2 q_1^2] + \frac{1}{m_2}\ln \int Dr_1 \{Ds_1 [2e^K \\ & \times \cosh(\sqrt{\lambda_1^1} r_1 + \sqrt{\lambda_1^2} s_1) + 2e^{-K} \\ & \times \cosh(\sqrt{\lambda_1^1} r_1 - \sqrt{\lambda_1^2} s_1)]^{m_2}\}^{m_2/m_1}. \end{aligned} \tag{28}$$

Here we have omitted the terms with q_0^1, q_0^2, λ_0^1 and λ_0^2, which are zero. In deriving Eq. (28) we have considered the case where the replication group initially breaks down into a subgroup of m_2 permutations and then into a subgroup of m_1. After integrating over ds_1 in Eq. (28), we arrive at an expression of the form

$$\begin{aligned} & \frac{1}{m_2}\ln \int Dr_1 [2\cosh(Km_1)]^{m_2/m_1} \{\theta(r_1)\exp[m_1\sqrt{\lambda_1^1} r_1 \\ & + \lambda_1^2 (m_1)^2] + \theta(-r_1)\exp[m_1\sqrt{-\lambda_1^1} r_1 \\ & + \lambda_1^2 (m_1)^2]\} = \frac{1}{m_2}\ln 2 + \frac{1}{m_1}\ln(2\cosh Km_1), \end{aligned} \tag{29}$$

where $\theta(x)$ is the Heaviside function. Substituting Eq. (29) in Eq. (28) and differentiating with respect to m_2 and m_1, we obtain

$$\frac{B^2 m_1^2}{4} - \ln 2 = 0, \tag{30}$$

and

$$\frac{J^2 B^2 m_1^2}{4} - h(C)\ln 2 = 0. \tag{31}$$

For the free energy we have

$$F = \frac{B}{B_1}\left[\frac{B_1^2}{4} + \ln 2\right] + \frac{B}{B_2}\left[\frac{J^2 B_1^2}{4} + h(C)\ln 2\right], \tag{32}$$

where

$$B_1 = \lambda_1 \sqrt{\ln 2}, \tag{33}$$

and

$$B_2 = \lambda_2 \sqrt{h(C)\ln 2}/J. \tag{34}$$

This solution exists for

$$B > B_2, \tag{35}$$

and

$$B_c > B_2 > B_1. \tag{36}$$

The latter inequality holds for

$$J^2 > h(C). \tag{37}$$

For

$$B_1 < B < B_2, \tag{38}$$

the first subsystem is in the spin-glass phase and the second, in the paramagnetic phase. In place of Eq. (28) we obtain

$$\begin{aligned} F = {} & \frac{1}{4}B^2(1+J^2) + \frac{1}{4}B^2J^2(m-1)q_1^P + (m-1)q_1^P - (m \\ & - 1)\lambda_1 q_1 - \frac{1}{2}\lambda_1 - KC + \frac{1}{m}\ln \\ & \times \int Dr \ln[(2\cosh k)^m 2\cosh(\sqrt{\lambda_1} m r)]. \end{aligned} \tag{39}$$

On choosing

$$q_1 = 1, \quad \lambda_1 \to \infty, \tag{40}$$

we obtain

$$F = \frac{1}{4}B^2J^2 + \frac{1}{4}B^2 m + \frac{1}{m}\ln 2 - KC + \ln\frac{1}{m}\cosh K. \tag{41}$$

Differentiating with respect to m and K, we obtain

$$\tanh K = C, \tag{42}$$

and

$$\frac{1}{4}B^2 m^2 = \ln 2. \tag{43}$$

Finally,

$$F = \frac{B}{B_1}\left[\frac{B_1^2}{4} + \ln 2\right] + \frac{B^2J^2}{4} + h(C)\ln 2, \tag{44}$$

where B_1 has been defined in Eq. (33). This regime exists when condition (38) is satisfied. Equation (44) can be treated as follows: when $h(C)$ is large enough (close to 1), the first subsystem slides into the spin-glass phase, as happens in the usual Derrida model. Since the entropy is zero in the spin-glass phase, there are exactly $\exp[Nh(C)]$ states in the first subsystem (in the paramagnetic phase). When the entropy vanishes in Eq. (44), the system slides into a spin-glass phase with a free energy given by Eq. (32).

5. FERROMAGNETIC PHASE WITH STRONG SUBSYSTEM CORRELATION

Thus far we have considered only a symmetric distribution for the coupling constants. We now consider the hamiltonian

$$H=-\sum_{1\leqslant i_1<i_2\ldots<i_P\leqslant N}\sigma^1_{i_1}\ldots\sigma^1_{i_P}(\sqrt{N/C_N^P}j^1_{i_1\ldots i_P}+J_0^1N/C_N^P)+\sigma^2_{i_1}\ldots\sigma^2_{i_P}(j^2_{i_1\ldots i_P}\sqrt{N/C_N^P}+J_0^2N/C_N^P) \tag{45}$$

with the coupling (1). The constants $j^1_{i_1\ldots 1_P}$ and $j^2_{i_1\ldots 1_P}$ have the distribution (3). We assume that condition (1) is satisfied for $\{\sigma_i^1\}=\{1\}$ and $\{\sigma_i^2\}=\{\xi_i\}$, where $\{\xi_i\}$ is a given configuration of the spins in the second subsystem. We shall be interested in the spin-glass–ferromagnetic transition under the condition

$$\frac{1}{N}\sum_{i=1}^{N}\xi_i=C. \tag{46}$$

In a situation where both systems are in the ferromagnetic phase, the average field method gives the following expression for the free energy:

$$F=(J_0^1+J_0^2)B. \tag{47}$$

We first consider the case of strong correlation between the subsystems and then, the case of weak correlation.

In order for the solution of Eq. (47) to be thermodynamically favorable, we must have

$$J_0^1+J_0^2>B_c(1+J^2)/2=\sqrt{\ln 2}[1+h(C)]. \tag{48}$$

This condition is necessary, but not sufficient. In principle, a situation can arise in the system where one of the subsystems is in a ferromagnetic phase and the other in a spin-glass phase. Let the first subsystem be in a ferromagnetic and the second in a spin-glass phase. Then,

$$F=\frac{1}{4}B^2(1+J^2)+J_0^1s^P+\frac{1}{4}J^2(m-1)(q_1^2)^P-(m-1)\lambda_1^2q_1^2/2-st-KC-\lambda_1^2/2-\frac{1}{4}(q_0^1)^PB^2J^2+\frac{\lambda_0^1q_0^1}{2}-\frac{\lambda_0^1}{2}+\frac{1}{m}\int_{-\infty}^{\infty}Dz_0\ln\int_{-\infty}^{\infty}Dr_1[2e^K\cosh(r_1\sqrt{\lambda_1^2}+t+z_0\sqrt{\lambda_0^2})+2e^{-K}\cosh(r_1\sqrt{\lambda_1^2}-t-z_0\sqrt{\lambda_0^2})]^m, \tag{49}$$

where s is the magnetization of the first subsystem and t is the conjugate variable. We leave out the term J_0^2 since it does not contribute to the free energy in the thermodynamic limit, when the magnetization (in the second subsystem) is equal to zero. Differentiating with respect to s, q_0^1, and q_1^2 at $s=1$, $q_0^1=1$, and $q_1^2=1$ gives

$$t\to\infty,\quad \lambda_0^1\to\infty,\quad \lambda_1^2\to\infty. \tag{50}$$

Transforming the integral in Eq. (49), we obtain

$$\frac{1}{m}\int Dz_0\ln\left\{\exp\left[mt+m\sqrt{\lambda_0^1}z_0+\frac{1}{2}\lambda_2^1m^2\right]2\cosh km\right\}=\frac{1}{m}\ln 2\cosh km+\frac{1}{2}\lambda_2^1m+t. \tag{51}$$

For the free energy we obtain

$$F=\frac{1}{4}B^2J^2m+\frac{1}{m}\ln 2\cosh Km-KC+J_0^1B, \tag{52}$$

where

$$\tanh Km=C,$$

$$Bm=\frac{2}{J}\sqrt{h(C)\ln 2}. \tag{53}$$

Substituting Eq. (53) in Eq. (52), we obtain

$$F=J_0^1B+BJ\sqrt{h(C)\ln 2}. \tag{54}$$

Thus, along with Eq. (48) we have the condition

$$J_0^2>J\sqrt{h(C)\ln 2}. \tag{55}$$

If the first subsystem underwent a phase transition, then we would obtain the condition

$$J_0^1>J\sqrt{h(C)\ln 2}. \tag{56}$$

The inequalities (48), (55), and (56) have two types of solutions:

$$\begin{cases}J_0^1>\sqrt{h(C)\ln 2},\\ J_0^2>\sqrt{(1+j^2)h(C)\ln 2}-J_0^1\end{cases} \tag{57}$$

and

$$\begin{cases}J_0^2>\sqrt{h(C)\ln 2}/J,\\ J_0^1>\sqrt{(1+j^2)/h(C)\ln 2}-J_0^2.\end{cases} \tag{58}$$

Here the following condition must be satisfied:

$$J^2>h(C). \tag{59}$$

If, on the other hand, the first inequalities in the systems (57) or (58) are violated, then the system shifts into a mixed phase where one subsystem is in a ferromagnetic phase and the other is in a spin-glass phase.

We now consider a ferromagnetic phase with a weak correlation between the subsystems. The expression for the free energy in a mixed (ferromagnet+spin glass) phase has the same form as with strong correlation. Instead of Eq. (48) we have the condition

$$J_0^2+J_0^1>\sqrt{h(C)\ln 2}/J+\sqrt{\ln 2}. \tag{60}$$

We again obtain a pair of solutions,

$$\begin{cases}J_0^1>\sqrt{h(C)\ln 2},\\ J_0^2>\sqrt{h(C)\ln 2}/J+\sqrt{\ln 2}-J_0^1\end{cases} \tag{61}$$

and

$$\begin{cases}J_0^2>\sqrt{h(C)\ln 2}/J,\\ J_0^1>\sqrt{h(C)\ln 2}/J+\sqrt{\ln 2}-J_0^2.\end{cases} \tag{62}$$

This occurs for

$$J^2 < h(C). \tag{63}$$

6. PROCEDURE FOR THE RANDOM-ENERGY MODEL

Let us consider two sets of 2^N levels with energies E^1_α, E^2_β with $1 \leqslant \alpha, \beta \leqslant 2^N$. The energy of the system is given by $E = E^1_\alpha + E^2_\beta$. Not all pairs of levels E^1_α, E^2_β are allowed. We introduce the coupling matrix $C_{\alpha\beta}$ whose elements equal 1 for the allowed pairs (E^1_α, E^2_β) and 0 for the forbidden pairs.

The levels with energies E^1_α, E^2_β have the following distributions:

$$P_x(E^1_\alpha) = \frac{1}{\sqrt{\pi N}} \exp\left[-\frac{(E^1_\alpha)^2}{N} \right], \tag{64}$$

and

$$P_y(E^2_\beta) = \frac{1}{\sqrt{\pi N J}} \exp\left[-\frac{(E^2_\beta)^2}{NJ^2} \right]. \tag{65}$$

Thus, the free energy is equal to

$$\langle \ln Z \rangle = \int \prod_{\alpha=1}^{M} dE^1_\alpha dE^2_\beta \ln\left[\sum_{\alpha,\beta=1}^{M} C_{\alpha\beta} \times \exp[E^1_\alpha + E^2_\beta] B \right] P_x(E^1_\alpha) P_y(E^2_\beta), \tag{66}$$

where $M = 2^N$. For the matrix $C_{\alpha\beta}$ we have the restriction

$$\sum_{\beta=1}^{M} C_{\alpha\beta} = e^{Nh \ln 2}, \quad \alpha = 1, \ldots, M. \tag{67}$$

Using the formula

$$\ln Z = \Gamma'(1) - \int_0^\infty \ln Z de^{-iZ} \equiv \Gamma'(1) - \int_{-\infty}^{\infty} u de^{-e^u Z} \tag{68}$$

from Ref. 1 (Γ' is the derivative of the Euler function), we obtain

$$\langle \ln Z \rangle = \Gamma'(1) - \int_{-\infty}^{\infty} u d\Psi(u), \tag{69}$$

where

$$\Psi(u) = \int_{-\infty}^{\infty} \prod_{\alpha=1}^{M} \frac{dx_\alpha dy_\alpha}{\pi} \exp\left\{ -\sum_{\alpha=1}^{M} (x^2_\alpha + y^2_\beta) - e^u \sum_{\alpha,\beta=1}^{M} C_{\alpha\beta} e^{\lambda_1 x_\alpha + \lambda_2 y_\beta} \right\}, \tag{70}$$

$\lambda_1 = N\sqrt{B}$, and $\lambda_2 = N\sqrt{B}J$. In the first approximation the monotonic function $\Psi(u)$ has the form of a step (with its center at the point u_0 which we shall define below), so that Eq. (68) yields

$$\ln Z \simeq -u_0. \tag{71}$$

For $u > -u_0$ the function $\Psi(u)$ falls off exponentially, while for $u < -u_0$ it equals 1 to within an exponential factor. Depending on the values of λ_1 and λ_2, different asymptotic forms will be obtained for the following similar function:

$$\Psi_1(u) = \int_{-\infty}^{\infty} \frac{dx\, dy}{\pi} \exp\{-(x^2+y^2) - e^u(e^{\lambda_1 x + \lambda_2 y})\}. \tag{72}$$

When $|u| \to \infty$, we have two cases.

In the first case we can expand the last term in the exponential and then

$$\Psi_1(u) \simeq 1 - \exp[u + (\lambda_1^2 + \lambda_2^2)/4]. \tag{73}$$

The main contribution is from the region near the point

$$x = \frac{\lambda_1}{2}, \quad y = \frac{\lambda_2}{2}. \tag{74}$$

In order for the calculations to be self consistent, the value of $\exp[u + (\lambda_1 x + \lambda_2 y)]$ at the point determined by Eq. (74) should be small. This condition is satisfied when

$$u < -(\lambda_1^2 + \lambda_2^2)/2. \tag{75}$$

Proceeding similarly with all the degrees of freedom, we obtain

$$\Psi(u) \simeq \exp\left\{ -\exp\left[u + \frac{\lambda_1^2 + \lambda_2^2}{4} + (1+h)N \ln 2 \right] \right\}. \tag{76}$$

Using Eq. (70), from this we have

$$\frac{\langle \ln Z \rangle}{N} = \frac{\lambda_1^2 + \lambda_2^2}{4} + (1+h)\ln 2 \equiv (1+J^2)\frac{B^2}{4} + (1+h)\ln 2. \tag{77}$$

Equation (75) gives

$$B < B_c \equiv 2\sqrt{\frac{(1+h)\ln 2}{1+J^2}}. \tag{78}$$

We have thus determined the critical temperature for the paramagnet–spin-glass transition.

We now consider the second case, where $\exp(u + \lambda_1 x + \lambda_2 y)$ is large compared to 1 in the region which makes the main contribution to the integral. Then,

$$\Psi(u) = \int_D \prod_{\alpha=1}^{M} dx_\alpha dy_\alpha \exp\left[-\sum_{\alpha=1}^{M} (x^2_\alpha + y^2_\alpha) \right], \tag{79}$$

where the convex region D is defined by the condition

$$\lambda_1 x_\alpha + \lambda_2 y_\beta < -u. \tag{80}$$

The value of the integral (79) is determined by the point on the boundary of region D at which $x^2_\alpha + y^2_\beta$ has a minimum. For this point we obtain

$$x_c = -\frac{\lambda_1 u}{\lambda_1^2 + \lambda_2^2}, \quad y_c = -\frac{\lambda_2 u}{\lambda_1^2 + \lambda_2^2}. \tag{81}$$

For one pair of configurations we obtain (omitting the preexponential factor)

$$\Psi_1(u) \simeq 1 - \exp[-(x_c^2 + y_c^2)]. \tag{82}$$

Similarly,

$$\Psi(u) \simeq \exp\left\{ -\exp\left(\frac{-u^2}{\lambda_1^2 + \lambda_2^2} + N(1+h)\ln 2 \right) \right\}. \tag{83}$$

For $B > B_c$ this yields

$$\langle \ln Z\rangle \simeq N\sqrt{(\lambda_1^2+\lambda_2^2)(1+h)\ln 2} = BN\sqrt{(1+h)(1+J^2)\ln 2}. \tag{84}$$

In general, for $\Psi_1(u)$ we have the following expression:

$$\Psi_1(u)\approx \int_{-\infty}^{\infty} dx\,\frac{e^{-x^2}}{\sqrt{\pi}}\left\{1-\theta\left(-u-\lambda_1 x-\frac{\lambda_2^2}{2}\right)\times\exp[u+\lambda_1+\lambda_2^2/4]-\theta\left(u+\lambda_1 x+\frac{\lambda_2^2}{2}\right)\times\exp[-(u+\lambda_1 x)^2/\lambda_2^2]\right\}. \tag{85}$$

Taking the integral over the variable x_α alone and over the 2^{Nh} variables y_β associated with it, we find that the expression in the curly brackets of Eq. (85) is of degree 2^{Nh}. This cuts off the region of integration in Eq. (85). Depending on the values of λ_1 and λ_2, one of the correction terms predominates in Eq. (85). As a result, we have two regimes:

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{x_1} e^{-x^2}dx \simeq 1-\frac{e^{-x_1^2}}{2x_1\sqrt{\pi}},$$

$$x_1=\left(\frac{\lambda_2^2}{4}+Nh\,\ln 2+u\right)\frac{1}{\lambda_1} \tag{86}$$

or

$$\frac{1}{\sqrt{\pi}}\int_{-\infty}^{x_2} e^{-x^2}dx \simeq 1-\frac{e^{-x_2^2}}{2\sqrt{\pi}x_2},\quad x_2=\frac{\lambda_2\sqrt{Nh\,\ln 2}+u}{\lambda_1}. \tag{87}$$

Raising Eqs. (86) and (87) to the 2^N-th power, we obtain the corresponding expressions for $\Psi(u)$:

$$\Psi(u)\simeq\exp\left\{-\exp\left[\frac{-(u+Nh\,\ln 2+\lambda_2^2/4)^2}{\lambda_1^2}+N\,\ln 2\right]\right\} \tag{88}$$

or

$$\Psi(u)\simeq\exp\left\{-\exp\left[-\frac{(u+\lambda_2\sqrt{Nh\,\ln 2})^2}{\lambda_1^2}+N\,\ln 2\right]\right\}. \tag{89}$$

From this, we find that for $2\sqrt{\ln 2} < B < 2\sqrt{h\,\ln 2}/J$,

$$\frac{\langle \ln Z\rangle}{N}=B\sqrt{\ln 2}+\frac{B^2J^2}{4}+h\,\ln 2, \tag{90}$$

and for $B > 2\sqrt{h\,\ln 2}/J$

$$\frac{\langle \ln Z\rangle}{N}=B\sqrt{\ln 2}\,\frac{B^2y^2}{4}+h\,\ln 2. \tag{91}$$

The solutions (90) and (91) exist for

$$h > J^2, \tag{92}$$

while for

$$h < J^2 \tag{93}$$

the solution is Eq. (84). At high temperatures we have the paramagnetic solution (77). Subsequently, when the condition (93) is satisfied and $B = B_c$, the system goes into the ferromagnetic phase (84). In the case of Eq. (92), we have a (spinglass+paramagnet) mixture and the spin glass.

7. A STUDY OF THE FERROMAGNETIC PHASE USING THE RANDOM-ENERGY MODEL

Now for each of the subsystems we have one ''ferromagnetic'' configuration apiece. Instead of Eqs. (64) and (65) for E_1^1 and E_1^2 we can write

$$P_1(E_1^1)=\frac{1}{\sqrt{\pi N}}\exp\left[-\frac{1}{N}\,(E_1^1+J_0^1N)^2\right], \tag{94}$$

and

$$P_2(E_1^2)=\frac{1}{\sqrt{\pi NJ}}\exp\left[-\frac{1}{J^2N}\,(E_1^2+J_0^2N)^2\right]. \tag{95}$$

For the other E_α^1 and E_α^2 with $\alpha \geqslant 2$, the old distributions (64) and (65) remain valid. The constants J_0^1 and J_0^2 are positive. Then, for $\Psi(u)$ we obtain

$$\Psi(u)=\int\prod_{\alpha=1}^{M}\frac{dx_\alpha dy_\alpha}{\sqrt{\pi}}\exp\left\{-\sum_{\alpha=1}^{M}(x_\alpha^2+y_\alpha^2)-\left\{\exp[u+u_1+u_2+\lambda_1x_1+\lambda_2y_1]-\sum_{\beta=2}^{M}C_{1\beta}\exp[u+u_1+\lambda_1x_1+\lambda_2y_\beta]-\sum_{\alpha=2}^{M}C_{\alpha1}\exp[u+u_2+\lambda_1x_\alpha+\lambda_2y_1]-\sum_{\alpha,\beta=2}^{M}C_{\alpha\beta}\exp[u+\lambda_1x_\alpha+\lambda_2y_\beta]\right\}\right\}, \tag{96}$$

where

$$u_1=J_0^1BN,\quad u_2=J_0^2BN. \tag{97}$$

For the ferromagnetic phase the first term in Eq. (96) is the principal term. To lowest order we can neglect the other three terms, so that $\Psi(u)$ can be regarded as a step with its center at the point $u_0 = -(J_0^1 + J_0^2)N_B$ to the right of which $\Psi(u)$ falls off as

$$\Psi(u)\simeq\exp\left[-\frac{(u+u_1+u_2)^2}{\lambda_1^2+\lambda_2^2}\right]. \tag{98}$$

From this we obtain

$$\frac{\langle \ln Z\rangle}{N}=(J_0^1+J_0^2)N_B. \tag{99}$$

This solution is correct when it exceeds the corresponding expression for the spin-glass and mixed (one subsystem in the ferromagnetic and the other in the spin-glass state) phases. Mixed phases occur when the second or third terms in the curly brackets of Eq. (96) are dominant. Then

$$\frac{\langle \ln Z\rangle}{N}=J_0^1 N_B+B\sqrt{h\ \ln 2}J \tag{100}$$

or

$$\frac{\langle \ln Z\rangle}{N}=J_0^2 N_B+B\sqrt{h\ \ln 2}. \tag{101}$$

For $J^2 \leqslant h$ we obtain the following conditions for the existence of the solution (99):

$$\begin{cases} J_0^1+J_0^2>\sqrt{\ln 2}+\sqrt{h\ \ln 2}J, \\ J_0^1>\sqrt{h\ \ln 2}, \\ J_0^2>\sqrt{h\ \ln 2}J \end{cases} . \tag{102}$$

In the case of $h^2 \leqslant J^2$, on the other hand, the existence of the solution (99) is determined by the conditions

$$\begin{cases} J_0^1+J_0^2>\sqrt{(1+h)(1+J^2)\ln 2}, \\ J_0^1>\sqrt{h\ \ln 2}, \\ J_0^2>\sqrt{h\ \ln 2}J \end{cases} . \tag{103}$$

8. FINITE VOLUME EFFECTS

We are interest in corrections to the expression for the free energy which show up when the finite volume in the ferromagnetic phase is taken into account. When calculating the magnetization in the ferromagnetic phase, on examining the relative weight of the ferromagnetic configuration we arrive at the expression

$$\left\langle \frac{1}{Z}\, e^{-B(E_1^1+E_1^2)}\right\rangle=\left\langle e^{-B(E_1^1+E_1^2)}\int_0^\infty dt e^{-tZ}\right\rangle$$

$$=\int_{-\infty}^{\infty}\Psi(u)du, \tag{104}$$

where $\Psi(u)$ is defined in Eq. (96). If we neglect the last three terms in the exponent of Eq. (96), then the integral in Eq. (104) will be equal to 1. We denote the expression for $\Psi(u)$ in this case by $\Psi_2(u)$.

The correction terms in the exponent cause the integral in Eq. (104) to be replaced by an expression of the form

$$\int_{-\infty}^{-u_c}\Psi_2(u)\Psi_3(u)du, \tag{105}$$

where the cutoff point $-u_c$ is defined as the point where the expressions (100), (101), and (84) have a maximum for $h > J^2$ or the expressions (100), (101), and (91) have one for $h < J^2$. The function $\Psi_3(u)$ is defined by the principal correction term in Eq. (97). It has the form of a step centered at the point $-u_c$. When Eq. (84) is dominant, we obtain

$$\int_{-\infty}^{-u_c}\Psi_2(u)\left\{1-\exp\left[-\frac{u^2}{N\beta^2(1+y^2)}+N(1+h)\ln 2\right]\right\}$$

$$\times du\simeq 1-\int_{-u_c}^{-\infty}\Psi_2(u)du-\int_{-\infty}^{-u_c}\Psi_2(u)$$

$$\times\exp\left[-\frac{u^2}{N\beta^2(1+y^2)}+N(1+h)\ln 2\right]du. \tag{106}$$

$\Psi_2(u)$ equals 1 for all $u \leqslant -(J_0^1+J_0^2)/NB$ and falls off exponentially beyond that:

$$\Psi_2(u)\simeq\frac{1}{\sqrt{\pi}}\frac{NB^2(1+J^2)}{2|u+(J_0^1+J_0^2)NB|}$$

$$\times\exp\left\{-\frac{(u+(J_0^1+J_0^2)NB)^2}{NB^2(1+J^2)}\right\}. \tag{107}$$

There are two possibilities for evaluating Eq. (107) by the method of steepest descent. In the first case, the saddle point lies within the domain of integration and in the second, outside it. In the second case the integral is determined by the value of the exponent at the boundary of the region. Given this, for the magnetization m we obtain

$$m=1-a\ \exp\left\{\left(\frac{J_0^1+J_0^2-\sqrt{(1+J^2)(1+h)\ln 2}}{1+J^2}\right)^2 N\right\}$$

$$\times\theta\left[-\frac{J_0^1+J_0^2}{\sqrt{1+J^2}}+2\sqrt{(1+h)\ln 2}\right]-b$$

$$\times\exp\left\{-\frac{(J_0^1+J_0^2)^2N}{2(1+J^2)}+N(1+h)\ln 2\right\}$$

$$\times\theta\left(\frac{J_0^1+J_0^2}{\sqrt{1+J^2}}+2\sqrt{(1+h)\ln 2}\right). \tag{108}$$

This is valid under the conditions

$$\sqrt{(1+J^2)(1+h)\ln 2}>J_0^1+J\sqrt{h\ \ln 2}, \tag{109}$$

and

$$\sqrt{(1+J^2)(1+h)\ln 2}>J_0^2+J\sqrt{h\ \ln 2}. \tag{110}$$

When the inequalities (109) or (110) are not satisfied, we obtain

$$m=1-a\ \exp\left\{-\left(J_0^1-\sqrt{h\ \ln 2}\right)^2 N\right\}\theta(-J_0^1$$

$$+2\sqrt{h\ \ln 2})$$

$$-b\ \exp\{[-(J_0^1)^2/2+h\ \ln 2]N\}$$

$$\times\theta[J_0^1-2\sqrt{h\ \ln 2}] \tag{111}$$

for

$$\begin{cases} J_0^1+J\sqrt{h\ \ln 2}>\sqrt{(1+J^2)(1+h)\ln 2}, \\ J_0^1+J\sqrt{h\ \ln 2}>J_0^2+\sqrt{h\ \ln 2} \end{cases} \tag{112}$$

or

$$m=1-b\ \exp\left\{\left[-\frac{1}{2}\left(\frac{J_0^2}{J}\right)^2+h\ \ln 2\right]N\right\}$$

$$\times\theta\left[\frac{J_0^2}{J}-2\sqrt{h\ \ln 2}\right]-a\ \exp\left\{-\left(\frac{J_0^2}{J}\right.\right.$$

$$\left.\left.-\sqrt{h\ \ln 2}\right)^2 N\right\}\theta\left[-\frac{J_0^2}{J}+2\sqrt{h\ \ln 2}\right] \tag{113}$$

for

$$\begin{cases} J_0^2 + J\sqrt{h\ \ln 2} > \sqrt{(1+J^2)(1+h)\ln 2}, \\ J_0^2 + J\sqrt{h\ \ln 2} > J_0^1 + \sqrt{h\ \ln 2} \end{cases}. \tag{114}$$

Similar calculations can be done for the case $J^2 > h$. We see that for $T = 0$ there is a set of subphases differing by small corrections for the finite volume.

9. CONCLUSION

We have solved a system of two Derrida models with fixed coupling (scalar products) between the spin configurations and independent random constants. A solution was found both by the average field method and by using the technique developed by Derrida for the random-energy model. The two solutions are the same and this confirms the reliability of the (very tedious!) calculations. An interesting interference effect develops between the subsystems. A phenomenon of this sort should be observed with other models of spin glasses (spherical model, Kirkpatrick–Sherrington model). Similar models have been examined recently by Parisi's group (Italy) in studying the properties of a spin-glass phase. In their work, however, both subsystems have the same set of random constants.

The methods used in the last two sections of this paper can be generalized to the case of weak coupling (diminished coupling constants), where the sum in Eq. (2) is not taken over the entire set of indices $i,\ldots,i_P$, but only over α^N of them, chosen randomly from all the possible C_N^P possibilities. The physical picture does not change. Besides the physical interest of our results, they can be used for encoding many-terminal systems in the case of channels with multiple access. More complicated examples of problems[7] involving many-terminal systems such as broadcasting channels have not been solved. This suggests than the concepts and methods used for spin glasses may also be of use there.

The most difficult task in this paper was to derive the corrections to the expression for the free energy which arise when the finite volume in the ferromagnetic phase is taken into account (Section 8). For the case of a rarefied system these corrections are analogous to those we have found before.[8] For couplings α near critical, they correspond to the maximum possible suppression of finite volume effects (probability of decoding error). In the case of large α, on the other hand, a limiting expression for this exponentially vanishing (proportional to $\exp[-\gamma\alpha N]$) correction has not yet been found. The coefficient γ in the exponent of the expression for the corrections is evidently some sort of universal function (such as the critical indices for phase transitions). Calculating this function has a significance close to that of calculating the critical indices in real dimensionalities. This fact, along with the above-noted possibility of advances in solving problems involving multi-terminal systems, establishes the importance of research in this area, which lies at the frontier between statistical physics and information theory.

D. B. S. thanks Yu Lu for an invitation to Trieste and E. L. Zhidkov for hospitality at the LVTA at the Joint Institute for Nuclear Research (Dubna). We are grateful to E. Arutyunyan for clarifying the multi-terminal systems problem, S. Frants for discussions of the case of real replicas, and M. Virasoro for discussions of the spin-glass problem.

This work was partially supported by the Ministry of Science and Technology of the Federal Republic of Germany (Grant No. 211-5231).

[1] B. Derrida, Phys. Rev. Lett. **45**, 79 (1980).
[2] D. Gross and M. Mezard, Nucl. Phys. B **240**, 43 (1984).
[3] N. Sourlas, Nature **239**, 693 (1989).
[4] D. B. Saakyan, JETP Lett. **55**, 192 (1992).
[5] D. Rujan, Phys. Rev. Lett. **701**, 2968 (1993).
[6] H. Nishimori, Physica A **205**, 1 (1994).
[7] I. Chisar and Ya. Kerner, *Information Theory* [Russian translation], Mir, Moscow (1985).
[8] A. Allakhverdyan and D. Saakian, E-prints archive cond. matt/95.06.007.

Translated by D. H. McNeill

Simultaneous gamma, x-ray, and electron Mössbauer spectroscopy of the volume and surface magnetic structure of hexagonal *M*-ferrites

A. S. Kamzin*) and V. L. Rozenbaum

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

L. P. Ol'khovik

Kharkov State University, 310077 Kharkov, Ukraine
(Submitted 21 May 1996)
Zh. Éksp. Teor. Fiz. **111**, 1426–1437 (April 1997)

Direct comparative studies are made between the magnetic structures of a surface layer of thickness $\sim$ 40 nm and the bulk magnetic structure of ferromagnetic single crystals of hexagonal *M* ferrites ($BaFe_{12}O_{19}$, $SrFe_{12}O_{19}$, $PbFe_{12}O_{19}$) with a magneto-plumbite structure. Measurements are made by simultaneous gamma, x-ray, and electron Mössbauer spectroscopy in order to investigate the properties of the surface layer and the bulk crystal simultaneously. Experimental data obtained with a depth resolution of $\sim$ 10 nm show that the orientation of the magnetic moments of the iron ions (along the crystallographic *c* axis) does not change on approaching the surface from the crystal volume. Thus, to within an experimental error of $\sim$ 10 nm, single crystals of the hexagonal ferrites $BaFe_{12}O_{19}$, $SrFe_{12}O_{19}$, and $PbFe_{12}O_{19}$ with a ferromagnetic structure do not have a ''transition'' surface layer whose magnetic structure differs from that of the bulk crystal such as that which exists, with a depth of several hundred nm, in antiferromagnetic materials with weak ferromagnetism. © *1997 American Institute of Physics.* [S1063-7761(97)02204-X]

1. INTRODUCTION

Since the beginning of the 1970's there has been increasing interest in research on the magnetic properties of crystal surfaces. From a scientific standpoint, the study of the surface properties is important because further development of the new branch of solid state physics which has come to be known as ''surface magnetism'' requires an understanding of the relationships among the properties and phenomena at the surface and those observed in the interior of a crystal. From an applied standpoint, research on surface properties is important because promising microelectronic devices can only be made from very thin films and powders. Thin powders are also widely used as magnetic memory carriers, as catalysts, and in magnetic fluids. An understanding of the reasons for changes in the magnetization and coercive force of thin powders will make it possible to determine ways of creating new magnetic materials to serve as information carriers with high and superhigh recording densities. Many researchers have pointed out that the properties of the surface govern the properties of thin films and powders because of the substantial increase in the importance of the surface in these materials. Thus, studies of the properties of surfaces and of the nature of the formation of the crystalline and magnetic structures in the surface layer of crystals are of current interest.

Research on the properties of surfaces began with the 1954 paper by Néel[1] in which he proposed the existence of a surface anisotropy based on the fact that ions lying on the surface have a smaller number of covalent bonds than do the ions occupying sites within the crystal volume. In later studies of surface properties[2,3] it was predicted theoretically that magnetization might exist on the surface of magnetic materials at temperatures above the Curie or Néel points. Later, a similar result was obtained using various theoretical approaches for studying the role of the surface.[4]

The influence of the surface was invoked for the first time in interpreting experimental data on γ-Fe_2O_3 (Ref. 5) and γ-Fe_3O_4 (Ref. 6) to explain the fact that the saturation magnetization of thin powders is lower than in macroscopic samples. This experimental fact has now attracted the close attention of researchers for a quarter of a century. Various models have been used to explain it. Berkowitz *et al.*[5] conjectured that there is a nonmagnetic (magnetically dead) surface layer at the facets of crystallites. The thickness of this layer was estimated to be $\sim$ 6 Å.[5]

Thin powders and films are widely used for studying the properties of surfaces because of a lack of experimental methods that can distinguish signals from a thin surface layer and those from the crystal volume. In the case of thin powders, it is known that the smaller the volume of the crystallite the greater the importance of its surface. In addition, for example, in Mössbauer spectroscopy it is possible to enhance or attenuate the signal from a surface layer by doping it with the isotopes Fe-57 or Fe-56.

Based on data from Mössbauer studies of fine γ-Fe_2O_3 powder it was found[7] that the reason for the reduction in the saturation magnetization of the powder is a change in the magnetic structure of the surface of the crystallites. In order to explain experimental data from studies of thin powders of γ-Fe_2O_3,[8,9] $CrFe_2O_4$,[10] $NiFe_2O_4$,[11] CrO_2,[12] $Y_3Fe_{12}O_{19}$ and $Dy_2BiFe_5O_{19}$,[13] and $BaFe_{12}O_{19}$,[14] a ''shell'' model has been used which describes the magnetic properties of a particle, in particular its outer shell and inner part.[15] The magnetic structure of the inner part is similar or, perhaps, entirely identical to the structure of a bulk crystal, while the magnetic mo-

 1063-7761/97/040788-06$10.00

ments in the outer shell (i.e., on the surface) are arranged noncollinearly.[14,15]

A number of other authors, however, do not explain the experimental data in terms of a shell model. Thus,[16] it has been proposed that the spin noncollinearity in particles of γ-Fe_2O_3 does not exist only on the surface. Others[17] assume that in order to overcome the bulk anisotropy and completely order the magnetic moments, much stronger magnetic fields must be imposed collinearly with the external magnetic field than the fields used to test the shell model. Studies of the behavior of the magnetic moments in textured γ-Fe_2O_3 samples in strong magnetic fields[18] have shown that the reason for the incomplete ordering of the magnetic moments in a particle cannot be the large assumed[17] bulk magnetic anisotropy. The experimental data[18] obtained for γ-Fe_2O_3 were explained by the noncollinearity of the spin ordering.

Polarized neutron scattering data for $CoFe_2O_4$ particles have been explained by the existence within a particle of a magnetically ordered volume and a magnetically disordered shell.[19] It has been assumed[20] that on the surface of $NiFe_2O_4$ particles there is also a surface layer with angular ordering of the spins with a set of stable configurations, one of which is separated out upon cooling in a field and that at temperatures below 50 K this layer is converted into a spin-glass phase.

The question of the existence of a thin layer on the surface of ferrimagnetic crystals with magnetic properties which differ from those observed in the volume, therefore, remains open. Furthermore, studies of surface properties employing such objects as fine powders are quite complicated because a number of factors such as the variation in the particle size distribution, superparamagnetic phenomena, a strong dependence on fabrication techniques, etc., make it much harder to interpret the experimental data. In order to understand the nature of the phenomena on a surface and their relationship with bulk phenomena, it is necessary to study the properties of the surface layers of macroscopic crystals.

The first experimental studies of the surface properties of macroscopic crystals were carried out in 1972 by Krinchik *et al.*,[21] who showed that the surface anisotropy predicted in Ref. 1 could be observed in antiferromagnetic materials with weak ferromagnetism with easy plane magnetic anisotropy. This is because in this type of magnet (in contrast to ferromagnets) the energy of the demagnetizing field is low owing to the resulting small magnetic moment and there is essentially no magnetic anisotropy in the basis plane. This latter circumstances leads to an increased role for the surface anisotropy in weakly ferromagnetic crystals. This is what made it possible to observe surface magnetism in hematite.[21,22] Based on these data, it was proposed[22] that on a hematite surface there is a ''transition''[22] layer within which the orientation of the magnetic moments undergoes a smooth transition from their direction in the volume to their direction at the surface. A phenomenological theory was developed[22] to calculate the thickness of the transition layer, which was a few nm in the case of hematite.[22] A thin transition layer within which the orientation of the magnetic moments differs from that in the interior of the crystallites has also been observed in Mössbauer studies of thin hematite powders.[23]

A transition layer has been observed on the surface of macroscopic crystals of $FeBO_3$,[24,25] $ErFeO_3$,[26] and $TbFeO_3$,[27] which, like hematite, are antiferromagnetic with weak ferromagnetism. The thickness of the surface layer whose properties differ from those in the bulk material has been found experimentally to be $\sim$ 500 nm for $FeBO_3$,[24] i.e., ten times greater than the calculated value for hematite.[22] Magnetooptical data[25] of the surface layer of $FeBO_3$ and an analysis of these data using the theory of surface magnetism[22] also showed that the thickness of the transition layer in $FeBO_3$ is $\sim$ 500 nm.

Direct experimental data on the existence of a transition layer on the surface of macroscopic crystals have been found.[28] These data were obtained using a new method for simultaneous gamma, x-ray, and electron Mössbauer spectroscopy first proposed by Kamzin *et al.*[29] and described in detail elsewhere.[30] This technique makes it possible to extract information from the surface layer and the crystal interior simultaneously.

Simultaneous gamma, x-ray, and electron Mössbauer spectroscopy has been used to study the properties of the surface and interior of bulk antiferromagnetic Fe_3BO_6 crystals, which have weak ferromagnetism and belong, as does $FeBO_3$, to the borate family. It was shown experimentally that there is a transition layer on the surface of Fe_3BO_6 in which the orientation of the magnetic moments of the iron ions differs from the orientation of the magnetic moments of iron ions occupying sites within the sample volume.[28] Layer-by-layer studies revealed that the thickness of the transition layer is $\sim$ 400 nm. As the crystal surface is approached, the angular deviation of the magnetic moments of the iron ions from their orientation in the volume (the lower boundary of the transition layer) increases smoothly to an angle corresponding to the deviation of the magnetic moments of the iron ions lying on the surface (the upper boundary of the transition layer).[28]

Thus, the existence in antiferromagnetic crystals with weak ferromagnetism of a surface layer whose magnetic structure differs from that in the bulk crystal (referred to as a transition layer in Refs. 21 and 22) has been convincingly demonstrated. In this paper we study the properties of the surface layer of ferromagnetic crystals for the case of hexagonal *M* ferrites.

2. MÖSSBAUER STUDIES OF THE MAGNETIC STRUCTURE OF CRYSTALS

Mössbauer spectroscopy is so widely used for studying the magnetic properties of materials because almost all magnetic materials contain the iron ion used in the overwhelming majority of Mössbauer studies. Mössbauer spectroscopy provides direct information on the magnetic structure of a crystal. In fact, the magnetic moment (M) of the iron ion is always oriented antiparallel to the effective magnetic field at the nucleus (H_{eff}), whose direction, as will be shown below, is easily calculated from the ratio of the intensities of the Mössbauer spectrum lines.

The location of the hyperfine structure sublevels is determined by the ratio of the energies of the magnetic and electric interactions, the symmetry of the electric field gradi-

ent, and, in general, the direction of the internal magnetic field relative to the axes of the tensor of this gradient.[31] We assume for simplicity that the energy of the electric interaction in the crystal is negligibly small and that the levels split owing to the energy of the magnetic hyperfine interaction. In the Mössbauer spectrum of a thin absorber, the intensities of the components of the Zeeman sextuplet $A_1:A_2:A_3:A_4:A_5:A_6$ depend on the angle θ between the direction of propagation of the γ-photons and the orientation of the effective magnetic field H_{eff} at the nucleus of the iron ion (or the magnetic moment of the iron ion) as follows:

$$3(1+\cos^2\theta):4\sin^2\theta:(1+\cos^2\theta):(1+\cos^2\theta):$$
$$4\sin^2\theta:4(1+\cos^2\theta). \quad (1)$$

In this way, the formula relating the intensity of the second or fifth ($A_{2,5}$) lines of the Zeeman sextuplet associated with transitions with $\Delta m = 0$ to the intensity of the first or sixth ($A_{1,6}$) lines,[31]

$$A_{2,5}/A_{1,6}=4\sin^2\theta/3(\cos^2\theta+1), \quad (2)$$

can be used to find the angle θ which gives the orientation of the magnetic moments relative to the γ-ray beam from the Mössbauer spectrum of a single crystal:

$$\theta=\sin^{-1}\sqrt{\frac{(3/2)A_{2,5}/A_{1,6}}{1+(3/4)(A_{2,5}/A_{1,6})}}. \quad (3)$$

If the angle θ varies over the sample, then Eqs. (2) and (3) yield an average $\langle\theta\rangle$. Thus, Mössbauer spectroscopy can be used to determine the orientation of the magnetic moments in a crystal with rather high accuracy.

3. SIMULTANEOUS MÖSSBAUER STUDIES OF THE SURFACE AND VOLUME OF CRYSTALS

The simultaneous gamma, x-ray, and electron Mössbauer spectroscopy method proposed in Ref. 29 and described in Ref. 30 was used in the measurements. This method is based on the simultaneous detection of Mössbauer spectra at wavelengths with different mean free paths in the material, specifically γ-rays and the characteristic x-radiation of both conversion and Auger electrons, which carry information on the volume and on surface layers with thicknesses of a few μm and $\sim$ 300 nm of bulk crystals, respectively. The energy of an electron leaving the sample is lower when the atom which produced the electron lies deeper in the material. The electrons could be selected for energy by using simple proportional detectors.[32–34] Of course, the accuracy of the analysis of the layers by thickness is considerably poorer than with electron or magnetic separators,[35] but in a number of cases this accuracy was entirely sufficient. Thus, simultaneous gamma, x-ray, and electron Mössbauer spectroscopy with a proportional detector was used to perform layer-by-layer analysis of the properties of surface layers with thicknesses below 300 nm.[33]

The advantages of simultaneous gamma, x-ray, and electron Mössbauer spectroscopy are, first, that the spectra containing information on the properties of the surface and volume are taken simultaneously under identical conditions for the sample and, second, the spectral data containing the information on the surface and volume can be compared directly because they have been obtained using a single effect, namely the Mössbauer effect.

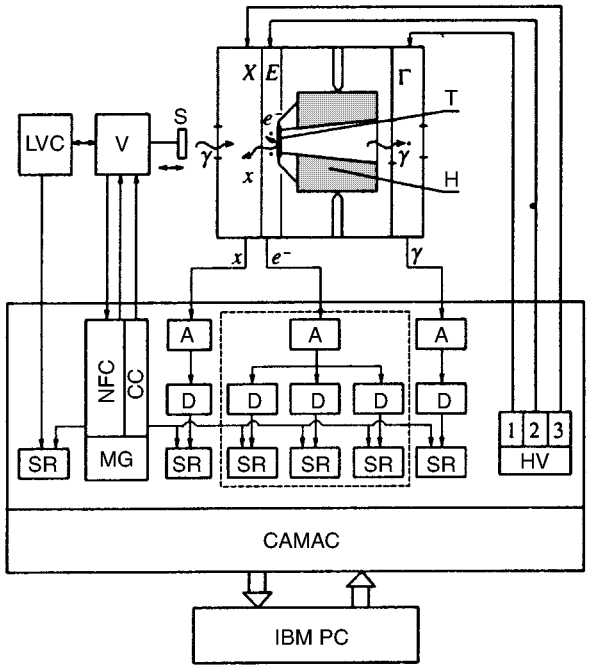

FIG. 1. Block diagram of the automated system for simultaneous gamma, x-ray, and electron Mössbauer spectroscopy. Γ, X, and E denote the detectors of γ-rays, characteristic x-rays, and secondary electrons, respectively. Other components are: (S) a γ-ray source, (T) the test sample, (H) a heater (oven), (V) a Doppler modulator (vibrator), (MG) a motion generator, (LVC) a laser velocity calibrator, (A) amplifiers, (D) discriminators, (SR) storage registers, (HV) high voltage supplies, (NFC) a negative feedback channel, and (CC) a computer correction channel.

Simultaneous gamma, x-ray, and electron Mössbauer spectroscopy was performed on the automated system illustrated in Fig. 1. The radiation (γ-photons, x-ray emission, and conversion and Auger electrons) is detected by counters labelled Γ, X, and E mounted in a universal three-chamber detector. The signals from the counters are fed to amplifiers (A) and then discriminators (D) to set the thresholds for distinguishing the corresponding energies; the resulting signals are then sent to corresponding storage registers (SR) for the Mössbauer spectra.

4. EXPERIMENTAL RESULTS

Studies were conducted on single crystals of $BaFe_{12}O_{19}$, $SrFe_{12}O_{19}$, and $PbFe_{12}O_{19}$. These compounds have several names, including hexagonal ferrites, M ferrites, or ferrites with a magnetoplumbate structure. The last term comes from the name of the mineral $PbFe_{12}O_{19}$.[36] The crystalline structure of these compounds can be regarded as a superposition of hexagonal and spinel units. The iron ions are distributed over five nonequivalent positions and the bulk magnetization of the M-ferrites is the vector sum of the magnetizations of the five sublattices. The hexagonal c axis is a magnetization easy axis, so that the magnetocrystalline anisotropy of these ferrites is very large.

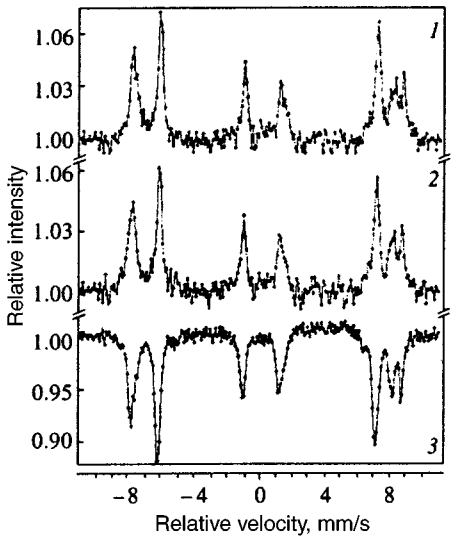

FIG. 2. Mössbauer spectra of Sr-*M* hexaferrite obtained at a temperature of 293 K with γ-ray detection (*3*), i.e., from the bulk of the crystal, with detection of the characteristic x-radiation from a surface layer several μm thick (*2*), and with detection of conversion and Auger electrons from a surface layer extending from 0 to 200 nm (*1*). The crystallographic *c* axis is oriented parallel to the wave vector of the γ-rays.

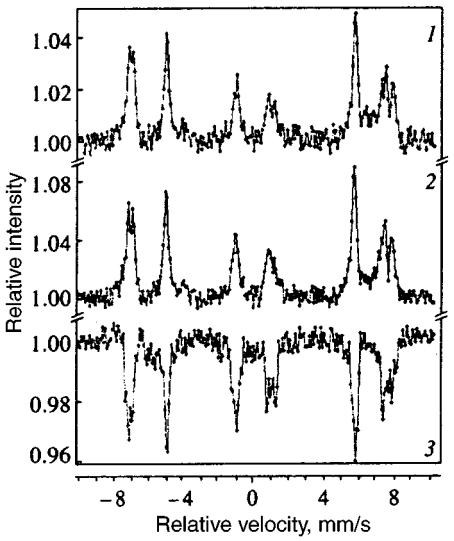

FIG. 3. Mössbauer spectra of Ba-*M* hexaferrite obtained at a temperature of 450 K with γ-ray detection (*3*), i.e., from the bulk of the crystal, and with detection of conversion and Auger electrons from surface layers extending from 50 to 100 nm (*2*) and from 0 to 40 nm (*1*). The crystallographic *c* axis is oriented parallel to the wave vector of the γ-rays.

Samples in the form of ~ 110 μm thick wafers with the crystallographic *c* axis perpendicular to the plane of the wafer were prepared for the measurements. This cut was chosen because at temperatures below the Curie point the magnetic moments are oriented along the *c* axis and if a beam of γ-rays is directed parallel to the crystallographic *c* axis then the second and fifth lines of the Zeeman sextuplet corresponding to transitions with $\Delta m = 0$ will be absent from the Mössbauer spectrum.[37] A deviation of the direction of the magnetic moments from that of the γ-ray beam causes the second and fifth lines of the sextuplet to appear.

Single crystals with a natural mirror facet were chosen for the experiments. Wafers were also subjected to chemical polishing by various methods, including (a) boiling in orthophosphoric acid, (b) holding for one minute in orthophosphoric acid at a temperature of 120 °C, and (c) holding in an $H_3PO_4+H_2SO_4$ acid mixture (1:1 proportions) heated to 90 °C for several minutes. The Mössbauer spectra were numerically processed to yield the intensities, linewidths of the Zeeman sextuplets, and the hyperfine interaction parameters.

Simultaneous gamma, x-ray, and electron Mössbauer spectroscopy yielded experimental spectra at temperatures of 300 to 700 K with the wave vector of the γ-rays oriented parallel to the crystallographic *c* axis. As an example, Fig. 2 shows spectra of $SrFe_{12}O_{19}$ at room temperature with detection of γ- and x-rays, as well as of conversion and Auger electrons escaping from a surface layer at least 200 nm thick.

Gamma-ray and conversion and Auger electron Mössbauer spectra of $BaFe_{12}O_{19}$ at 450 K and of $PbFe_{12}O_{19}$ at 473 K are shown in Figs. 3 and 4. In these figures, as opposed to Fig. 2, the spectra have been obtained by detecting electrons originating from a surface layer with thickness less than 100 nm (Fig. 3 curve *2*, and Fig. 4 curve *2*) or 40 nm (Fig. 3 curve *1* and Fig. 4 curve *1*). It should be noted that the γ-ray spectra of $SrFe_{12}O_{19}$, $BaFe_{12}O_{19}$, and $PbFe_{12}O_{19}$ (Fig. 2 curve *3*, Fig. 3 curve *3*, and Fig. 4 curve *3*) are similar to Mössbauer spectra taken previously[38–40] by the authors under the same experimental conditions. The spectrum lines are well resolved in these figures and this made it possible to calculate the angles θ which determine the direction of the magnetic moments relative to the wave vector of the γ-rays with high resolution using Eq. (3) and the intensity ratio of the first and second (fifth and sixth) lines of the sextuplets.

Data from an analysis of the γ-ray spectra obtained when the crystallographic *c* axis was oriented parallel to the wave vector of the γ-rays (Fig. 2 curve *3*, Fig. 3 curve *3*, and

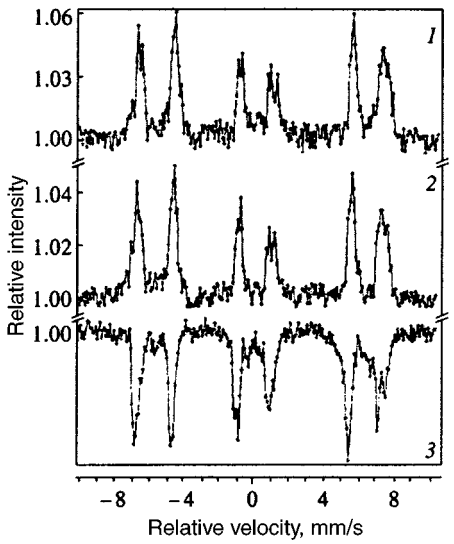

FIG. 4. Mössbauer spectra of Pb-*M* hexaferrite obtained at a temperature of 473 K with γ-ray detection (*3*), i.e., from the bulk of the crystal, and with detection of conversion and Auger electrons from surface layers extending from 50 to 100 nm (*2*) and from 0 to 40 nm (*1*). The crystallographic *c* axis is oriented parallel to the wave vector of the γ-rays.

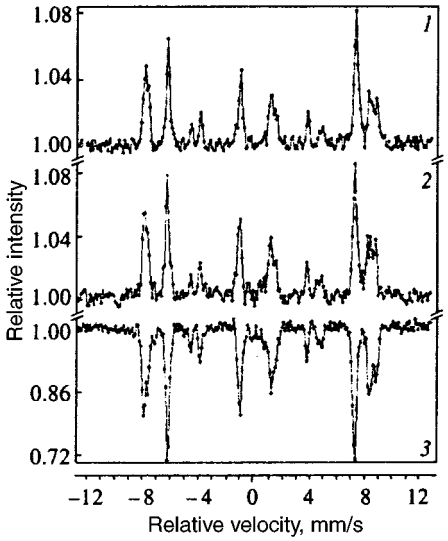

FIG. 5. Mössbauer spectra of Ba-*M* hexaferrite obtained at a temperature of 293 K with γ-ray detection (*3*), i.e., from the bulk of the crystal, and with detection of conversion and Auger electrons from surface layers extending from 50 to 100 nm (*2*) and from 0 to 40 nm (*1*). The crystallographic *c* axis is oriented at an angle $\theta = 25°$ to the wave vector of the γ-rays.

Fig. 4 curve *3*) showed that in the Zeeman sextuplets of each nonequivalent position, the intensities of the second and fifth lines, which correspond to transitions with $\Delta m = 0$, were zero. This means (see Eq. (3)) that the angle θ equals zero and, therefore, the magnetic moments of the iron ions at sites in the crystal volume are collinear with the wave vector of the γ-rays and parallel to the crystallographic *c* axis.

An analysis of the spectra taken by detecting conversion and Auger electrons when the crystallographic *c* axis was oriented parallel to the wave vector of the γ-rays (Fig. 2 curve *1*, Fig. 3 curves *1* and *2*, and Fig. 4 curves *1* and *2*) showed that the intensities of the second and fifth lines of the sextuplets are zero. This means that the magnetic moments of the iron ions occupying sites in a surface layer $\sim$ 200 nm thick layer (Fig. 2) are collinear with the wave vector of the γ-rays and the crystallographic *c* axis. A similar result was obtained for iron ions located in 100 and 40 nm thick surface layers (Figs. 3 and 4). This picture is seen up to temperatures of 600 K, beyond which it is difficult to determine θ owing to the poor resolution of the lines.

The following experiments were done as a control: a test single-crystal wafer was positioned so that the crystallographic *c* axis was at an angle θ with respect to the wave vector of the γ-rays. Figure 5 shows some examples of Mössbauer spectra for $\theta = 25° \pm 2°$. As can be seen from Fig. 5, the deviation of the magnetic moments from the propagation direction of the γ-rays causes the second and fifth lines of the Zeeman sextuplets, which correspond to transitions with $\Delta m = 0$, to appear in the γ-ray and conversion and Auger electron Mössbauer spectra. The angle θ given by Eq. (3) is 25° and is consistent with the specified experimental conditions.

For the purpose of analyzing the experimental data, we have modelled the Mössbauer spectra assuming that a surface layer with a different magnetic structure from the

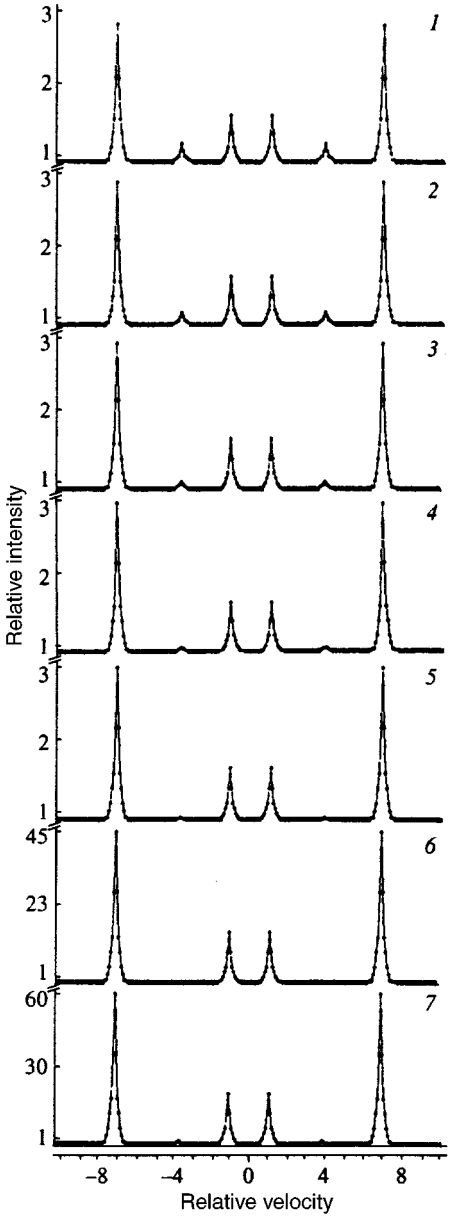

FIG. 6. Model Mössbauer spectra calculated for a surface layer extending from 0 to 40 nm with a 10 nm transition layer: (*1*) model spectrum from a layer extending from 0 to 2 nm with a deviation of $\theta = 25°$ for the magnetic moment, (*2*) from 2 to 4 nm, $\theta = 20°$, (*3*) from 4 to 6 nm, $\theta = 15°$, (*4*) from 6 to 8 nm, $\theta = 10°$, (*5*) from 8 to 10 nm, $\theta = 5°$, (*6*) from 0 to 40 nm, $\theta = 0$, (*7*) combined model spectrum of all layers from 0 to 40 nm.

bulk exists on the surface of these hexaferrites (and, by analogy with antiferromagnetic materials with weak ferromagnetism,[22] we refer to it as a transition layer). Let us assume that within this transition layer the magnetic moments veer continuously relative to the normal on approaching the crystal surface by an amount, ranging from 0° to 25°, determined for antiferromagnetic crystals with weak ferromagnetism.[32] Let us suppose that the transition layer has a thickness of 10 nm and the turning of the magnetic moments takes place in 2 nm layer steps, i.e., over five layers. The magnetic structure of the layers lying below the transition layer is similar to that in the bulk crystal. Figure 6 shows Mössbauer spectra constructed for a 40 nm thick layer as-

suming the existence of this sort of transition layer. Curves *1–5* of Fig. 6 show Mösssbauer spectra of 2 nm thick layers in which the magnetic moments deviate from the *c* axis by amounts ranging from 25° to 0°. The Mössbauer spectrum of a layer lying at a depth of from 10 to 40 nm below the surface of the model crystal is shown in curve *6* of Fig. 6 and the parallelism of the magnetic moments to the crystallographic *c* axis is indicated by the absence of the second and fifth lines of the Zeeman sextuplet from the spectrum. Curve *7* of Fig. 6 shows the model Mössbauer spectrum of a 40 nm thick surface layer obtained by summing (with appropriate weighting) the spectra of curves *1–6* of Fig. 6. As can be seen from curve *7* of Fig. 6, the second and fifth lines, which correspond to transitions with $\Delta m = 0$ essentially have zero intensities. Thus, an analysis of the model spectra showed that when Mössbauer spectroscopy is used to study a surface to a depth of $\sim$ 40 nm, the accuracy of the method is not adequate to establish the existence of a transition layer with a thickness of $\sim$ 10 nm. This implies that, in order to determine the presence or absence of a magnetic structure on the surface of the Pb-*M* hexaferrites which differs from the bulk structure, methods will have to be used which permit measurements on surface layers thinner than 20 nm.

To summarize, we have found experimentally that in a surface layer with a thickness of $\sim$ 40 nm on macroscopic hexagonal *M*-ferrites, specifically $BaFe_{12}O_{19}$, $SrFe_{12}O_{19}$, and $PbFe_{12}O_{19}$, the magnetic moments of the iron ions are oriented just as in the interior of the crystal, i.e., parallel to the crystallographic *c* axis. Thus, to within the limits of accuracy of the experiment ($\sim$ 10 nm), no transition surface layer of the sort which exists in antiferromagnetic crystals and having a thickness of several hundred nm[24,25,28] has been found in the *M*-hexaferrites ($BaFe_{12}O_{19}$,[41] $SrFe_{12}O_{19}$, and $PbFe_{12}O_{19}$).

This work was supported by the Russian Fund for Fundamental Research (Grant No. 95-02-04069).

*E-mail: KAMZIN@DIR.SHUV.IOFFE.RSSI.RU

[1] L. Neel, J. Phys. Radium **15**, 225 (1954).
[2] M. I. Kaganov and A. M. Omel'yanchuk, Zh. Éksp. Teor. Fiz. **61**, 1679 (1971) [Sov. Phys. JETP **34**, 895 (1972)].
[3] D. L. Mills, Phys. Rev. **3**, 3887 (1971).
[4] K. Binder, in: Phase Transitions and Critical Phenomena, Acad. Press, N.Y. (1983), Vol. 3, p. 325.
[5] A. E. Berkowitz, W. J. Schuele, and P. J. Flanders, J. Appl. Phys. **39**, 1261 (1968).
[6] L. Liebermann, D. R. Fridkin, and H. B. Shorre, Phys. Rev. Lett. **22**, 539 (1969); L. Liebermann, J. Clinton, D. M. Edwards, and J. Mathon, Phys. Rev. Lett. **25**, 232 (1970).
[7] J. M. D. Coey, Phys. Rev. Lett. **27**, 1140 (1971).
[8] A. H. Morrish, K. Haneda, and P. J. Schurer, J. de Physique, Colloque C6 **37**, C6-301 (1976).
[9] P. M. de Bakker, E. DeGrave, R. E. Vandenberghe, and L. H. Bowen, Hyperfine Interactions **54**, 493 (1990).
[10] A. E. Berkowitz, J. A. Lahut, and C. E. van Buren, IEEE Trans. Magn. **MAG-16**, 184 (1980).
[11] A. H. Morrish, and K. Haneda, J. Appl. Phys. **52**, 2496 (1981).
[12] K. Haneda, J. Kojima, A. H. Morrish, P. J. Picone, and K. Wakai, J. Appl. Phys. **53**, 2686 (1982).
[13] A. H. Morrish and K. Haneda, IEEE Trans. Magn. **MAG-25**, 2597 (1989).
[14] K. Haneda and A. H. Morrish, Nucl. Instrum. and Methods in Phys. Res. B **76**, 132 (1993).
[15] K. Haneda, Canad. J. Phys. **65**, 1233 (1987).
[16] F. T. Parker, M. W. Foster, D. Margulis, and A. E. Berkowitz, Phys. Rev. B **47**, 7885 (1993).
[17] Q. A. Pankhurst and P. J. Pollard, Phys. Rev. Lett. **67**, 325 (1991).
[18] P. V. Hendriksen, S. Linderoth, C. A. Oxborrow, and S. Morup, J. Phys. Cond. Matter **6**, 3091 (1994).
[19] D. Lin, A. C. Nunes, C. F. Majkrzak, A. E. Berkowitz, and M. B. Maple, J. Mag. Magn. Mat. **45**, 343 (1995).
[20] R. H. Kodama, A. E. Berkowitz, E. J. McNiff Jr., and S. Foner, Submitted to J. Mag. Magn. Mat. (1996).
[21] G. S. Krinchik, A. P. Khrebtov, A. A. Askochenskiĭ, and V. E. Zubov, JETP Lett. **17**, 335 (1973).
[22] G. S. Krinchik and V. E. Zubov, Zh. Éksp. Teor. Fiz. **69**, 707 (1975) [Sov. Phys. JETP **42**, 359 (1976)].
[23] A. M. van der Kraan, Phys. Stat. Sol. (a) **18**, 215 (1973).
[24] V. G. Labushkin, V. V. Rudenko, É. R. Sarkisov, V. A. Sarkisyan, and V. N. Seleznev, JETP Lett. **34**, 544 (1981).
[25] V. E. Zubov, G. S. Krinchik, V. N. Seleznev, and M. B. Strugatskiĭ, Zh. Éksp. Teor. Fiz. **94**(10), 290 (1988) [Sov. Phys. JETP **67**, 2122 (1988)]; V. E. Zubov, G. S. Krinchik, V. N. Selesnev, and M. B. Strugatsky, J. Magnetism Magnetic Mat. **86**, 105 (1990).
[26] E. A. Balykina, E. A. Gan'shina, and G. S. Krinchik, Zh. Éksp. Teor. Fiz. **93**, 1879 (1987) [Sov. Phys. JETP **66**, 1073 (1987)].
[27] E. A. Balykina, E. A. Gan'shina, and G. S. Krinchik, Fiz. Tverd. Tela **30**, 570 (1988) [Sov. Phys. Solid State **30**, 326 (1988)].
[28] A. S. Kamzin, and L. A. Grigor'ev, JETP Lett. **57**, 557 (1993); S. Kamzin and L. A. Grigor'ev, Zh. Éksp. Teor. Fiz. **104**, 3489 (1993) [JETP **77**, 658 (1993)].
[29] A. S. Kamzin, V. P. Rusakov, and L. A. Grigoriev, in *Proc. Phys. of Trans. Metals., International Conf*., USSR, part II, p. 271 (1988).
[30] A. S. Kamzin, and L. A. Grigor'ev, Pis'ma Zh. Tekh. Fiz. **16**(16), 38 (1990) [Sov. Phys. Tech. Phys. Lett. **16**, 616 (1990)].
[31] V. I. Gol'danskiĭ, L. I. Krizhanskiĭ, and V. V. Khrapov, eds., *Mössbauer Effect and Its Applications in Chemistry*, Consultants Bureau, New York (1964).
[32] A. P. Kuprin and A. A. Novakova, Nucl. Instrum. and Methods in Phys. Res. B **62**, 493 (1992).
[33] A. S. Kamzin and L. A. Grigor'ev, Pis'ma, Zh. Tekh. Fiz. **19**(4), 50 (1993) [Tech. Phys. Lett. **19**, 245 (1993)].
[34] T. Kobayashi, T. Fukumura, and A. Nakanishi, Nuc. Instrum. and Methods in Phys. Res. B **76**, 204 (1993).
[35] G. Klingelhofer, U. Imkeller, E. Kankeleit, and B. Stahl, Hyperfine Interactions **69**, 819 (1991).
[36] G. Amonoff, and Geof. Foren. Stockholm Forhandl. **47**, 283 (1925).
[37] Sh. Sh. Bashkirov, A. B. Liberman, and V. I. Sinyavskiĭ, Magnetic Structure of the Ferrites [in Russian], Izdatel'stvo Kazanskogo Universiteta (1978).
[38] A. M. Balbasbov, V. A. Makarov, E. F. Makarov, V. A. Povitskiĭ, and A. Ya Chervonenkis, Izv. Akad. Nauk SSSR. Ser. Fiz. **34**, 943 (1979).
[39] Ch. Houpert, M. Nguyen, F. Studer, D. Groult, and M. Toulemonde, Nucl. Intrum. and Methods in Phys. Res. B **34**, 228 (1988).
[40] Z. Haba, and D. Broz, JMMM **124**, 27 (1993).
[41] A. S. Kamzin, L. P. Ol'khovik, and V. L. Rozenbaum, JETP Lett. **61**, 936 (1995).

Translated by D. H. McNeill

Mössbauer absorption by soft ferromagnets in radio-frequency magnetic field

A. Ya. Dzyublik*) and V. Yu. Spivak*)

Institute for Nuclear Research, Academy of Sciences of Ukraine, 252028 Kiev, Ukraine
(Submitted 9 July 1996)
Zh. Éksp. Teor. Fiz. **111**, 1438–1448 (April 1997)

The simultaneous influence of periodical magnetic field reversals at the nucleus between the values $\pm\mathbf{h}_0$ and of magnetostrictive vibrations on the shape of the Mössbauer absorption spectrum is analyzed. The effect of a constant external magnetic field is taken into account by assuming unequal durations of the states $-\mathbf{h}_0$ and $+\mathbf{h}_0$. It is shown that such asymmetric reversals of the magnetic field lead to splitting of the absorption lines into Zeeman patterns corresponding to the time-averaged magnetic field $\mathbf{h}_0 R$, where R is the asymmetry parameter of the reversals. The calculations agree well with experiment. © *1997 American Institute of Physics.* [S1063-7761(97)02304-4]

1. INTRODUCTION

As is well known (see the surveys in Refs. 1–4), an external radio-frequency (RF) magnetic field generates magnetostrictive vibrations in ferromagnets. The corresponding Mössbauer absorption spectrum for hard ferromagnets[1–12] consists of a central Zeeman pattern and additional lines (sidebands) shifted by $n\Omega$, where n is an integer and $\Omega = 2\pi\nu$ is the circular frequency of the alternating magnetic field. The corresponding theory is the standard one for Mössbauer absorption in a vibrating crystal.[11,15] In the case of soft ferromagnets with low anisotropy fields $\mathbf{H}_a$, the absorption spectrum is a fringe of equidistant lines (doublets if the quadrupole interaction is important, i.e., $Q \neq 0$). It collapses to single or double lines when the frequency Ω greatly exceeds the Larmor frequency Ω_L. Such RF collapse has been qualitatively explained by Pfeiffer,[1,16] who assumed that the crystal magnetization $\mathbf{M}(t)$, induced by the RF magnetic field and the magnetic field $\mathbf{h}(t)$ at the nucleus, related to $\mathbf{M}$, periodically reverse direction, so that at high frequencies the nucleus feels only a zero average magnetic field and the absorption spectrum degenerates to a single line. In the opposite case of vanishing Ω the spectrum[1–4] coincides with the typical spectrum for constant magnetic field.

Two main model approaches to this problem are known: the harmonic model of Olariu *et al.*,[17] in which $\mathbf{h}(t) = \mathbf{h}_0 \cos\Omega t$ is assumed, and second, the coherent stepwise model,[18,19] assuming that $\mathbf{h}(t)$ performs periodical instantaneous jumps between the values $+\mathbf{h}_0$ and $-\mathbf{h}_0$. General formulae for both the absorption and scattering spectra, taking into account time-dependent effects, were derived in Ref. 20. In particular, it was shown that the harmonic model of Olariu[17] leads to incorrect result in the stationary limit, when $\Omega \to 0$.

The stepwise model only qualitatively reproduces the observations, yielding collapse at high frequencies and a stationary Zeeman spectrum at vanishing frequencies. In the intermediate case $\Omega \sim \Omega_L$, this model predicts a much more rapid drop in satellite intensities with the order n than is found experimentally.[1,16] But experimental data for Permalloy with damped magnetostrictive vibrations are well described by the stepwise model.[21] So it is natural to attribute the great discrepancy between calculations and experimental data to the contribution of magnetostrictive vibrations. An attempt to take into account simultaneously the effect of both the magnetic field reversals and magnetostrictive vibrations has been described in Ref. 19. Unfortunately, the final result was stated without any derivation, and as will be shown below, it contains some mistakes.

Hence one goal of this paper is to derive a theory simultaneously treating both reversals of the magnetic field and magnetostrictive vibrations. Another goal is to analyze the effect of a superimposed constant magnetic field $\mathbf{H}_0$ on the Mössbauer absorption in an RF magnetic field.

There are clear indications[22,23] that soft ferromagnets represent cluster structure. Every such cluster behaves as a superparamagnetic particle, within which all the spins are strongly coupled. At the same time, the intercluster interaction is weak. The main role of an external RF field is probably to destroy this intercluster cooperation.[19] The magnetization $\mathbf{M}_c$ of the superparamagnetic cluster may be oriented along the easy magnetization axis or in the opposite direction. Consequently, its potential energy W has the form of two potential wells separated by a potential barrier.[24] The minima of these potential wells correspond to the two values of the magnetization, $\mathbf{M}_c$ and $-\mathbf{M}_c$. If the easy magnetization axis is parallel to the external magnetic field $\mathbf{H}_{\mathrm{RF}}(t)$, then the additional potential energy of the cluster will be $V(t) = -\mathbf{M}_c \cdot \mathbf{H}_{\mathrm{RF}}(t)$. Hence the complete potential energy $W + V(t)$ will be an asymmetric time-dependent curve. In a strong field, $\mathbf{H}_{\mathrm{RF}}(t)$ one of the potential wells vanishes at some point, which forces $\mathbf{M}_c$ to jump into the opposite potential well. Another jump occurs in the backward direction at time $T/2$.

When we superimpose a constant magnetic field $\mathbf{H}_0$ parallel to $\mathbf{H}_{\mathrm{RF}}(t)$, the superparamagnetic cluster receives an additional contribution $M_c H_0$ or $-M_c H_0$ to the potential energy. The static field $\mathbf{H}_0$ itself without $\mathbf{H}_{\mathrm{RF}}(t)$ ensures the asymmetry of the potential curve, deepening the potential well corresponding to $\mathbf{M}_c$ oriented along $\mathbf{H}_0$. The switched-on field $\mathbf{H}_{\mathrm{RF}}(t)$, being larger than $\mathbf{H}_0$, will again produce jumps of the magnetization, but in this case the time T_1 spent by the cluster in the potential well with $\mathbf{M}_c$ parallel

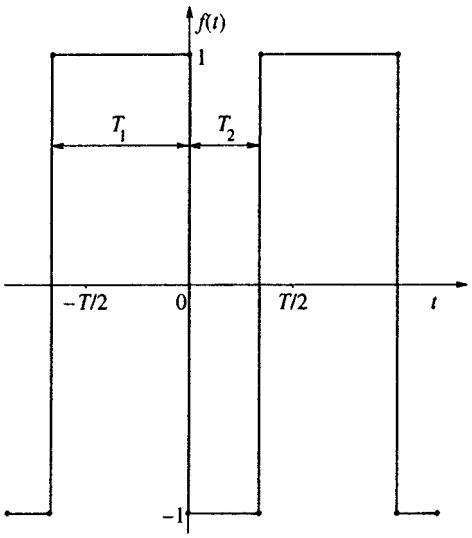

FIG. 1. Time-dependence of the magnetic field at the nucleus.

to $\mathbf{H}_0$ will be greater than the time T_2 spent in that with $\mathbf{M}_c$ antiparallel to $\mathbf{H}_0$.

We do not touch on any stochastic problems here (see also Ref. 19) and suppose the crystal magnetization to be completely governed by the external magnetic field, i.e., we deal only with coherent reversals of the magnetization at definite times, which imply corresponding coherent reversals of the magnetic field at the nucleus.

2. WAVE FUNCTIONS

Let the magnetic field $\mathbf{h}(t)$ at the Mössbauer nucleus periodically change its direction to the opposite one, i.e.,

$$\mathbf{h}(t)=\mathbf{h}_0 f(t), \quad f(t)=f(t+T), \tag{1}$$

where T is the period of the RF field and $\Omega=2\pi/T$ is the circular frequency. We suppose that the reversals of the magnetic field occur as abrupt jumps from $+\mathbf{h}_0$ to $-\mathbf{h}_0$ and vice versa. Let one of such jumps from $+\mathbf{h}_0$ to $-\mathbf{h}_0$ be at $t=0$. If we have $\mathbf{H}_0\neq 0$, then the time T_1 during which $\mathbf{h}(t)=\mathbf{h}_0$ holds is greater than the time T_2 when $\mathbf{h}(t)=-\mathbf{h}_0$. Then

$$f(t)=\begin{cases} 1, & -T_1<t<0 \ \text{ or } \ T_2<t<T_1+T_2, \\ -1, & 0<t<T_2, \end{cases} \tag{2}$$

as shown in Fig. 1.

In order to describe such asymmetric reversals, we introduce the dimensionless parameter

$$R=\frac{T_1-T_2}{T_1+T_2}, \tag{3}$$

which varies in the interval $0\leqslant R\leqslant 1$. The value $R=0$ corresponds to symmetric reversals at equal times $T/2$, and $R=1$ to the constant field $+\mathbf{h}_0$ during the whole time. From (3) it follows that

$$T_1=\frac{1+R}{2}\,T, \quad T_2=\frac{1-R}{2}\,T. \tag{4}$$

The Floquet wave function of the nucleus in periodical field (1), and (2) may be written as

$$\Psi^N_{I_\kappa M_\kappa}(t)=|I_\kappa M_\kappa\rangle\Phi^N_{I_\kappa M_\kappa}(t)e^{i\mathscr{E}^N_{M_\kappa}t/\hbar}, \tag{5}$$

where $|I_k M_k\rangle$ is the stationary wave function of the nucleus in the κth state ($\kappa=g$ for the ground state and $\kappa=e$ for the excited one) with spin I_κ and projection M_κ in the direction of $\mathbf{h}_0$; the periodic function of time $\Phi^N(t)$ may be defined in the interval from $-T/2$ to $T/2$ as

$$\Phi^N_{I_\kappa M_\kappa}(t)=\begin{cases} \exp[i\gamma_\kappa M_\kappa h_0(1-R)t], & -T/2\leqslant t\leqslant 0, \\ \exp[-i\gamma_\kappa M_\kappa h_0(1+R)t], & 0\leqslant t\leqslant T_2, \\ \exp[i\gamma_\kappa M_\kappa h_0(1-R)(t-T)], & T_2\leqslant t\leqslant T/2. \end{cases} \tag{6}$$

The corresponding quasi-energies are

$$\mathscr{E}^N_{M_\kappa}=E^N_\kappa-\gamma_\kappa M_\kappa h_0 R+Q[3M^2_\kappa-I_\kappa(I_\kappa+1)], \tag{7}$$

where $E^N_g=0$ and $E^N_e=E'_0$ is the energy of the unsplit resonant level, γ_κ specifies the gyromagnetic ratio, and Q is the quadrupole constant. The complete set of quasi-energies is obtained by adding $n\hbar\Omega$ to (7), where n is an integer; the corresponding functions are given by (6) multiplied by $\exp(in\Omega t)$.

The RF magnetic field also induces magnetostrictive vibrations in a ferromagnetic crystal. If the crystal magnetization induced by the RF field coherently changes direction at definite moments of time, then the magnetostrictive vibrations occur with frequency twice Ω.[17,19] Strictly speaking, the wave function of such a lattice in the harmonic approximation is represented by a product of wave functions to describe the quantum oscillations of uncoupled oscillators about their instantaneous equilibrium positions, which perform classical vibrations.[14,15] The displacement of the Mössbauer atom owing to such classical vibrations is

$$\mathbf{X}(t)=\mathbf{A}\cos(2\Omega t+\varphi_0), \tag{8}$$

where $\mathbf{A}$ and φ_0 are the amplitude and initial phase of these magnetostrictive vibrations, respectively.

3. MÖSSBAUER ABSORPTION

A simple generalization of the method developed in Refs. 14, 15, and 20 gives the following result for the phononless absorption cross-section of a γ-ray with energy E by the Mössbauer nucleus:

$$\sigma_a(E)=\frac{\sigma_0\Gamma^2}{2}\,e^{-2W_a} \times\sum_{n=-\infty}^{\infty}\sum_{M_e,M_g}\frac{J_{eg}(\vartheta)|a_{eg}(n)|^2}{(E-\mathscr{E}^N_e+\mathscr{E}^N_g-n\hbar\Omega)^2+(\Gamma/2)^2}, \tag{9}$$

where σ_0 is the resonant cross-section, Γ is the width of the resonant level, e^{2W} is the Debye–Waller factor, and the functions $J_{eg}(\vartheta)$ determine the relative intensities of the lines versus the angle ϑ between the wave vector $\mathbf{k}$ of the γ-quantum and h_0:

$$\frac{1}{3}\,J_{\pm 3/2,\pm 1/2}(\vartheta)=J_{\pm 1/2,\pm 1/2}(\vartheta)=\frac{1}{16}\,(1+\cos^2\vartheta),$$

$$J_{\pm 1/2,\mp 1/2}(\vartheta)=\frac{1}{4}\sin^2\vartheta. \tag{10}$$

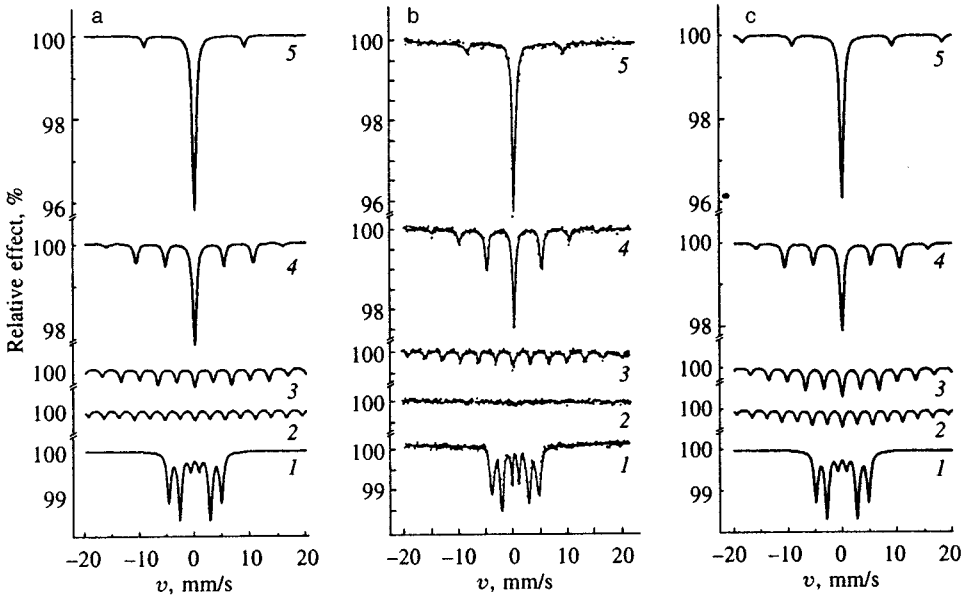

FIG. 2. Simultaneous effect of magnetic field reversals and magnetostrictive vibrations on the shape of Mössbauer spectra: a) our calculations without any constraints to the modulation index; b) Pfeiffer's data;[16] c) our calculations using the frequency dependence (24) of the modulation index. Curves *1, 2, 3, 4,* and *5* correspond to $\gamma=0$, 32, 39, 61, and 106 MHz.

The quantities $a_{eg}(n)$ stand for the following Fourier coefficients:

$$a_{eg}(n)=\frac{1}{T}\int_{-T/2}^{T/2} dt e^{-in\Omega t}\Phi_e^N(t)^*\Phi_g^N(t) \times\exp[i\mathbf{k}\cdot\mathbf{A}\cos(2\Omega t+\varphi_0)]. \tag{11}$$

The cross-section (9) must be averaged over the phononless energy distribution of the incident γ-rays

$$w_e(E)=e^{-2W_e}\frac{\Gamma/2\pi}{(E-E_0-s)^2+(\Gamma/2)^2}, \tag{12}$$

where $s=(v/c)E_0$ is the Doppler shift and v is the velocity of the emitter relative to the absorber. This averaging gives for ^{57}Fe

$$\sigma_a(s)=\frac{\sigma_0\Gamma^2}{2}e^{-2W_e-2W_a}\sum_{n=-\infty}^{\infty}\sum_{M_e,M_g} \times\frac{J_{eg}(\vartheta)|a_{eg}(n)|^2}{\{s-\Delta-\hbar\alpha_{eg}R-Q[3M_e^2-15/4]-n\hbar\Omega\}^2+\Gamma^2}, \tag{13}$$

where $\Delta=E_0'-E_0$ is an isomer shift; the quantities

$$\hbar\alpha_{eg}=(\gamma_g M_g-\gamma_e M_e)h_0 \tag{14}$$

determine the Zeeman splitting in the constant field $\mathbf{h}_0$. Furthermore, $\hbar\alpha_{eg}$ multiplied by R determines the magnetic hyperfine structure provided by the time-averaged field $\langle\mathbf{h}(t)\rangle=\mathbf{h}_0R$.

In order to calculate $\alpha_{eg}(n)$ we use the familiar expansion

$$e^{ix\cos t}=\sum_{n=-\infty}^{\infty} i^n J_n(x)e^{int}, \tag{15}$$

where $J_n(x)$ is the Bessel function of order n. Then

$$a_{eg}(n)=\sum_{m=-\infty}^{\infty} i^m e^{im\varphi_0}J_m(\mathbf{k}\cdot\mathbf{A})b_{eg}(n-2m), \tag{16}$$

where the coefficients $b_{eg}(n)$, given by (11) with $\mathbf{A}=0$, may be written as

$$b_{eg}(n)=b_{eg}^{(+)}(n)+b_{eg}^{(-)}(n),$$

$$b_{eg}^{(+)}(n)=\frac{1}{T}\int_0^{(1-R)T/2} dt\ \exp[-in\Omega t-i\alpha_{eg}(1+R)t],$$

$$b_{eg}^{(-)}(n)=\frac{1}{T}\int_{-(1+R)T/2}^{0} dt\ \exp[-in\Omega t+i\alpha_{eg}(1-R)t]. \tag{17}$$

A simple calculation yields

$$b_{eg}(n)=\frac{2x_{eg}}{[(1-R)x_{eg}-n\pi][(1+R)x_{eg}+n\pi]} \times\sin\left\{\frac{1+R}{2}[(1-R)x_{eg}-n\pi]\right\} \times\exp\left\{i\frac{1+R}{2}[(1-R)x_{eg}-n\pi]\right\}, \tag{18}$$

where we have used the notation

$$x_{eg}=\frac{\alpha_{eg}T}{2}. \tag{19}$$

The forced vibrations are important only for $\mathbf{k}\cdot\mathbf{A}\neq 0$, i.e., when $\mathbf{A}$ is not perpendicular to $\mathbf{k}$. In all the experiments the beam of incident Mössbauer radiation is perpendicular to the absorber surface and the external alternating magnetic field $\mathbf{H}_{RF}(t)$ is parallel to it. Therefore the original magnetostrictive vibrations generated along $\mathbf{H}_{RF}(t)$ are perpendicular to $\mathbf{k}$. As pointed out in Ref. 16, these vibrations are scattered by defects of the crystal, giving rise to vibrations along $\mathbf{k}$,

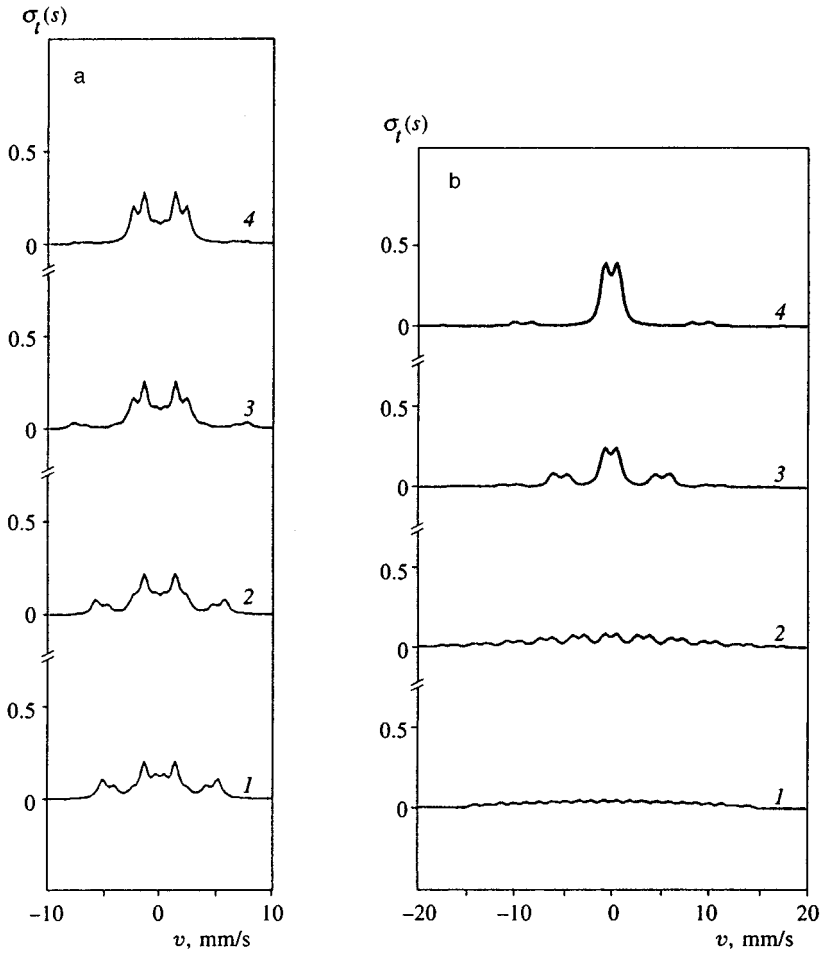

FIG. 3. Hyperfine structure of Mössbauer absorption lines due to the asymmetry of the magnetic field reversals: a) when vibrations are absent ($Q=m=0$) and $R=0.5$; b) the modulation indices are the same as in Fig. 2a and $R=0.2$; $\sigma_t(s)$ is in units $(\sigma_0\Gamma^2/2)\exp(-2W_e-2W_a)$; curves *1, 2, 3,* and *4* correspond to $\nu=32$, 39, 62, and 106 MHz.

which manifest themselves in experiment. Then such vibrations have random phases φ_0. Averaging $|\alpha_{eg}(n)|^2$ over φ_0 one gets

$$\overline{|a_{eg}(n)|^2}=\sum_{m=-\infty}^{\infty} J_m^2(kx_0)|b_{eg}(n-2m)|^2, \tag{20}$$

where $x_0=\mathbf{k}\cdot\mathbf{A}/k$ is the amplitude of vibrations along $\mathbf{k}$. Besides, we must average (20) over the distribution of the amplitudes x_0. We shall take the Rayleigh distribution, which describes Mössbauer data most accurately:

$$P(x_0)=\frac{x_0^2}{\overline{x_0^2}}\exp\left\{-\frac{x_0^2}{2\overline{x_0^2}}\right\}, \tag{21}$$

where $\overline{x_0}$ corresponds to the maximum of the distribution. Then the average cross section may be written as

$$\sigma_t(s)=\sum_{n=-\infty}^{\infty} e^{-m^2} I_n(m^2)\sigma_t(s-n2\hbar\Omega)|_{\overline{x_0}=0}, \tag{22}$$

where $I_n(x)$ is the modified Bessel function, $m=k\overline{x_0}$ represents the so-called modulation index, and the cross-section $\sigma_t(s)|_{\overline{x_0}=0}$ for the case without vibrations is given by

$$\sigma_t(s)|_{\overline{x_0}=0}=\frac{\sigma_0\Gamma^2}{2}e^{-2W_e-2W_a}\sum_{k=-\infty}^{\infty}\sum_{M_e,M_g}$$

$$\times\frac{J_{eg}(\vartheta)|b_{eg}(k)|^2}{[s-\Delta-\hbar\alpha_{eg}R-Q(3M_e^2-15/4)-k\hbar\Omega]^2+\Gamma^2}. \tag{23}$$

Previous calculations[19] for $R=0$ were based on an equation similar to (22), which contained the incorrect factor $I_{2n}(m^2)$ instead of $I_n(m^2)$. Therefore we first analyzed the case of symmetric reversals with $R=0$. Using Eqs. (22) and (23) we found the best fit to Pfeiffer's data[1,16] using the modulation index $m=0.05$ for $\nu=106$ MHz, $m=0.6$ for $\nu=61$ MHz $m=2.4$ for $\nu=39$ MHz, and $m=4$ for $\nu=31$ MHz. As can be seen from Fig. 2, the agreement of these calculations with experiment is good. However, Pfeiffer[1,16] proposed the following frequency dependence of the modulation index:

$$m^2\propto 1/\nu^{3.5}. \tag{24}$$

Calculations employing this law are shown in Fig. 2(c). In this case the fitting parameters are $m=0.28$ for $\nu=106$ MHz, $m=0.74$ for $\nu=61$ MHz, $m=1.6$ for $\nu=39$ MHz, and $m=2.4$ for $\nu=31$ MHz. Here the agreement with experiment is worse, so that the law (24) is doubtful.

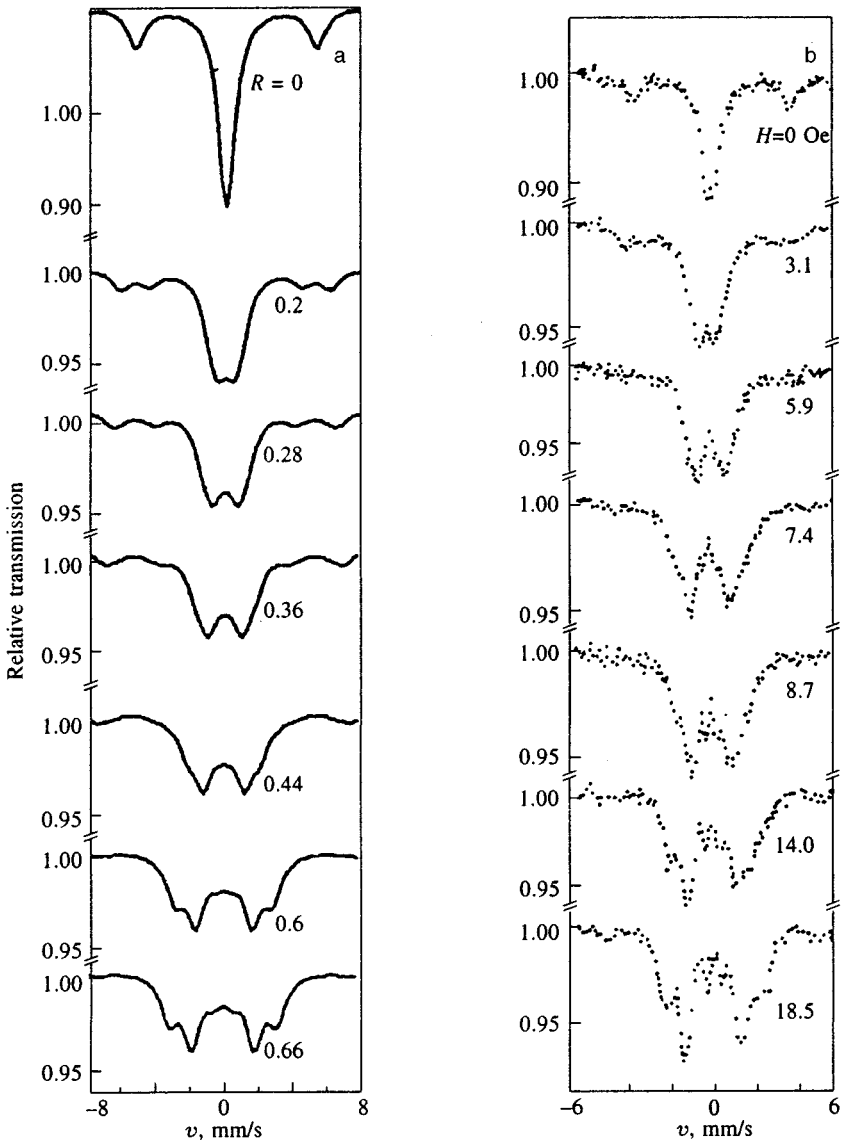

FIG. 4. Splitting of the collapsed line due to a superimposed static magnetic field: a) our calculations; b) experimental data (Ref. 26).

4. LIMITING CASES

In the high-frequency case we have

$$\lim_{\Omega\to\infty} b_{eg}(n)=\delta_{n0}. \tag{25}$$

The average amplitude $\overline{x_0}$ of magnetostrictive vibrations decreases with growing frequency Ω. For $\Omega\gg|\alpha_{eg}|$ one therefore has

$$|a_{eg}(n)|^2\simeq\delta_{n0}. \tag{26}$$

Thus, at such high frequencies the absorption cross-section becomes

$$\sigma_t(s)_{av}=\frac{\sigma_0\Gamma^2}{2}\,e^{-2W_e-2W_a}\times\sum_{M_e,M_g}\frac{J_{eg}(\vartheta)}{[s-\Delta-\hbar\alpha_{eg}R-Q(3M_e^2-15/4)]^2+\Gamma^2}. \tag{27}$$

Hence we can state that the hyperfine structure of the Mössbauer spectrum collapses to a single line at $Q=0$ or a quadrupole doublet for $Q\neq0$ only in the case of symmetric reversals with $R=0$. When the magnetic field reverses asymmetrically with $R>0$, the nucleus feels the high-frequency reversing magnetic field as a static field with magnitude $\langle h(t)\rangle$ smaller than h_0 by the factor R. Then the spectrum retains its hyperfine structure corresponding to the field $\langle h(t)\rangle$.

In the opposite low-frequency case, $(1-R)T\to\infty$, using the definition of the ζ-function[25] one finds

$$b_{eg}^{(\pm)}\approx\pm\frac{1}{iT}\,[n\Omega\pm(1\pm R)\alpha_{eg}\mp i\eta]^{-1}, \tag{28}$$

where $\eta\to+0$. Putting $n\Omega=\omega_n$ and $\Omega=\Delta\omega$, one can treat the sums over n as integrals,

$$\frac{1}{T}\sum_{n=-\infty}^{\infty}\to\frac{1}{2\pi}\int_{-\infty}^{\infty}d\omega, \tag{29}$$

in the limit $T\to\infty$. Substitution of (28) and (29) into (13) gives after a contour integration the six-line pattern produced by the constant magnetic field h_0 (see the curves *1* in Fig. 2).

5. DISCUSSION

An external RF magnetic field influences the nucleus in a soft ferromagnet via the reversing magnetic field $\mathbf{h}(t)$ at the nucleus and magnetostrictive vibrations. For the constant magnetic field $\mathbf{H}_0=0$ the nucleus has infinite sets of quasi-energetic levels separated by the interval $\hbar\Omega$ both in the ground and excited states. In the absence of quadrupole interaction $(Q=0)$ each such level is degenerate with respect to the magnetic quantum number M_κ. Such quasi-energetic structure produces transitions of the nucleus absorbing γ-rays, from the ground state to any quasi-energetic level corresponding to the excited nuclear state. As a consequence, the Mössbauer spectrum will consist of a set of equidistant lines (doublets if $Q=0$). There nth line (doublet) is associated with the transitions from the ground state with quasi-energy $\mathscr{E}^N_{M_g}$ to excited states with quasi-energies $\mathscr{E}^N_{M_e}+n\hbar\Omega$, which are specified by the quantum number M_e. In high- and low-frequency cases, such spectra approach single or double lines and a standard sextet, respectively.

Our calculations well agree with Pfeiffer's data on Permalloy[16] (see Fig. 2) supporting the point of view that the RF magnetic field causes both reversals of the magnetization and magnetostrictive vibrations. Note that Julian and Daniels,[19] who have been using the incorrect formula, predicted unreasonably high modulation indices fitting the same data: for 32 MHz they used $m=100$.

When a constant magnetic field is superimposed, the quasi-energetic sublevels split as indicated by Eq. (7). It is of interest that such split quasi-energies coincide with the energies of the nucleus placed in the constant magnetic field $\mathbf{h}_0R$. Thus, the quantum system not only feels the oscillations of the magnetic field exchanging photons with frequency Ω, but also sees its time-averaged value $\langle\mathbf{h}(t)\rangle=\mathbf{h}_0R$, where the asymmetry parameter R depends on the magnitude of the external field $\mathbf{H}_0$. The splitting of quasi-energies leads to a corresponding splitting of the Mössbauer lines. Such a splitting arises at all frequencies and is typical both for the central line and for sidebands. This effect is illustrated by Fig. 3a, showing $\sigma_t(s)$ in units $(\sigma_0\Gamma^2/2)\exp(-2W_e-2W_a)$ for $R=0.5$ and $Q=m=0$. The changes owing to the magnetostrictive vibrations are shown in Fig. 3b, where we used the same modulation indices as in Fig. 2a and took $R=0.2$. Our calculations with $Q=m=0$ are compared with data[26] obtained at $\nu=62$ MHz with different gradually increasing values of $\mathbf{H}_0$ for amorphous $Fe_{78}Si_9B_{13}$ alloy. In Fig. 4 we can see satisfactory agreement with these observations when the line splits under the influence of the static magnetic field. From top to bottom the fitting parameters are $R=0$, 0.2, 0.28, 0.36, 0.44, 0.6, 0.66. In Figs. 2 and 3 we used $\Gamma=0.4$ mm/s, and in Fig. 4 $\Gamma=0.6$ mm/s.

It is of great interest to observe such a splitting, caused by an external constant magnetic field H_0, not only for the collapsed line but also for the sidebands existing at intermediate frequencies. It would be the most direct evidence for the nuclear quasi-energetic picture (7) and respectively to the cluster structure of soft ferromagnets.

*)e-mail: kinr@sovam.com

[1] L. Pfeiffer, in *Mössbauer Effect Methodology*, Vol. 7, ed. by I. J. Gruverman, Plenum Press, New York (1972), pp. 263–298.
[2] J. K. Srivastava, in Advances in *Mössbauer Spectroscopy*, ed. by B. V. Thosar, P. K. Iengar, J. K. Srivastava, and S. C. Bhargava, Elsevier Publ. Amsterdam (1983), pp. 761–813.
[3] M. Kopcewicz, in *Mössbauer Spectroscopy Applied to Inorganic Chemistry*, Vol. 3, ed. by G. J. Long and F. Grandjean, Plenum Press, New York (1989), pp. 243–287.
[4] M. Kopcewicz, Struct. Chem. **2**(105), 313 (1991).
[5] G. Asti, G. Albanese, and C. Bucci, Nuovo Cimento L **7**, 531 (1968).
[6] N. D. Heiman, L. Pfeiffer, and J. C. Walker, Phys. Rev. Lett. **21**, 93 (1968).
[7] G. Albanese, G. Asti, and S. Rinaldi, Lett. Nuovo Cimento **4**, 220 (1972).
[8] M. Kopcewicz, U. Gonser, and H. G. Wagner, Nucl. Instrum. Methods **199**, 163 (1982).
[9] M. Kopcewicz, H. G. Wagner, and U. Gonser, J. Magn. Magn. Mater. **40**, 139 (1983).
[10] G. Asti, G. Albanese, and C. Bucci, Phys. Rev. **184**, 260 (1969).
[11] M. Kopcewicz, Phys. Stat. Sol. **46**, 675 (1978).
[12] L. Pfeiffer, N. D. Heiman, and J. C. Walker, Phys. Rev. B **6**, 74 (1972).
[13] A. Abragam, Acad. Sci. (France) **250**, 4334 (1960).
[14] A. Ya. Dzyublik, Phys. Stat. Sol. B **123**, 53 (1984).
[15] A. Ya. Dzyublik, Phys. Stat. Sol. B **134**, 503 (1986).
[16] L. Pfeiffer, J. Appl. Phys. **42**, 1725 (1971).
[17] S. Olariu, I. Popesku, and C. B. Collins, Phys. Rev. C **23**, 50, 1007 (1981).
[18] Yu. V. Baldokhin, S. A. Borshch, L. M. Klinger, and V. A. Povitsky, Zh. Éksp. Teor. Fiz. **63**, 708 (1972) [Sov. Phys. JETP **36**, 374 (1972)].
[19] S. R. Julian and J. M. Daniels, Phys. Rev. B **38**, 4394 (1988).
[20] A. Ya. Dzyublik, Phys. Stat. Sol. B **194**, 699 (1996).
[21] F. G. Vagizov, PhD Thesis, Institute of Physics and Technology, Kazan (1988).
[22] D. G. Rancourt, S. R. Julian, and J. M. Daniels, J. Magn. Magn. Mat. **51**, 83 (1985).
[23] D. G. Rancourt, H. H. A. Smit, and R. C. Thiel, J. Magn. Magn. Mat. **66**, 121 (1987).
[24] W. Kündig, H. Bömmel, G. Constabaries, and R. H. Lindquist, Phys. Rev. **142**, 327 (1966).
[25] W. Heitler, *Quantum Theory of Radiation*, Clarendon Press, Oxford (1954).
[26] M. Kopcewicz, H. G. Wagner, and U. J. Gonser, J. Magn. Magn. Mat. **51**, 225 (1985).

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

Magnetic ordering of Fe atoms in icosahedral $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals

I. S. Lyubutin

Institute of Crystallography, Russian Academy of Sciences, 117333 Moscow, Russia

Ch. R. Lin and S. T. Lin

Department of Physics, National Cheng-Kung University, 70101 Tainan, Taiwan
(Submitted 28 June 1996)
Zh. Éksp. Teor. Fiz. **111**, 1449–1464 (April 1997)

Novel icosahedral quasicrystals, in which Fe atoms possess a magnetic moment, have been found in $Al_{70-x}B_xPd_{30-y}Fe_y$ compounds with $5<x<10$ and $10<y<20$. The compounds have ferromagnetic properties, and their Curie temperature ranges between 280 and 340 K, the saturation magnetization $\sigma_s(5\text{ K})\approx 7.5$ emu/g. It follows from Mössbauer spectra that only a fraction of Fe atoms (12 to 15%) are magnetically ordered at 4.2 K, and the mean saturation field $\langle H_{hf}\rangle = 96$ kOe. The isomer shift values confirm that the atomic volume of magnetic Fe sites is larger than that of nonmagnetic Fe sites. The magnetic properties of these quasicrystals can be interpreted in terms of large magnetic clusters with a size of 185 to 290 Å. This size correspond to about 4×10^4 ''unit cells,'' hence the magnetic state can be described in terms of bulk parameters. The localized magnetic moment of Fe atoms is tentatively ascribed to bonding between Fe and B, similarly to that in the amorphous $Fe_{\sim 50}B_{\sim 50}$ alloy. © *1997 American Institute of Physics.* [S1063-7761(97)02404-9]

1. INTRODUCTION

Magnetic icosahedral quasicrystals (i-QC) containing iron have not been known until recently.[1–4] Unlike i-QC with Mn, quasicrystals containing even a notable quantity of iron are usually nonmagnetic.[3] For example, in the $Al_{65}Cu_{20}Fe_{15}$ i-QC iron atoms bear no magnetic moment, at least at $T>1.5$ K.[3] $Al_{70}Pd_{20}Fe_{10}$[5] and$Al_{70}Pd_{16}Fe_{14}$[6] icosahedral quasicrystals are also nonmagnetic at temperatures higher than 4.2 K. Quasicrystals based on Al–Mn alloys with a small content of Fe are paramagnetic.[7–9] In Al–Si–Mn ferromagnetic i-QC, iron doping results in spin-glass ordering, which coexists at low temperatures with ferromagnetism, but iron atoms in these materials have zero magnetic moment.[7] In ferromagnetic $Al_{40}Ge_{25}Mn_{25}Cu_{10-x}Fe_x$ i-QC, iron atoms have zero magnetic moment at 100 K, and a small hyperfine magnetic field at iron nuclei ($H_{hf}=9$ kOe) is transferred from manganese atoms.[10] Recently Nasu and co-workers[11–13] detected in Al–Ge–Mn–Cu–Fe quasicrystals a magnetic transition at 30 K, and the presence of magnetic and nonmagnetic iron atom sites below 30 K. Stadnik and Stroink[14,15] have found that in Al–Ge–Mn and Al–Ge–Cu–Mn i-QC lightly doped with iron, a fraction of iron atoms have a small magnetic moment at 4.2 K, but there are also nonmagnetic sites, at which Fe atoms have zero magnetic moment.

Recently Yokoyama and co-workers[16,17] discovered that boron doping of nonmagnetic quasicrystals based on Al–Mn compounds leads to appearance of magnetic moments at Mn atoms. The magnetization value of Al–Mn–Pd–B quasicrystals depends on the B content and has a maximum at 10% B. The magnetic state of Al–Mn–Pd–B system is a feature of the quasicrystalline phase, and the magnetization vanishes when the icosahedral phase transforms to a crystalline state. By analogy with the Mn-based systems, we have discovered recently the novel $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ magnetic i-QC containing iron.[18] This is a ferromagnet with a Curie temperature $T_C=305$ K. At $T=5$ K its magnetization is $\sigma_s=4.2$ emu/g, the coercive force is 1.2 kOe, and the magnetic moment μ_{eff} is estimated to be $1.63\mu_B$ per Fe atom. The conclusion that the magnetic state is a feature of the quasicrystalline phase of $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ and is not due to impurities or crystalline phases was supported by various arguments.[18]

Mössbauer spectroscopy is one of the most efficient techniques in studies of atomic and magnetic structure of quasicrystals containing Fe atoms. This topic was discussed in several recent reviews.[3,4,12] This paper reports on magnetic measurements and Mössbauer spectroscopy studies of novel $Al_{70-x}B_xPd_{30-y}Fe_y$ ferromagnetic quasicrystals containing iron.

2. EXPERIMENT

We attempted to fabricate quasicrystals of various $Al_{70-x}B_xPd_{30-y}Fe_y$ compositions. Icosahedral quasicrystals were produced by arc melting of high-purity initial components, namely Al (99.999 wt. %), B (99.9 wt. %), Pd (99.9 wt. %), and Fe (99.995 wt. %) in purified argon atmosphere. The melt was rapidly quenched by pouring it onto a copper roller with a diameter of 15 cm rotated at 6000 rpm. X-ray diffraction patterns were recorded using an x-ray tube with a rotating copper anode (we used the K_α line, the tube was operated at 50 kV and 200 mA) with a graphite monochromator using the (002) reflection. Magnetization was measured by an MPMS-5 Quantum Design SQUID magnetometer in the temperature range 5 to 400 K in a magnetic field H_{ext} of up to 5.5 T. Electrical resistivity of some of the samples was measured in the temperature range 4.2–300 K. Mössbauer spectra of ^{57}Fe nuclei were recorded in the temperature range 4.2–400 K using conventional spectrometers

 1063-7761/97/040800-08$10.00

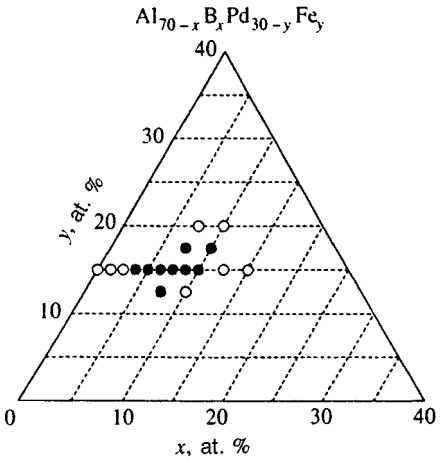

FIG. 1. Regions of the composition diagram of the Al–B–Pd–Fe system where single-phase icosahedral quasicrystals occur: (●) only single icosahedral phase; (○) crystalline phase and/or decagonal quasicrystalline phases coexist with the icosahedral phase, the icosahedral phase being dominant.

operating in the transmission configuration at constant acceleration. The ^{57}Co(Rh) gamma-ray source was at room temperature. Measurements at temperatures 4.2 to 295 K were performed in a helium-flow cryostat of Oxford Instruments CF-506 type, and at 295–400 K in a VF-1000 vacuum oven of Austin Instruments Inc. The effective thickness of our absorbers was estimated to be ~ 0.08.[19] This means that in analyzing our spectra we could use the ''thin-absorber approximation.''[19] All the values of the Mössbauer isomer shift are given in the paper with respect to metallic iron at room temperature.

3. RESULTS AND DISCUSSION

3.1. Formation and stability of quasicrystalline phases

We have succeeded in obtaining the $Al_{70}Pd_{20}Fe_{10}$ single-phase quasicrystal of the $Al_{70}Pd_{30-y}Fe_y$ system and a similar $Al_{70}Pd_{20}Fe_8Mn_2$ compound, in which some iron atoms were replaced with manganese. In addition to *i*-QC phase, $Al_{70}Pd_{15}Fe_{15}$ contained a fraction of the crystalline phase. Several single-phase *i*-QCs, however, were fabricated by replacing some aluminum atoms with boron. We plotted a phase diagram for the $Al_{70-x}B_xPd_{30-y}Fe_y$ system and found that compounds with the single-phase icosahedral structure were formed in the region $5<x<10$, $10<y<20$ (Fig. 1). In particular, we fabricated $Al_{62.5}B_{7.5}Pd_{30-y}Fe_y$ single-phase *i*-QC at $y=12.5$, 15 and 17.5.

X-ray diffraction patterns of these materials are shown in Fig. 2. All the reflections are identified as icosahedral peaks, according to Elser's scheme[20,21] suggested for identification of icosahedral quasicrystalline phases. This means that the entire sample is composed of a single phase, and its structure can be described as a six-dimensional face-centered icosahedral quasilattice. According to Elser,[20] the quasilattice constant a_R is defined as the length of edges of rhombohedral cells making up Penrose's three-dimensional tiling. We derived from X-ray diffraction patterns of$Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ the value $2\theta=42.3°$ for the intense reflection with the indices

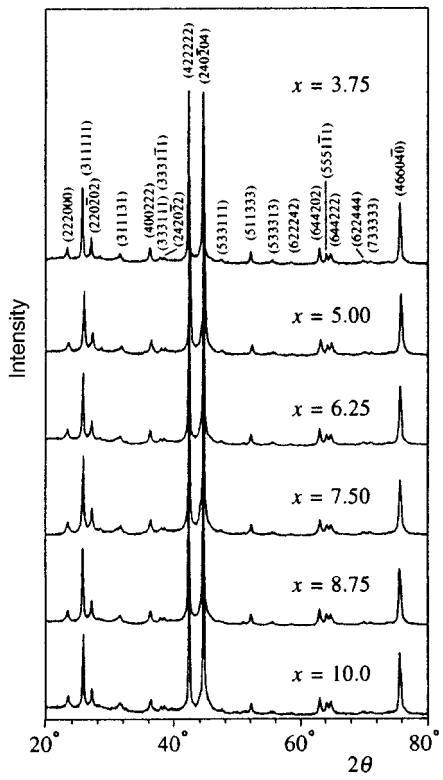

FIG. 2. X-ray diffraction patterns of $Al_{70-x}B_xPd_{15}Fe_{15}$ icosahedral quasicrystals.

(422222) (Fig. 2) characterized by the relation $Q_{100000}=4\pi\lambda^{-1}\sin\theta$.[20] By using the equation $Q_{100000}\,a_R=13.308$, which applies to the face-centered icosahedral structure,[20,21] we obtained $a_R=4.5217$ Å.

We have also studied transformation of the *i*-QC phase to equilibrium crystalline phases. The $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal was annealed at a fixed temperature T_a ranging between 400 and 1000 K for two hours, and after each anneal stage, x-ray diffraction patterns and Mössbauer spectra were recorded. We have found that no phase transformations occur at $T_a \leqslant 800$ K. At annealing temperatures of up to 1000 K, the *i*-QC phase starts to decompose into crystalline phases. From the x-ray data, these phases were identified as Al–Fe, Fe_2B, and FeB. The Mössbauer spectra of these crystalline phases are radically different from those of the quasicrystalline phase, so they can be used as an additional test of the sample phase composition.

3.2. Magnetic measurement data

Curves of magnetization versus magnetic field, $\sigma(H)$, are given in Fig. 3. In strong magnetic fields, the curves are almost saturated, and they show the typical ferromagnetic behavior of all icosahedral compositions. Hysteresis properties of magnetization typical of the ferromagnetic state were observed at all temperatures up to the Curie point T_C.[18] We have found that the saturation magnetization σ_s increases with the concentration of Fe and B (Fig. 4). The values of σ_s are equal for ribbon and powder samples. However, the values of σ_s and T_C for samples of identical composition,

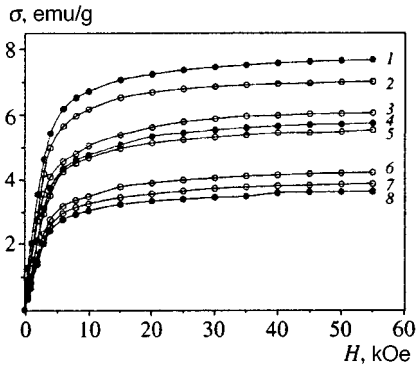

FIG. 3. Magnetization of $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals with various compositions versus magnetic field at 5 K: (1) $x=7.50$, $y=17.5$; (2) $x=10.0$, $y=15.0$; (3) $x=8.25$, $y=15.0$; (4) $x=7.5$, $y=15.0$; (5) $x=6.25$, $y=15.0$; (6) $x=5.0$, $y=15.0$; (7) $x=3.75$, $y=15.0$; (8) $x=7.50$, $y=12.5$.

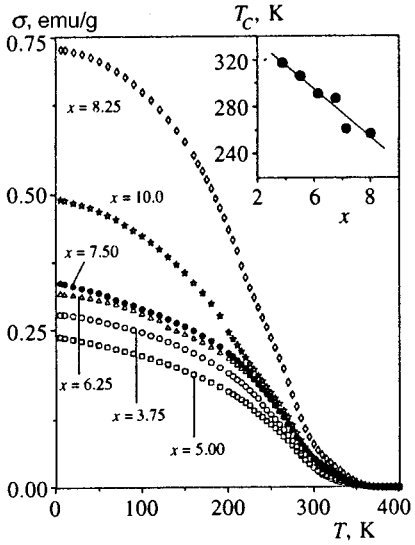

FIG. 5. Magnetization of $Al_{70-x}B_xPd_{15}Fe_{15}$ quasicrystals with various boron concentrations x in a magnetic field of 20 Oe versus temperature. The insert shows the Curie temperature as a function of the boron concentration x.

fabricated in the same conditions but at different times are slightly different. In particular, σ_s and T_C for the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ samples measured in this work are different from the results of Ref. 18. The temperature dependences of magnetization (Figs. 5 and 6) are very complicated and it is difficult to find a proper fitting model which is valid throughout the studied temperature range. For rough estimation of the Curie temperature, we used the linear extrapolation of the $\sigma^2(T)$ curves obtained by measuring the field-cooled magnetization. The resulting Curie temperature increases with the Fe content, but decreases with the content of boron (see Table II and inserts in Figs. 5 and 6).

Ferromagnetic properties and an increase in the magnetization with the boron concentration were also detected in the Mn-analog of our quasicrystals, i.e. $Al_{70-x}B_xPD_{15}Mn_{15}$,[17] although the B concentrations at which single-phase quasicrystals exist are different for these two systems. The ferromagnetism of $Al_{70-x}B_xPd_{15}Mn_{15}$ quasicrystals has been explained in terms of magnetic Mn–Mn clusters, whereas the role of B is to enhance the magnetic coupling among the Mn–Mn pairs due to Mn–B bonding and an increase in the cluster size.[17]

3.3. Analysis of Mössbauer spectra

Typical ^{57}Fe Mössbauer spectra of $Al_{70-x}B_xPd_{30-y}Fe_y$ icosahedral quasicrystals are shown in Fig. 7. At low temperatures, the spectra contain an intense central component typical of the nonmagnetic state of Fe atoms in quasicrystals and a highly split component of lower intensity. The temperature dependence of the spectrum shape indicates that the less intense component belongs to a subspectrum due to the magnetic hyperfine interaction. This means that a fraction of Fe atoms at low temperatures are in a magnetically ordered state. In the temperature range 4.2–250 K the magnetic and nonmagnetic components coexist. The coexistence of mag-

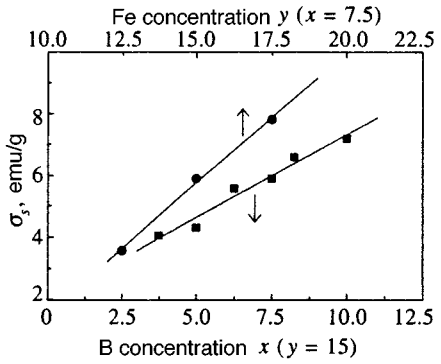

FIG. 4. Saturation magnetization σ_s of the $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals at 5 K as a function of the iron concentration (at a fixed boron concentration $x=7.5$) and of the boron concentration (at a fixed iron concentration $y=15$).

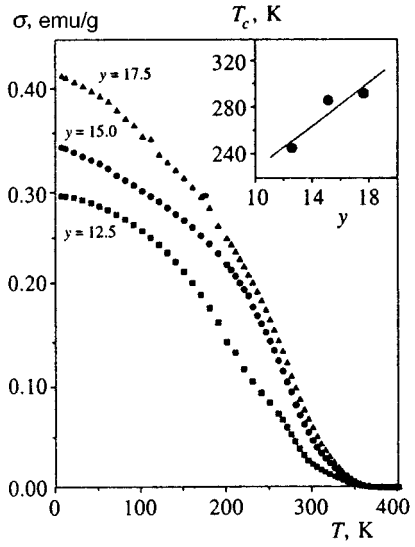

FIG. 6. Magnetization of $Al_{62.5}B_{7.5}Pd_{30-y}Fe_y$ quasicrystals with various content of iron y in a magnetic field of 20 Oe as a function of temperature. The insert shows the Curie temperature T_C as a function of the iron content y.

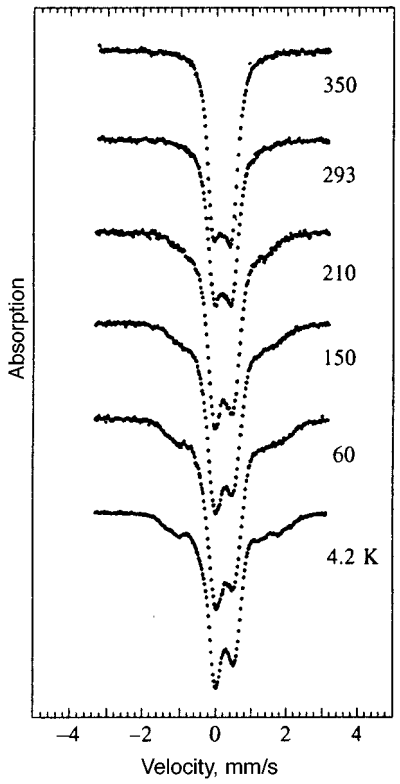

FIG. 7. Mössbauer spectra of ^{57}Fe nuclei in the$Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal at several temperatures ranging between 4.2 and 350 K.

netic and nonmagnetic components was also recently detected in NMR spectra of ^{55}Mn nuclei in Al–B–Pd–Mn quasicrystals.[22]

We have tested several models for processing measured spectra of $Al_{70-x}B_xPd_{30-y}Fe_y$. Table I lists as an example hyperfine parameters derived from spectra of the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ sample using different models. The nonmagnetic component is fairly well described by the model of quadrupole-splitting distribution, $P(QS)$ (Fig. 8a, b, and Table I), in which we have used Window's approximation.[23] In order to account for the asymmetry of experimental spectra, we have introduced a linear relation between the quadrupole splitting QS and isomer shift IS[24,25]: $IS=IS_0+nQS$, where IS_0 corresponds to $QS=0$, and n is the correlation parameter. We have noted, however, that the center of gravity of the magnetic component of the spectrum is shifted to higher velocities with respect to the nonmagnetic component. Therefore we have attempted to describe the nonmagnetic component by two doublets, D1 and D2, with different IS but close QS values (Table 1). This model yields a good fit to the experimental spectra (Fig. 9), but the areas of the doublets are nonmonotonic functions of temperature.

In order to improve the resolution of the magnetic component, the nonmagnetic component (previously fitted to the QS-distribution model) was subtracted from the measured spectrum. Figure 9 demonstrates that this procedure allows us to resolve magnetic components clearly. The mean values of the hyperfine parameters H_{hf}, QS, and IS for the magnetic components of quasicrystals of various compositions are listed in Table II. In what follows, we will use the mean

TABLE I. Parameters of hyperfine interaction derived from Mössbauer spectra of the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal using various models. IS is the isomer shift with respect to metallic iron at room temperature, $QS=e^2qQ/2$ is the quadrupole splitting, H_{hf} is the hyperfine magnetic field, $A_{\rm rel}$ is the relative spectrum area, Γ is the half-maximum linewidth, M_i are magnetic sextets, D_i are paramagnetic doublets, $\langle M\rangle$ and $\langle D\rangle$ are the mean values for the magnetic and nonmagnetic components

T, K	M,D	IS, mm/s	QS, mm/s	H_{hf}, kOe	$A_{\rm rel}$, %	Γ, mm/s	χ^2	Model
	$\langle D\rangle$	0.340	0.503	–	88.0	0.280	2.304	1
	$\langle M\rangle$	0.512	−0.069	95.5	12.0	0.334		
	D	0.368	0.495	–	87.7	0.454	6.393	
	$M1$	0.514	−0.067	105.6	5.4	0.351		2
4.2	$M2$	0.528	−0.107	85.5	6.9	0.250		
	$D1$	0.398	0.526	–	52.5	0.369	2.148	
	$D2$	0.256	0.425	–	35.5	0.415		3
	$M1$	0.509	−0.054	106.6	4.9	0.342		
	$M2$	0.515	−0.084	85.9	7.1	0.330		
293	$D1$	0.399	0.417	–	18.2	0.310	2.314	4
	$D2$	0.213	0.445	–	81.8	0.429		
350	$D1$	0.368	0.395	–	20.0	0.241	2.269	4
	$D2$	0.187	0.439	–	80.0	0.385		
	$\langle D\rangle$	0.210	0.454	–	100.0	0.280	2.191	5

Note: 1 is the fitting model with a distribution $P(QS)$ of quadrupole splittings for the central, nonmagnetic component of the spectrum and with a mean hyperfine magnetic field $\langle H_{hf}\rangle$ for the magnetic component of the spectrum; 2 is the fitting model based on one asymmetric paramagnetic doublet and two magnetic sextets; 3 is the fitting model with two symmetric doublets and two sextets; 4 is the fitting model with two symmetric doublets (at 293 K the area of the magnetic components is ~0.4% and can be neglected); 5 is the model with the distribution $P(QS)$.

value of the hyperfine magnetic field $\langle H_{hf}\rangle$, which is defined as $\Sigma H_iA_i/\Sigma A_i$, where A_i is the relative area of the ith component.

At higher temperatures the magnetic component transforms to a nonmagnetic one typical of the quasicrystalline

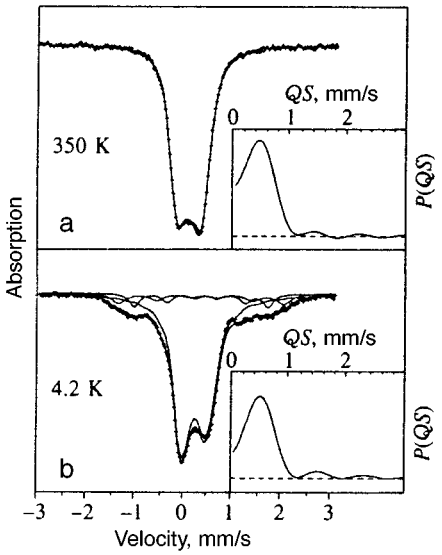

FIG. 8. Mössbauer spectra of the$Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal at (a) 350 K and (b) 4.2 K. The experimental spectra are approximated by the model with a distribution of quadrupole splitting $P(QS)$. The inserts show $P(QS)$ for the paramagnetic components of the spectra.

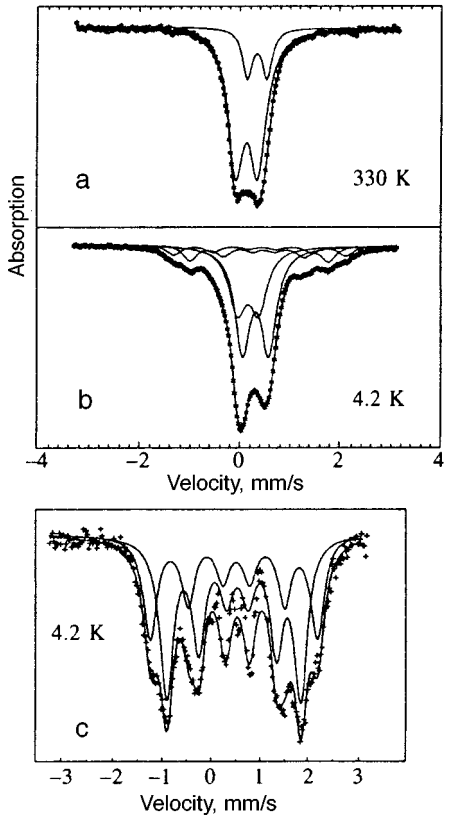

FIG. 9. Mössbauer spectra of the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal at (a) 330 K and (b) 4.2 K. (c) The spectrum after subtraction of the paramagnetic component in the form of the $P(QS)$ distribution from the experimental spectrum recorded at 4.2 K. The experimental data are approximated by (a) two symmetric doublets, (b) two doublets and two magnetic sextets, and (c) two magnetic sextets.

phase, and no additional lines due to extra phases are detected. This conclusion can be tested to high accuracy owing to the high sensitivity of the χ^2 factor to the fitting procedure, and this is an additional argument against the presence of extra iron phases in $Al_{70-x}B_xPd_{30-y}Fe_y$.

3.4. Comparative magnetic and Mössbauer parameters in quasicrystals and related amorphous and crystalline materials

An important point is comparison of macromagnetic characteristics and hyperfine parameters of Mössbauer spectra of studied quasicrystals to those of corresponding amorphous and crystalline materials. The center shift of the Mössbauer spectrum, CS, and H_{hf} are functions of temperature. At $T>0$ CS includes the isomer shift IS and the second-order Doppler shift, whereas H_{hf} can change owing to relaxation effects. Therefore parameters of different materials should be compared at the same temperature close to 0 K. In the Al–Fe, Al–B–Fe, Al–Pd–Fe, and Al–B–Pd–Fe systems, we have found neither crystalline nor amorphous compounds with values of IS, H_{hf}, and T_C close to those measured in the studied quasicrystals. It is known, in particular, that the iron solubility in the face-centered metallic aluminum is very low, and at large Fe concentrations the solid solution of Fe and Al coexists with the Fe_4Al_{13} compound.[26] In both Fe–Al dilute solution and Fe_4Al_{13}, iron atoms are nonmagnetic, and $IS(4.2\ \mathrm{K})\approx 0.6$ mm/s and 0.12 mm/s, respectively.[26] Given these parameters, the presence of these phases in the studied quasicrystals is out of the question.

$Al_{70}Pd_{20}Fe_{10}$ (Ref. 5) and $Al_{70}Pd_{16}Fe_{14}$ (Ref. 6) quasicrystals are nonmagnetic at temperatures above 4.2 K. Therefore the ferromagnetic properties of $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals, which contain boron, are probably due to local bonds between Fe and B, as in Mn-bearing i-QC.[16,17] The hyperfine Mössbauer parameters and temperatures of magnetic ordering in amorphous and crystalline Fe–B compounds of various compositions were studied in detail by Chien and co-workers.[27–29]

There are three crystalline Fe–B compounds, namely FeB, Fe_2B, and Fe_3B. Their values of H_{hf} at 4.2 K and T_C equal (131 kOe, 598 K), (242 kOe and 252 kOe, 1015 K for two sites in the lattice), and (242 kOe, 284 kOe, and 305 kOe, 800 K for three sites in the lattice).[27] It is obvious that these values of H_{hf} and T_C in crystalline materials are notably higher than in $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals (Table II). The crystalline FeB has the parameter $H_{hf}(4.2\ \mathrm{K})=131$ kOe close to that of quasicrystals, but its isomer shift $IS(4.2\ \mathrm{K})=0.37$ mm/s is considerably lower than $IS(4.2\ \mathrm{K})=0.51$ mm/s typical of magnetic type Fe atoms in quasicrystals (Table II). In the crystalline Fe_2B the value of $IS(4.2\ \mathrm{K})$ is even lower (~ 0.22 mm/s). Thus, the ferromagnetism detected in $Al_{70-x}B_xPd_{30-y}Fe_y$ icosahedral quasicrystals is not caused by crystalline Fe–B inclusions in the studied materials.

In the Fe_xB_{100-x} amorphous system,[28] we have found the composition $Fe_{45-50}B_{55-50}$ with the parameters H_{hf} and T_C close to those in the studied quasicrystals, but the parameter $IS(4.2\ \mathrm{K})$ in amorphous materials is much lower than in the quasicrystals. In this system, the amorphous $Fe_{47}B_{53}$ compound has the highest $IS(4.2\ \mathrm{K})\approx 0.33$ mm/s,[28] which is, nevertheless, considerably lower than in quasicrystals. Moreover, it is well known that Fe_xB_{100-x} amorphous alloys with $x\leqslant 50$ can be fabricated only by the sputtering method,[28] but cannot be produced by the melt quenching technique used in our experiments. In actuality, we could not find any signs of amorphous phase in x-ray diffraction patterns (Fig. 2). Thus, we have come to the conclusion that ferromagnetism in icosahedral $Al_{70-x}B_xPd_{30-y}Fe_y$ is an inherent property of the quasicrystalline phase, and local atomic configurations

TABLE II. Parameters of hyperfine interaction for the magnetic component of the Mössbauer spectrum of $Al_{70-x}B_xPd_{30-y}Fe_y$ icosahedral quasicrystal at $T=4.2$ K. IS is the isomer shift with respect to metallic iron at room temperature, $\varepsilon=e^2qA/4$ is the quadrupole shift, $\langle H_{hf}\rangle$ is the mean hyperfine magnetic field, A_{rel} $A_{\mathrm{m}}/A_{\mathrm{tot}}$ is the relative area of the magnetic component. The Curie temperature T_c and saturation magnetization σ_s are derived from magnetic measurements, μ_{eff} is the magnetic moment per Fe atom.

x	y	IS, mm/s	ε, mm/s	$\langle H_{hf}\rangle$, kOe	A_{rel}, %	T_C, K	σ_s, emu/g	μ_{eff} (μ_B)
5.0	15	0.516(5)	−0.097(5)	97.5(5)	14.1(5)	304	4.3	1.55
10.0	15	0.502(5)	−0.062(5)	96.4(5)	14.3(5)	258	7.2	2.50
7.5	12.5	0.520(5)	−0.087(5)	95.0(5)	10.0(5)	247	3.6	1.86
7.5	15.0	0.508(5)	−0.082(5)	95.3(5)	12.3(5)	287	5.8	2.37
7.5	17.5	0.497(5)	−0.068(5)	95.5(4)	14.1(5)	294	7.8	2.69

around magnetic Fe atoms in these *i*-QC might be similar to those in the $Fe_{45-50}B_{55-50}$ amorphous phase.

3.5. Electronic properties

Our analysis of Mössbauer spectra indicates that iron atoms in $Al_{70-x}B_xPd_{30-y}Fe_y$ *i*-QC are distributed among two types of lattice sites, magnetic and nonmagnetic, and 12 to 15% of iron atoms are magnetically ordered at low temperatures. The average value of *IS*(4.2 K) for nonmagnetic iron atoms is 0.34 mm/s, which is typical of most quasicrystals,[3–5,12] whereas this parameter of magnetic iron atoms is notably higher (0.51 mm/s). In the general case, an increase in *IS* indicates a lower density of *s*-electrons at iron nuclei, and this implies that the material becomes less metallic and more insulating. Our measurements of the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ resistivity in the 4.2–300 K range yielded $\sim 3\times 10^{-3}\Omega\cdot$cm. This value is two or three orders of magnitude higher than that of metallic iron and close to Mott's minimal conductivity $\sim 200\,(\Omega\cdot\text{cm})^{-1}$. A material with a lower conductivity becomes insulating at 0 K. This suggests electronic localization in quasicrystals.[30] The higher value of *IS* for magnetic iron atoms is consistent with the concept of electron localization at these sites and it may correlate with appearance of the magnetic state in these sites. We have found out that the temperature dependencies of isomer shifts of magnetic and nonmagnetic iron atoms are similar, but cannot be described in the Debye approximation.

Previously[8,31–33] an assumption about two types of lattice sites for transitional atoms in *i*-QCs was made. The dimension of a nonmagnetic site is smaller than that of a magnetic one. The presence of two types of sites can be ascribed to the intrinsic disorder typical of icosahedral quasicrystals[15] and follows from the general principles of the theory of energy bands in transitional metals and their alloys.[34,35] This concept has been confirmed by some experiments.[11–15,32]

The presence of two types of iron sites in $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals is directly confirmed not only by the coexistence of magnetic and nonmagnetic atoms, but also by the observed values of isomeric shifts. Tables I and II clearly demonstrate that, irrespective of the data processing technique, the parameter *IS* for magnetic atoms is much higher than for nonmagnetic ones. This means that the density of *s*-electrons at ^{57}Fe nuclei is lower for magnetic Fe atoms. This may be caused (a) by a chemical effect of different local atomic environment and/or (b) by a larger atomic volume of magnetic sites as compared to that of nonmagnetic ones. The comparison with the Fe_xB_{100-x} amorphous system indicates that the mechanism (a) cannot provide the observed effect, hence the mechanism (b) is the most probable. Table II shows that the parameter *IS* in $Al_{70-x}B_xPd_{30-y}Fe_y$ quasicrystals in the region $5\leqslant x\leqslant 10$, $12.5\leqslant y\leqslant 17.5$ is almost constant. This indicates that the iron–ligand bond in these compounds has a similar character.

3.6. Magnetic properties derived from Mössbauer spectra

The hyperfine magnetic field H_{hf} changes slowly with both *x* and *y* (Table II), and its average saturation value is approximately 96 kOe. This low value of $\langle H_{hf}\rangle$ can be ascribed to the low magnetic moment of Fe atoms, which is typical of quasicrystals. The field H_{hf} at an iron nucleus in a metallic system can be expressed as[37,38]

$$H_{hf}=a\mu_{\text{Fe}}+b\mu, \tag{1}$$

where the first term on the right is the local contribution of the electronic magnetic moment μ_{Fe} of this iron atom, and the second term is the contribution of conduction electrons spin polarized by neighboring magnetic moments. The first term is usually negative, and the second can be either negative or positive, depending on the magnetic structure and type of interaction.[39] The typical value of the constant *a* for most alloys is -142 kOe/μ_{Fe}.[37,38] Given that the conductivity of the studied quasicrystals is close to the metal–dielectric transition, we neglect the parameter *b* on the right of Eq. (1). Then using the relation $|H_{hf}|=142\mu_{\text{Fe}}$ and the experimental value $\langle H_{hf}\rangle\approx 96$ kOe for the material with $x=7.5$ and $y=15$, we obtain $\mu_{\text{Fe}}=0.68\ \mu_B$. This result, however, is much lower than the magnetic moment per iron atom $\mu_{\text{eff}}=2.37\ \mu_B$ derived from our magnetic measurements (Table II). Similar results were obtained by Shinohara and co-workers[22] for manganese magnetic moments in $Al_{64}B_6Pd_{15}Mn_{15}$ *i*-QC. The value of μ_{Mn} derived from the NMR frequency of ^{55}Mn was notably different from the value derived from magnetization measurements. This probably means that Eq. (1), derived for metals, is not valid for highly resistive quasicrystals. The values of μ_{eff} listed in Table II were derived from the values of σ_s under the assumption that only a fraction A_{rel} (%) of all Fe atoms are magnetic, where $A_{\text{rel}}=A_{\text{m}}/A_{\text{tot}}$ is the relative area of the magnetic component. In the $Al_{62.5}B_{7.5}Pd_{30-y}Fe_y$ system, we have detected correlation between the values of σ_s and A_{rel}, which indicates an increase in the magnetic-phase volume with the Fe content.

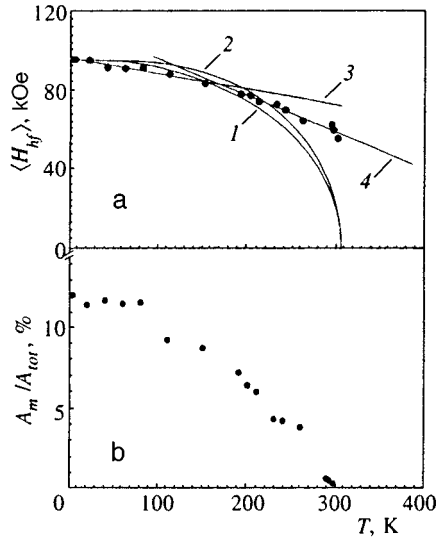

FIG. 10. (a) Mean hyperfine magnetic field $\langle H_{hf}\rangle$ and (b) relative area $A_{\text{m}}/A_{\text{tot}}$ of the magnetic component of the Mössbauer spectrum of the $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$ quasicrystal as functions of temperature. Solid lines are theoretical curves calculated by (1 and 2) Brillouin functions with $S=5/2$ and $S=1$, respectively, (3) model of collective spin excitations, and (4) model of critical indices for $\beta=1$ in the spin-glass approximation.

Figures 10a and 10b show the field $\langle H_{hf}\rangle$ and area $A_{\rm rel}$ as functions of temperature for $Al_{62.5}B_{7.5}Pd_{15}Fe_{15}$. We have found that $A_{\rm rel}$ does not abruptly drop to zero at the Curie temperature. In the range 4.2–80 K this area is approximately constant, but above 80 K it gradually decreases with temperature and vanishes near T_C (Fig. 10b). The curve of $\langle H_{hf}\rangle$ versus temperature also has a peculiar shape (Fig. 10a).

We tried to describe the curve $\langle H_{hf}\rangle = f(T)$ using various theoretical models, such as the model of critical indices, Brillouin functions, two-dimensional Ising model, spin-glass model, and collective-spin-excitations model, but none of these models could approximate the experimental curve throughout the studied temperature range. The features described above, however, suggest that the magnetic properties of these *i*-QCs can be described in terms of superparamagnetic behavior of small magnetic particles or magnetic clusters.

3.7. Superparamagnetic approach

The behavior of an assembly of small magnetic particles is governed by the particle volume V and depends on the relaxation time τ[40]:

$$1/\tau = f_0 \exp[-KV/kT], \tag{2}$$

where K is the anisotropy constant and f_0 is a frequency factor approximately equal to 10^9 s^{-1}.[41,42] The parameter K can be derived from magnetic measurements[18] using the relationship

$$H_c = (2KM_s)[1-(25kT/KV)^{1/2}],$$

where H_c is the coercivive force and M_s is the saturation magnetization.[40] For the quasicrystals with $x=7.5$ and $y=15$ at 4.2 K we have $M_s = 4.2$ emu/g and $H_c = 1200$ Oe.[18] From these data we have in the low-temperature approximation $K = 5.38\times 10^3$ erg$\cdot$cm^{-3}.

With certain assumptions, we can derive from Mössbauer spectra the particle size in this sample. The characteristic time in Mössbauer experiments is the time τ_L of Larmor precession in the field H_{hf}, and at $H_{hf}\approx 100$ kOe it equals $\sim 0.5\times 10^{-8}$ s. From Eq. (2) we obtain $V\approx 1.6$ kT/K. On the other hand, there is a so-called blocking temperature T_b for particles of a constant size. Below this temperature, the measured sample magnetization is stable, and Mössbauer spectra yield the hyperfine magnetic splitting. For uniaxial particles $T_b \approx KV/1.6k$. The parameter T_b can be evaluated from the curve of $A_{\rm rel} = f(T)$. Figure 10b shows that $A_{\rm rel}$ is almost constant in the temperature range of 4.2–80 K and drops at higher temperatures. By taking $T_b^{\min} = 80$ K for the lowest blocking temperature, we can estimate the minimal volume of magnetic particles, $V_{\min}\approx 3.30\times 10^{-18}$ cm^3, which corresponds approximately to the particle diameter $D_{\min}\approx 185$ Å. The temperature at which $A_{\rm rel}=0$ can be defined as the maximal blocking temperature $T_b^{\max}$. It corresponds to the maximum volume $V_{\max}$ of particles tested in a Mössbauer experiment. We have derived from the curve $A_{\rm rel} = f(T)$ the values of $T_b^{\max} = 300$ K and $V_{\max}\approx 1.24\times 10^{-17}$ cm^3, which corresponds to $D_{\max}\approx 290$ Å.

Thus, the model of small magnetic particles yields a minimum particle size of about 185 Å. Given the dimension of the ''unit cell parameter'' quoted above $a_R \simeq 4.52$ Å, a particle of the minimum size contains about 3×10^4 ''unit cells,'' and this particle can be treated as bulk material. Moreover, it turned out that $T_b^{\max}$ value is very close to the Curie temperature of this material (305 K). All this means that its magnetic properties can be described in the bulk material approximation.

3.8. Temperature dependence of hyperfine magnetic field

Figure 10a demonstrates that the experimental curve $\langle H_{hf}\rangle = f(T)$ cannot be approximated by the Brillouin functions, but it can be described by two linear laws, one in the low-temperature range 4.2–150 K and the other in the range 160–300 K (Fig. 10a). Morup *et al.*[43,44] showed that in isolated particles with uniaxial anisotropy, in the low-temperature range of collective spin excitations at $kT/KV < 0.1$ (which corresponds to $\tau > 10^{-7}$ s in Mössbauer experiments), the observed dimensionless hyperfine field should be described by a linear function of temperature:

$$H_{\rm obs}/H_0 = 1 - kT/2KV. \tag{3}$$

We have found that in the quasicrystalline sample with $x=7.5$ and $y=15$ the curve of $\langle H_{hf}\rangle = f(T)$ can be approximated in the range 4.2–150 K by the linear function given by Eq. (3) with a slope of -0.07473 kOe/deg (curve *3* in Fig. 10a). With the experimental value of K, Eq. (3) yields an estimate of the particle volume, $V = 1.65\cdot 10^{-17}$ cm^3. It is interesting that this independent estimate is close to $V_{\max}$ quoted in Sec. 3.7. This agreement supports the applicability of the cluster approximation to magnetic behavior of the studied materials.

In the high-temperature range 160–300 K, curve *4* in Fig. 10a corresponds to the model of critical indices, $H_{\rm obs}/H_0 = (1-T/T_N)^{\beta}$, where $\beta \approx 1$, which is typical of materials with spin-glass (SG) ordering. A linear decrease of magnetization at $T > T_{\rm SG}$ was proposed for concentrated spin glasses on the basis of theoretical concepts.[45–47] The theory assumes existence of magnetic clusters of various sizes and interaction between the clusters depending on the magnetic phase concentration.

Similar behavior $\langle H_{hf}\rangle = f(T)$ was observed in other $Al_{70-x}B_xPd_{30-y}Fe_y$ compounds. Thus, the magnetic properties of these materials can be understood in terms of small magnetic particles or magnetic clusters. A similar interpretation was suggested for ferromagnetic properties of $Al_{70-x}B_xPd_{15}Mn_{15}$.[17] Feasibility of clustering of Mn atoms in icosahedral and decagonal Al–Mn quasicrystals was also analyzed by Machado *et al.*[48,49]

4. CONCLUSIONS

We have discovered a new system of iron-bearing icosahedral quasicrystals, i.e., $Al_{70-x}B_xPd_{30-y}Fe_y$ ($5<x<10$ and $10<y<20$) with ferromagnetic properties, T_C ranging between 280 and 340 K, and maximum magnetization $\sigma_s \approx 7.5$ emu/g. Magnetic measurements indicate that the quasicrystals are inhomogeneous magnetic systems. Möss-

bauer spectra show that only about 12–15% of all iron atoms have a magnetic moment, and their magnetic behavior can be interpreted in terms of large magnetic Fe clusters with a size of 185–290 Å. A cluster of such a size contains about 4×10^4 rhombohedral ''unit cells'' that make up the Penrose three-dimensional tiling, and it can probably be treated as a bulk material.

None of the known crystalline or amorphous materials has magnetic or Mössbauer parameters close to those of these quasicrystals. Only the$Fe_{\sim50}B_{\sim50}$ amorphous alloy has close values of T_C and H_{hf}, but its isomer shift is different from that of the quasicrystals. The saturation magnetization of these quasicrystals increases with the concentrations of both Fe and B, but the hyperfine magnetic field and isomer shift values are almost constant. This implies that both the local crystal-chemical and magnetic environment of iron atoms in a cluster do not change, and only the cluster size and/or number of clusters change with the content of Fe or B. Bonding between Fe and B, similar to that in amorphous $Fe_{\sim50}B_{\sim50}$, is thought to be mainly responsible for the creation of the local magnetic moment at Fe atoms.

Two circumstances which can lead to coexistence of magnetic and nonmagnetic atoms in quasicrystals have been discussed in the literature, namely, the existence of specific large and small aluminum ''caves'' for atoms of transition elements on the one hand, and the continuous distribution of interatomic distances on the other (see Refs. 3 and 4 and references therein). The larger value of isomer shift for Fe atoms at magnetic sites, as compared to nonmagnetic sites, found in our experiments, is in agreement with the concept of a larger atomic volume of magnetic sites.[36]

We are indebted to T. V. Dmitrieva, D. M. Lin, and Ch. R. Wang for help in computer processing of experimental data and fruitful discussions. The work was supported by the National Scientific Council of the Republic of China and the Russian Fund for Fundamental Research.

[1] *The Physics of Quasicrystals*, P. J. Steinhardt and S. Ostlund (eds.), World Scientific, Singapore (1987).
[2] *Quasicrystals*, K. H. Kuo and T. Ninomiya, (eds.) World Scientific, Singapore (1991).
[3] Z. M. Stadnik, G. Stroink, H. Ma, and G. Williams, Phys. Rev. B **39**, 9797 (1989).
[4] R. A. Dunlap and D. W. Lawther, Materials Science and Engineering **10**, 141–185 (1993).
[5] C. R. Wang, S. T. Lin, and Y. C. Chen, J. Phys.: Condensed Matter **6**, 10747 (1994).
[6] K. Fukamichi, T. Kikuchi, Y. Hattori, A. P. Tasi, A. Inoue, and T. Masumoto, in *Proc. China–Japan Seminars on Quasicrystals*, K. H. Kuo and T. Ninomiya (eds.), World Scientific, Singapore (1990), p. 256.
[7] V. Srinivas, R. A. Dunlap, M. E. McHenry, and R. C. O'Handley, J. Appl. Phys. **67**, 5879 (1990).
[8] M. Eibschütz, M. E. Lines, H. S. Chen, J. V. Waszczak, G. Papaefthymiou, and R. B. Frankel, Phys. Rev. Lett. **59**, 2443 (1987).
[9] M. E. McHenry, V. Srinivas, D. Bahadur, R. C. O'Handley, D. J. Lloyd, and R. A. Dunlap, Phys. Rev. B **39**, 3611 (1989).
[10] V. Srinivas and R. A. Dunlap, J. Phys.: Condensed Matter **3**, 2411 (1991).
[11] S. Nasu, M. Miglierini, and T. Kuwano, Phys. Rev. B **45**, 12778 (1992).
[12] M. Miglierini and S. Nasu, Materials Transactions, JIM **34**, 178 (1993).
[13] M. Miglierini and S. Nasu, J. Phys. Soc. Jap. **60**, 2135 (1991).
[14] Z. M. Stadnik and G. Stroink, Phys. Rev. B **43**, 894 (1991).
[15] Z. M. Stadnik and G. Stroink, Phys. Rev. B **44**, 4255 (1991).
[16] Y. Yokoyama, A. Inoue, and T. Masumoto, Materials Transactions, JIM **33**, 1012 (1992).
[17] Y. Yokoyama, A. Inoue, H. Yamauchi, M. Kusuyama, and T. Masumoto, Jpn. J. Appl. Phys. **33**, 4012 (1994).
[18] C. R. Lin, C. M. Lin, S. T. Lin, and I. S. Lyubutin, Phys. Lett. A **196**, 365 (1995).
[19] H. Frauenfelder, *The Mössbauer Effect*, Benjamin, New York (1962), p. 45.
[20] V. Elser, Phys. Rev. B **32**, 4892 (1985).
[21] S. Matsuo, H. Hakano, T. Ishimasa, and M. Mori, J. Phys. Soc. Jap. **62**, 4044 (1993).
[22] T. Shinohara, Y. Yokoyama, M. Sato, A. Inoue, and T. Masumoto, J. Phys.: Condensed Matter **5**, 3673 (1993).
[23] B. Window, J. Phys. E. **4**, 401 (1971).
[24] Z. M. Stadnik and G. Stroink, Hyperfine Interactions **69**, 643 (1992).
[25] V. Srinivas and R. A. Dunlap, Philos. Mag. B **64**, 475 (1991).
[26] R. S. Preston and R. Gerlach, Rhys. Rev. B **3**, 1519 (1971).
[27] C. L. Chien, D. Musser, E. M. Gyorgy, R. C. Sherwood, H. S. Chen, F. E. Luborsky, and J. L. Walter, Phys. Rev. B **20**, 283 (1979).
[28] C. L. Chien and K. M. Unruh, Phys. Rev. B **25**, 5790 (1982).
[29] N. A. Blum, K. Moorjani, T. O. Poehler, and F. G. Satkiewicz, J. Appl. Phys. **53**, 2074 (1982).
[30] T. Klein, C. Berger, D. Mayou, and F. Cyrot-Lackmann, Phys. Rev. Lett. **66**, 2907 (1991).
[31] W. W. Warren, H. S. Chen, and G. P. Espinosa, Phys. Rev. B **34**, 4902 (1986).
[32] M. Eibshutz, M. E. Lines, H. S. Chen, and J. V. Waszczak, Phys. Rev. B **38**, 10038 (1988).
[33] K. Edagawa, H. Ino, S. Nasu, K. Kimura, S. Takeuchi, T. Shinjo, K. Koga, T. Shimizu, and H. Yasuoka, J. Phys. Soc. Japan **56**, 2629 (1987).
[34] V. L. Moruzzi and P. M. Marcus, Phys. Rev. B **38**, 1613 (1988).
[35] V. L. Moruzzi, Phys. Rev. B **41**, 6939 (1990).
[36] L. R. Walker, G. K. Wertheim, and V. Jaccarino, Phys. Rev. Lett. **6**, 98 (1961).
[37] O. Eriksson and A. Svane, J. Phys.: Condens. Matter **1**, 1589 (1989).
[38] Z. M. Stadnik and G. Stroink, Hyperfine Interactions **47**, 275 (1989).
[39] A. J. Freeman and R. E. Watson, in *Magnetism*, G. T. Rado and H. Suhl (eds.), Academic Press, New York (1965), Vol. IIA, p. 167.
[40] B. D. Cullity, *Introduction to Magnetic Materials*, University of Notre Dame, Indiana (1972) pp. 383, 415.
[41] W. Kundig, H. Bommel, G. Constabaris, and R. H. Lindquist, Phys. Rev. **142**, 327 (1966).
[42] M. F. Thomas and C. E. Johnson, in *Mössbauer Spectroscopy*, D. P. E. Dickson and F. J. Berry (eds.), Cambridge University Press (1986), p. 143.
[43] S. Morup, J. Mag. Mag. Mater. **37**, 39 (1983).
[44] S. Morup, J. A. Dumesic, and H. Topsoe, in *Applications of Mössbauer Spectroscopy*, Vol. II, R. L. Cohen (ed.), Academic Press, New York (1980), Ch. 1, p. 1.
[45] S. B. Liao, S. M. Bhagat, M. A. Manheimer, and K. Moorjani, J. Appl. Phys. **63**, 4354 (1988).
[46] E. M. Jackson, S. B. Liao, S. M. Bhagat, and M. A. Manheimer, J. Magn. Magn. Mater. **80**, 229 (1989).
[47] M. Lubecka and L. J. Maksymowicz, Phys. Rev. B **44**, 10106 (1991).
[48] F. L. A. Machado, W. G. Clark, L. J. Azevedo, D. P. Yang, W. A. Hines, J. I. Budnik, and M. X. Quan, Solid State Commun. **61**, 145 (1987).
[49] F. L. A. Machado, W. G. Clark, D. P. Yang, W. A. Hines, L. J. Azevedo, B. C. Giessen, and M. X. Quan, Solid State Commun. **61**, 691 (1987).

Translation provided by the Russian Editorial office.

Emergence of heterophase structures near phase transitions in photoferroelectric materials

R. F. Mamin

Kazan Physicotechnical Institute, Kazan Scientific Center of the Russian Academy of Sciences, 420029 Kazan, Russia
(Submitted 22 December 1995; resubmitted 28 November 1996)
Zh. Éksp. Teor. Fiz. **111**, 1465–1476 (April 1997)

The present study shows that under intense illumination, various spatially nonuniform states may emerge near structural phase transitions in systems with semiconducting properties. Different traveling and stationary autosoliton states are studied, and so is a periodic heterophase structure with alternating para- and ferroelectric phases, which results from a redistribution of electrons on trapping levels. The problem is solved for the one-dimensional case. The emergent behavior is related to the dynamics of the ferroelectric semiconductor as a whole and is an example of synergistic behavior. © *1997 American Institute of Physics.* [S1063-7761(97)02504-3]

1. INTRODUCTION

Lately there has been an upsurge of interest in various fields of science in studies of complex systems consisting of several subsystems.[1,2] On the one hand, this is due to the realization that systems consisting of two or more subsystems may have properties that cannot be attributed to any subsystem separately. On the other hand, studies of complex systems in various fields of science have revealed qualitatively similar motives in the behavior in the entire range of emergent new qualities. In particular, it appears that new states emerge when there is an influx of energy. All these facts have led to the development of a new science, synergetics,[1,2] devoted to establishing the laws governing the emergence of new qualities and to developing methods for analyzing complex systems. One example of complex systems is presented by ferroelectric semiconductors, a class of substances manifesting new properties related to the interaction of the electron and lattice subsystems.

Another interesting problem in this connection is the study of heterophase structures and their relationship to defects in crystals.[3–12] The objects that were investigated include large-scale periodic structures emerging because of the effect of elastic stresses in the presence of random fields of the local transition temperature[6] and heterophase states emerging because of the screening of polarization by free electrons in heavily doped semiconductors.[7,8]

A special line of research deals with the situation in which heterophase states emerge as a result of external illumination. Bursian *et al.*[3] and Baryshnikov *et al.*[4] studied the effects when a periodic heterophase structure is formed in the field of intense laser radiation. The formation of such a structure is directly related to the periodic electric field of the laser light. Another situation emerges when crystals are illuminated by ordinary white light.[12] In this case periodic structures also emerge. But since in the given situation the external agent (light) is characterized by no distinct periodicity, the formation of structures and their spatial size and periodicity are due to the properties of the given crystals as complex systems and to the interaction between the constituent subsystems, additionally stimulated by external illumination.

The present study is devoted to the case where the light possesses no distinct periodicity (white light), so that the nonuniform states emerge because of the cooperative effects of the electron and lattice subsystems in ferroelectric semiconductors. The result of illumination of the semiconductors is the generation of conduction electrons, which in turn fill the trapping levels and hence change the balance in the crystal. Such an approach corresponds to the situation studied in Refs. 11 and 12.

As is known, in ferroelectric semiconductors near photostimulated phase transitions there is strong interaction between the electron and lattice subsystems.[13] This manifests itself in such effects as vibrations of the phase boundary in SbSI (Ref. 14) and order-parameter oscillations in proustite (Ag_3AsS_3).[15] The dynamic behavior of ferroelectric semiconductors and the synergistic behavior of the given system were established on the basis of the model suggested in Ref. 16. But in addition to the dynamic states mentioned earlier, a synergistic system may have nonuniform states related to the correlation of the behavior of the interacting subsystems in space.[17,18] So-called autosoliton states may be formed.[17] The evolution of an autosoliton may lead to the formation of a periodic heterophase structure with a period determined by the diffusion length of the electron subsystem. This could be the mechanism of formation of a periodic heterophase structure observed in photoferroelectric materials.[12]

The present work studies the formation of various heterophase structures in ferroelectric semiconductors under intense illumination. The study is done for the one-dimensional case.

2. THE MODEL AND THE METHOD OF BUILDING THE SOLUTION

Cooperataive effects in the electron and lattice subsystems in ferroelectric semiconductors manifest themselves in the shift in the transition temperature as the concentration

of electrons in the traps changes and in the variation of the energy intervals of the electron spectrum as the order parameter varies.[16]

We write the equation describing the dynamics of the order parameter η in the form of a Landau–Khalatnikov relaxation equation:

$$\frac{d\eta}{dt}=-\Gamma\left[(\alpha+am)\eta+\beta\eta^3+\gamma\eta^5-\delta\frac{\partial^2\eta}{\partial x^2}\right]. \tag{1}$$

Here α, β, γ, and δ are the coefficients in the expansion of the lattice part of the thermodynamic potential in powers of the order parameter and its derivatives ($\alpha=\alpha'(T-T_0)$), Γ is a kinetic coefficient, and the term am determines the shift of the transition temperature due to electrons on trapping levels (traps).

The equation that describes the dynamics of the concentration m of the electrons on trapping levels and depends on the order parameter has the following form:[16]

$$\frac{dm}{dt}=D\frac{\partial^2 m}{\partial x^2}-Q(\eta,m), \tag{2}$$

$$Q(\eta,m)\equiv -J(M-m)+mA(\eta),$$

$$J=\gamma_n n_0, \quad A(\eta)=\gamma_n N_c\exp\left(-\frac{u_0+\tilde{a}\eta^2}{kT}\right).$$

Here J is proportional to the illumination intensity I, n_0 is the concentration of conduction electrons, with $n_0\propto I$, M is the concentration of trapping levels, γ_n is a kinetic coefficient, N_c is the density of states in the conduction band, $u=u_0+\tilde{a}\eta^2(x)$ is the energy interval between the bottom of the conduction band and the trapping levels, which depends on the order parameter,[16] and D is the diffusion coefficient.

Let us examine the possibility of emergence and the dynamics of nonuniform states caused by the redistribution of electrons in the traps. Here we are dealing with a typical synergistic system consisting of two interacting subsystems (the lattice subsystem and the electron subsystem) in the presence of an energy influx. Nonuniform states emerge when the only stable state in the system becomes unstable because of illumination. The existence of nonuniform states and the distribution of the order parameter in such states are directly related to the distribution of the trapped electrons in space.

The dynamics of the system is determined by the fact that, first, the characteristic times of variation of the order parameter η and the electron concentration m at the trapping levels differ substantially ($\epsilon=\tau_\eta/\tau_m\ll 1$) and, second, the characteristic diffusion lengths for these variables also differ substantially ($\lambda=L_\eta/L_m\ll 1$). This makes it possible to solve the problem by separating fast and slow processes.

In the region of rapid variations of the order parameter at the phase boundary, there exists an exact solution[11,18] for the order parameter η with a constant electron concentration at the trapping levels:

$$\eta^2(\xi)=\frac{\eta_0^2}{1+\exp(-\xi/\Delta)},$$

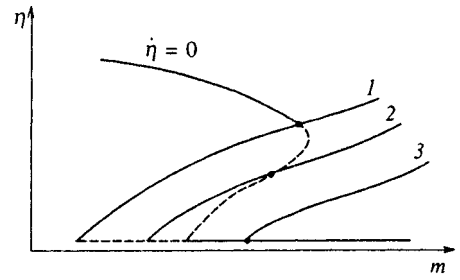

FIG. 1. Relative positions of the zero-isoclines of Eqs. (1) and (2) (curves *1*, *2*, and *3* correspond to $\dot{m}=0$) for different values of α. The cases *1* and *3* correspond to uniform metastable states, and the case *2* corresponds to an uniform unstable state.

$$\Delta=\left(-\frac{4\delta\gamma}{3\beta^2}\right)^{1/2}, \quad \eta_0^2=-\frac{3\beta}{4\gamma}, \tag{3}$$

where ξ is measured from the center of the corresponding domain wall. The electron concentration at the trapping levels in the vicinity of a static domain wall, m_0, is given by

$$m_0=\frac{3\beta^2-16\gamma\alpha}{16\gamma a}. \tag{4}$$

The value of the order parameter in the slow-motion region depends on the electron concentration at the trapping levels at the given point and can be found by solving Eq. (1) by employing the m-dependence of the stationary value of the order parameter, which in the paraelectric phase is zero and in the ferroelectric phase is given by

$$\eta^2(m)=\frac{-\beta+\sqrt{\beta^2-4\gamma(\alpha+am)}}{2\gamma} \tag{5}$$

for a given value of the electron concentration at the trapping levels.

The value of $m(x)$ in the slow-motion region can be found from diffusion equations. These equations are similar to the equation of motion of a particle in the potential field

$$U_1=-J\left(M-\frac{m}{2}\right)m+\frac{m^2}{2}A(0) \tag{6}$$

in the region where the order parameter is zero and in the potential field

$$U_3=J\left(M-\frac{m}{2}\right)m - \frac{\gamma^2 A(\eta(m))}{4a^2e^4}[y^2-(y+1)(2am\gamma e^2-3)], \tag{7}$$

$$y=e(2\gamma\eta^2(m)+\beta), \quad e=\frac{\tilde{a}}{2kT},$$

in the region where the order parameter is nonzero.

In nonequilibrium conditions, which emerge when there is an influx of energy, the system may become metastable and unstable in the uniform case. The possible unstable states are depicted in Fig. 1 in the form of intersections of zero-isoclines of the uniform system. In the nonuniform case, however, these unstable states may manifest them-

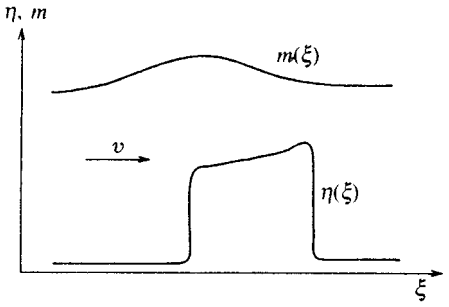

FIG. 2. Distributions of the order parameter η and the concentration m of the electrons on the trapping levels in a traveling autosoliton with a ferroelectric-phase domain in a metastable paraelectric phase, in a moving coordinate system ($\xi = x - vt$, where v is the velocity of the traveling autosoliton).

selves in the form of spatially nonuniform distributions of the parameters of the system: the order parameter and the electron concentration at the trapping levels.

Now let us describe the possible nonuniform states that can appear because of intense illumination. Such illumination may lead to the formation of nonuniform localized states, or autosolitons, and nonuniform states in the entire bulk with periodic phase alternation. Autosolitons emerge when under illumination the only steady state in the system becomes metastable. Traveling and static autosolitons may emerge in such a system.

3. A TRAVELING AUTOSOLITON

As noted earlier, a traveling autosoliton may emerge when the only steady state of the system becomes metastable under illumination. Such an autosoliton is the combination of a traveling domain (region) of one phase in another (a ferroelectric-phase domain in a metastable paraelectric phase, as depicted in Fig. 2, or a paraelectric-phase domain in a metastable ferroelectric phase) and the related moving spatial nonuniformity of the electron concentration at the traps.

A traveling autosoliton emerges in the following way. At a certain point in the sample, say, at the sample's boundary, as a result of an external agent or fluctuations there forms a region of the opposite phase, which begins to propagate since the state is metastable under variations of the order parameter. In the region of the newly formed phase the electron concentration at the traps changes, with the change increasing with the distance from the wavefront of the moving phase boundary, until the phase becomes unstable for the given value of the electron concentration at the traps. As a result the system goes back to the initial phase and the trailing edge of the domain in the form of a moving domain wall begins to emerge. Initially the back domain wall moves with a greater velocity than the front wall. But in the process of ''pursuit'' the velocity of the back wall decreases and becomes equal to that of the front wall. The result is a stable autosoliton moving with a constant velocity. The trailing edge of the domain is followed by an autosoliton tail in the form of a relaxation of the electron concentration at the traps to a metastable state and a related change in the order parameter in the case of a metastable ferroelectric phase.

In the region of rapid variations of the order parameter at the phase boundary with m_0 (Eq. (4)) constant there exists an exact solution $\eta(\xi)$ (Eq. (3)), where ξ is measured from the center of the corresponding phase boundary in a moving system of coordinates ($\xi = x - vt$). The velocity v of the traveling autosoliton is given by

$$v = v_0(1 - \sqrt{1 + \mu(m_0 - m_s)}),$$

$$v_0 = -\beta\Gamma\sqrt{\frac{\delta}{\gamma}}, \quad \mu = \frac{16\gamma a}{\beta^2}. \tag{8}$$

The order parameter in the slow-motion region depends as usual on the concentration of electrons on the trapping levels at the given point: it is determined by Eq. (5) in the ferroelectric phase and is zero in the paraelectric phase.

The values of $m(x)$ in the slow-motion region can be found by solving the diffusion equations in the potential fields (6) and (7). The boundary conditions for these equations are fixed by the following considerations. Directly in front of the autosoliton (at the front boundary) and far from the back boundary of the moving autosoliton the values of the system parameters are those of a uniform steady state, η_s and m_s, defined as the stationary solution of Eqs. (1) and (2). The values of the parameters at the back boundary (trailing edge) of the autosoliton domain, η_e and m_e, can be obtained by making the velocities at the front and back boundaries of the autosoliton domain equal:

$$v(m_e) = v(m_s), \quad \eta_e = \eta(m_e)$$

(see Eq. (5)).

The final shape of the spatial distribution of the electron concentration at the trapping levels is fixed by numerical integration of Eq. (2) in the moving system of coordinates. The result of such integration is depicted in Fig. 2.

4. STATIC AUTOSOLITON

A static autosoliton may emerge when as a result of an external agent or a fluctuation at defects there forms a region of the opposite phase and a related nonuniformity in the electron concentration at traps. If in the space of states of the system the given state lies in the attracting basin of the autosoliton state, the system relaxes to a state with a static autosoliton. A stable static autosoliton has certain dimensions and is characterized by certain distributions in space of the order parameter and the electron concentration at the trapping levels, both being determined primarily by diffusion processes in the electron subsystem.

In the region of rapid variations of the order parameter at the phase boundary with m_0 (Eq. (4)) constant there exists an exact solution $\eta(x)$ (Eq. (3)), where x is measured from the center of the corresponding stationary phase boundary.

The values of $m(x)$ in the slow-motion region can be found by solving the diffusion equations in the potential fields (6) and (7). The boundary conditions for these equations are fixed by the following considerations. Far from the autosoliton the system parameters are those of the uniform steady state, η_s and m_s, defined as the stationary solution of

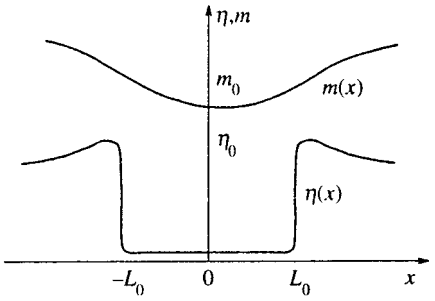

FIG. 3. Distribution of the order parameter η and the concentration m of the electrons on the trapping levels in a stationary autosoliton with a paraelectric-phase domain in a metastable ferroelectric phase.

Eqs. (1) and (2). The parameters at the autosoliton's center, η_c and m_c, can be obtained from the condition that the distribution of the concentration of electrons on the trapping levels at the phase boundary is smooth:

$$\int_{m_s}^{m_0} Q(\eta(m),m)\ dm + \int_{m_c}^{m_0} Q(\eta(m),m)\ dm = 0, \qquad (9)$$

where $\eta(m)$ corresponds to the value of the order parameter in the autosoliton. The halfwidth L_0 of the autosoliton is

$$L_0 = L\int_{m_c}^{m_0} G(m_c,m)\ dm, \qquad (10)$$

$$G(m_1,m_2) = \left(\frac{A(0)}{2}\right)^{1/2}\left[\int_{m_1}^{m_2} Q(\eta(m),m)\ dm\right]^{-1/2},$$

and for the case of an autosoliton paraelectric-phase domain in a ferroelectric phase has an exact solution,

$$L_0 = L\frac{2}{(J+1)^{0.5}}(l(m_s)-l(m_0)), \qquad (11)$$

$$l(m) = \ln[2(J+A(0))(U_3(m)-U_3(m_0))^{1/2} - J(M-m) + mA(0)],$$

where L is the characteristic diffusion length of the electron subsystem ($L^2 = DA^{-1}(0)$).

In the slow-motion region the distribution of the electron concentration at the trapping levels can be found by integrating the stationary part of Eq. (2):

$$|x| = L\int_{m_c}^{m} G(m_c,m)\ dm \quad \text{for } 0<|x|<L_0,$$

$$|x| - L_0 = L\int_{m_0}^{m} G(m_s,m)\ dm \quad \text{for } L_0<|x|<\infty. \qquad (12)$$

Since Eqs. (9)–(12) cannot be solved entirely in analytic form, the final result must be obtained by numerical integration. Figure 3 depicts the distribution of the order parameter η and the concentration m of the electrons on the trapping levels in an autosoliton for the case of a ferroelectric-phase autosoliton in the paraelectric phase.

The autosolitons described above are stationary entities, but the parameters of the system are such ($\epsilon \ll 1$ and $\lambda \ll 1$) that, according to Ref. 17, static autosolitons may become pulsed autosolitons, which requires studying the system in greater detail.

5. PERIODIC HETEROPHASE STRUCTURE

Now let us examine the possibility of emergence and the dynamics of a nonuniform state in the form of a periodic heterophase structure caused by a redistribution of electrons on the trapping levels. Periodic nonuniform states can only emerge when the uniform steady state of the system becomes absolutely unstable (see Fig. 1) as a result of illumination. Here the periodic heterophase structure and the periodic distribution of the electron concentration at the traps emerge in a self-consistent manner. The structure of the periodic state consists of the following elements: the regions of phase boundaries, where the order parameter rapidly changes while the electron concentration at the trapping levels remains constant, and the regions inside the domain of a single phase, where the variation of the order parameter follows the smooth variation of the electron concentration at the trapping levels.

The concentration m_0 of electrons on the trapping levels in the region of a static phase boundary (Fig. 4) is given by Eq. (4), and the variation of the order parameter $\eta(x)$ in the region of the phase boundary with m_0 constant is, as before, determined by Eq. (3).

The order parameter in the slow-motion region is, as before, a guiding variable. Its behavior in space follows the behavior of the concentration of electrons on the trapping levels at a given point: $\eta = \eta(m)$ (Eq. (5)) for $m > m_0$ and $\eta = 0$ for $m < m_0$.

The electron concentration at trapping levels, $m(x)$, in the slow-motion region can be found by solving the diffusion equations (2). The boundary conditions for these equations are fixed by the following considerations. The parameters m_b of the system at the center of a paraelectric-phase domain, $\eta_b = 0$, are determined by the conditions at the boundaries of the sample or by the structure at the defects. The parameters at the center of a ferroelectric-phase domain, η_c and m_c, can be obtained from the condition that the distri-

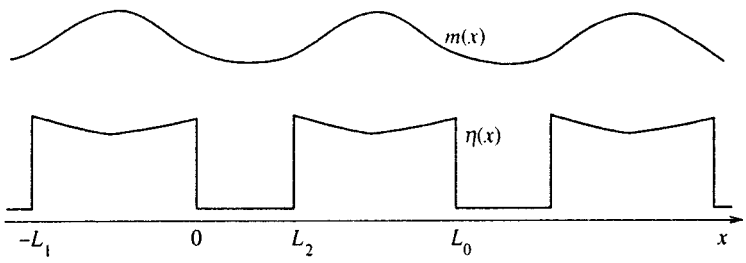

FIG. 4. Distribution of the order parameter η and the concentration m of the electrons on the trapping levels in a periodic heterophase structure with phase alteration.

bution of the concentration of electrons on the trapping levels at the phase boundary be smooth:

$$\int_{m_b}^{m_0} Q(\eta=0,m)\ dm+\int_{m_c}^{m_0} Q(\eta(m),m)\ dm=0, \qquad (13)$$

where $\eta(m)$ (Eq. (4)) corresponds to the order parameter at the given point. The period of the structure, L_0 ($L_0=L_1+L_2$; see Fig. 4), is given by

$$L_1=L\int_{m_c}^{m_0} G(m_c,m,\eta)dm,$$

$$L_2=L\int_{m_0}^{m_b} G(m_b,m,0)dm, \qquad (14)$$

$$G(m_1,m_2,\eta)=\left(\frac{A(0)}{2}\right)^{1/2}\left[\int_{m_1}^{m_2} Q(\eta(m),m)\ dm\right]^{-1/2},$$

where L is still the characteristic diffusion length for the electron subsystem.

The final expression for the distribution of the electron concentration at the trapping levels can be found by integrating the stationary part of Eq. (2) and allowing for the boundary conditions:

$$y=|x-L_2/2-nL_0|$$
$$=L\int_{m_b}^{m} G(m_b,m,0)\ dm, \quad 0<y<L_2/2,$$
$$y=|x-L_1/2-L_2-nL_0|$$
$$=L\int_{m_c}^{m} G(m_c,m,\eta)dm, \quad 0<y<L_1/2. \qquad (15)$$

Since Eq. (2) cannot be solved entirely in analytic form, the final result must be obtained by numerical integration. Figure 4 depicts the distributions of the order parameter η and the concentration m of the electrons on the trapping levels in a periodic heterophase structure.

6. EXISTENCE REGIONS

The region where an autosoliton can exist coincides with the region where the only state of a uniform system becomes metastable under variations in the order parameter in conditions of illumination. The region where a periodic structure can exist coincides with the region of instability of all states of a uniform system. All these regions in the space whose coordinates are the illumination intensity and the temperature are bounded by lines given by the following relationship for different values of the parameters ΔT and P:

$$J(T)=J(\Delta T,P)\equiv\frac{(\alpha-\Delta T)A(P)}{\Delta T-\alpha-aM}. \qquad (16)$$

Here the temperature dependence is hidden in the variable α (Eq. (1)) and the function $A(P)$, the latter determining the probability of a thermal jump in Eq. (2). The lines

$$J_1(T)=J\left(\frac{3\beta^2}{16\gamma},-\frac{3\beta}{4\gamma}\right), \quad J_2(T)=J\left(\frac{\beta^2}{4\gamma},-\frac{\beta}{2\gamma}\right),$$

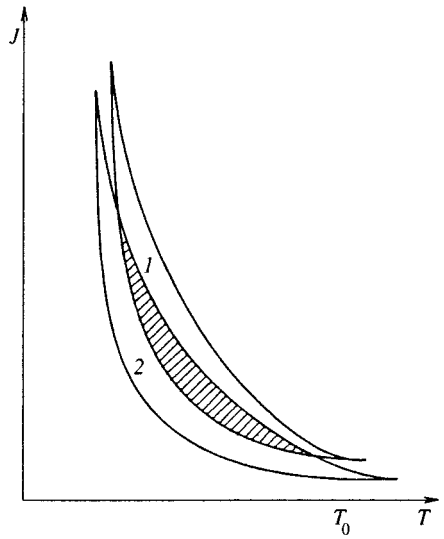

FIG. 5. Soliton existence regions in the space whose coordinates are the illumination intensity and the temperature: *1* and *2* are the regions of existence of stationary and traveling autosolitons (*1* for autosolitons with ferroelectric-phase domains in the metastable paraelectric-phase region, and *2* for autosolitons with paraelectric-phase domains in the metastable ferroelectric-phase region), and the hatched area corresponds to a periodic heterophase structure.

$$J_3(T)=J(0,0), \quad J_4(T)=J\left(\frac{3\beta^2}{16\gamma},0\right)$$

bound the region of existence of traveling and stationary paraelectric-phase autosolitons in the ferroelectric phase and ferroelectric-phase autosolitons in the paraelectric phase, respectively. The region of existence of a periodic structure in the space whose coordinates are the illumination intensity and the temperature are bounded by the lines $J_2(T)$ and $J_3(T)$. Figure 5 depicts the various regions of existence of these states. We see that nonuniform correlated states emerge in the system in a critical manner at high enough illumination levels. We also note that all these modes can emerge only when the electron and lattice subsystems have a strong effect, which is expressed by the following conditions imposed on the interaction parameters a and $\tilde{a}$:

$$aM(\alpha'\tilde{a}-\beta)^2>4\beta T\tilde{a}(\alpha')^2 \qquad (17)$$

for the emergence of autosolitons, and

$$a\tilde{a}M\gg\frac{\Gamma\beta^3 T}{\gamma A(0)} \qquad (18)$$

for the emergence of periodic structures.

Figure 6 depicts the variation of the autosoliton size L_0 with temperature in terms of the parameter α under constant external illumination ($J=0.6$).

7. DISCUSSION

So far no phenomena have been observed in ferroelectric semiconductors that could be associated with autosoliton states. Apparently, no special research in observing such states has ever been done, since the present paper is the first study in this area of research and establishes the importance and the subject of such investigations.

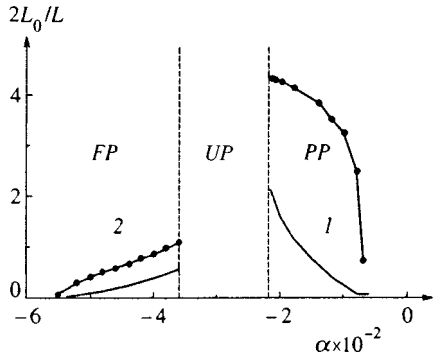

FIG. 6. The dependence of the soliton size on the parameter α: *1* corresponds to an autosoliton with a ferroelectric-phase domain in the metastable paraelectric phase, and *2* to an autosoliton with a paraelectric-phase domain in the metastable ferroelectric phase. The lines with dots correspond to traveling autosolitons, and those without dots to stationary autosolitons. *FP* stands for the ferroelectric phase, *PP* for the paraelectric phase, and *UP* for the unstable phase. Here ($\beta=-7\times10^4$, $\gamma=3\times10^{10}$, $a=10^{-19}$, $\tilde{a}=4\times10^8$, $M=10^{18}$, $T=300$ K, and $J=0.6$).

In contrast to autosolitons, the mechanism of formation of a periodic domain structure may be analogous to the mechanism of formation of a similar structure observed by Grekov *et al.*[12] in the SbSI photoferroelectric material. Another mechanism of this phenomenon was suggested by Larkin and Khmel'nitskiĭ.[11] They related this effect to a difference in conduction electron concentrations near the phase boundary. In contrast to the results of Ref. 11, in conditions of high illumination intensities the conduction electron concentration is uniform over the sample and is proportional to the irradiation intensity. In addition, since the publication of Ref. 11, new experimental and theoretical data have appeared, and it now obvious that phase transitions in ferroelectric semiconductors are affected mainly by the electrons in the traps rather than by conduction electrons. But even so, it remains a fact that it was Larkin and Khmel'nitskiĭ who predicted that the size of the structure is related to the dynamics of the electron subsystem in space.

The mechanism of formation of a periodic heterophase structure suggested in the present paper has another distinctive feature, in addition to the obvious fact that electrons on the trapping levels rather than conduction electrons play an important role in the formation of such a structure. In the ferroelectric-phase region the electron concentration at traps is higher than it is in the paraelectric phase, a result differing from that of Ref. 11. This is certainly a nontrivial fact, since as before, an increase in electron concentration shifts the transition temperature toward lower temperatures, and in the uniform case, as a rule, the paraelectric phase corresponds to a higher concentration of electrons on the trapping levels, and the ferroelectric phase to a lower concentration. It is therefore obvious that the suggested mechanism differs drastically from the one suggested in Ref. 11. The emergent behavior is related to the dynamics of the ferroelectric semiconductors as a whole and is an example of synergistic behavior.

In a real system both aperiodic domain structures and separate domains can emerge. The mechanism of formation of a periodic heterophase structure in real systems depends on the mechanism of autosoliton realization. It may be related to a local breakdown at the border of an autosoliton, where there is instability under small but finite fluctuations of the order parameter.[17]

Thus, we have studied the formation of various heterophase structures near a photostimulated phase transition. We found that this is caused by the fact that a uniform state becomes totally unstable. The related phenomena are linked to the self-consistent emergence of new-phase regions and to the fact that electron concentration at trapping levels is nonuniform.

The author would like to express his gratitude to A. P. Levanyuk, B. Z. Malkin, G. B. Teĭtel'baum, and V. M. Fridkin for useful discussion of the results. The present work was supported financially by the International Science Foundation and the Russian Government (Grant No. Rh 9300) and the Russian Fund for Fundamental Research (Grant No. 95-02-0382a).

[1] H. Haken, *Synergetics: An Introduction*, Springer, Berlin (1978).
[2] W. Ebeling, *Structurbildung bei irreversiblen Prozessen*, Teubner, Leipzig (1976).
[3] É. V. Bursian, A. B. Vall, and N. N. Trunov, Fiz. Tverd. Tela (Leningrad) **19**, 1890 (1977) [Sov. Phys. Solid State **19**, 1108 (1977)].
[4] S. V. Baryshnikov, É. V. Bursian, Ya. G. Girshberg, V. A. Lyakhovitskaya, and V. V. Maslov, Fiz. Tverd. Tela (Leningrad) **26**, 1872 (1984) [Sov. Phys. Solid State **26**, 1135 (1984)].
[5] E. V. Burcian, Ya. G. Girshberg, and N. N. Trunov, Ferroelectrics **36**, 305 (1981).
[6] A. L. Korzhenevskiĭ, Izv. Akad. Nauk SSSR, Ser. Fiz. **48**, 1069 (1984).
[7] A. I. Olemskoĭ, Fiz. Tverd. Tela (Leningrad) **25**, 2591 (1983) [Sov. Phys. Solid State **25**, 1490 (1983)].
[8] M. A. Krivoglaz and A. I. Karasevskiĭ, Zh. Éksp. Teor. Fiz. **69**, 297 (1975) [Sov. Phys. JETP **42**, 153 (1975)].
[9] A. Gordon, Phys. Lett. A **99**, 329 (1983).
[10] A. Gordon, Physica B **138**, 239 (1983).
[11] A. I. Larkin and D. E. Khmel'nitskiĭ, Zh. Éksp. Teor. Fiz. **55**, 2345 (1968) [Sov. Phys. JETP **28**, 1245 (1969)].
[12] A. I. Grekov, V. A. Lyakhovitskaya, A. I. Rodin, and V. M. Fridkin, Dokl. Akad. Nauk SSSR **169**, 810 (1966) [Sov. Phys. Dokl. **11**, 692 (1966)].
[13] V. M. Fridkin, *Photoferroelectrics*, Springer, Berlin (1979).
[14] V. M. Fridkin, M. I. Gorelov, A. A. Grekov, V. A. Lyakhovitskaya, and A. I. Rodin, JETP Lett. **4**, 310 (1966).
[15] I. M. Shmyt'ko, V. Sh. Shekhtman, V. I. Ivanov, and S. S. Khasanov, JETP Lett. **29**, 386 (1979).
[16] P. F. Mamin and G. B. Teĭtel'baum, JETP Lett. **44**, 420 (1986).
[17] B. S. Kerner and V. V. Osipov, Zh. Eksp. Teor. Fiz. **83**, 2201 (1982) [Sov. Phys. JETP **56**, 1275 (1982)].
[18] R. F. Mamin, JETP Lett. **60**, 52 (1994).

Translated by Eugene Yankovsky

Orbital and spin effects in the low-temperature behavior of the magnetoresistance of doped CdTe crystals

N. V. Agrinskaya and V. I. Kozub

A. F. Ioffe Physicotechnical Institute, Russian Academy of Sciences, 194021 St. Petersburg, Russia

R. Rentzsch

Freie Universität Berlin, Berlin, Germany

M. J. Lea and P. Fozooni

Royal Holloway, University of London, Egham, Surrey TW20 0EX, England
(Submitted 13 August 1996; resubmitted 10 November 1996)
Zh. Éksp. Teor. Fiz. **111**, 1477–1493 (April 1997)

An observation of the suppression of negative magnetoresistance in samples of doped CdTe that are far from the metal–insulator transition as the temperature is lowered in the temperature range 3–0.4 K was previously reported [N. V. Agrinskaya, V. I. Kozub, and D. V. Shamshur, JETP **80**, 1142 (1995)]. The results of an investigation of samples that are closer to the transition in the low-temperature region below 36 mK are presented. It is discovered that the samples investigated (which do not exhibit the suppression of negative magnetoresistance at comparatively high temperatures) display this effect at low temperatures and that, as previously, the suppression of the negative magnetoresistance correlates with the transition to conduction via Coulomb-gap states. A plateau-like magnetoresistance feature is displayed at low temperatures for the sample that is closest to the metal–insulator transition. The results obtained are analyzed within existing theoretical models that take into account the role of both the orbital and spin degrees of freedom. In particular, the low-temperature feature indicated is interpreted as a manifestation of positive magnetoresistance caused by spin effects. Nevertheless, it is shown within a detailed analysis supplemented by numerical calculations that the observed suppression of the negative magnetoresistance cannot be attributed only to the appearance of spin positive magnetoresistance. Moreover, the possibility of observing spin positive magnetoresistance is determined to a certain extent specifically by the suppression of the negative magnetoresistance competing with it.
[S1063-7761(97)02604-8]

1. INTRODUCTION

As has recently been shown in numerous theoretical and experimental studies, the picture of magnetoresistance in a hopping conduction regime is fairly complicated and diverse and depends on numerous physical factors. The most obvious among them is the ''contraction'' of the wave functions of the centers in a magnetic field, which leads to positive magnetoresistance. As was first shown by Shklovskiĭ,[1,2] this effect is modified significantly when scattering centers are present. Another important factor is determined completely by the scattering centers and is caused by the interference of various tunneling ''trajectories,'' which include subbarrier scattering events. Since the logarithmic character of the averaging over the different configurations stresses the configurations with ''destructive'' interference, suppression of the interference contribution by a magnetic field leads to negative magnetoresistance. These two factors have a very general character, and the competition between them determines the observed picture of the magnetoresistance.

The special mechanisms include, in particular, the narrowing of the impurity band in a magnetic field predicted by Raikh,[3] which also leads to negative magnetoresistance and has a purely orbital character.

Finally, the possible role of spin effects has been discussed for a long time. In particular, under certain conditions they can influence the interference contribution.[4] The corresponding mechanism operates when scattering occurs on filled centers (which, because of the statistics, should make the dominant contribution[2]). When the spin of an electron in a scattering center and the spin of an electron participating in ''direct'' (without scattering) hopping are antiparallel, the final spin states of the systems for the ''direct'' channel and the channel with scattering are different, and interference is prevented. Accordingly, a sufficiently strong magnetic field aligns the spins and thereby eliminates the ''antiparallel'' configurations. Since, according to the foregoing statements, the latter do not undergo destructive interference (and, thus, the hopping probability is higher), this can lead to positive magnetoresistance.

We note that the observation of features associated with ''spin alignment'' was reported in a recently published paper.[5] The competition of the ''spin'' factor under consideration with the suppression of interference by a magnetic field can significantly complicate the picture observed.[4] We note that the presence of ''built-in fields'' (due to exchange interaction), which fix the orientations of the spins at lattice sites, can suppress this mechanism.[2]

Another mechanism that leads to a direct contribution of spin effects to the magnetoresistance operates, if doubly filled sites take part in the conduction along with the singly filled and empty sites (which is possible if the Hubbard energy is lower than the width of the filled part of the impurity band).[6] In fact, the Zeeman splitting of the spin levels in a magnetic field leads to suppression of the hopping involving doubly filled sites in sufficiently strong fields, and thus to a decrease in the effective density of states, i.e., to positive magnetoresistance.

In addition, the type of magnetoresistance can differ under different mechanisms of hopping conduction: hopping between neighboring sites, variable-range hopping under the conditions of a constant density of states at the Fermi level, and similar hopping between Coulomb-gap states.

The question of the temperature dependence of the magnetoresistance is an interesting problem that has not been investigated in sufficient detail. It should be noted that various experimental studies have demonstrated behavior differences, and there has not been a systematic analysis of the available data.

In our recent studies,[7,8] we investigated the temperature dependence of the magnetoresistance in doped CdTe crystals in the temperature range 0.4–4 K. It was shown that samples which are sufficiently far from the metal–insulator transition on the insulator side exhibit crossover from conduction corresponding to Mott's law to conduction via Coulomb-gap states (the Éfros–Shklovskiĭ law) when the temperature is lowered. The transition to Éfros–Shklovskiĭ conduction is accompanied by suppression of the negative magnetoresistance in weak magnetic fields as the temperature is lowered.

In the theoretical papers by Schirmacher[9] and Raikh *et al.*[10] a significant increase in the negative magnetoresistance with decreasing temperature according to a $T^{-7/8}$ law was predicted within perturbation theory for hopping in the presence of scattering centers in the Mott regime in the ''weak-scattering'' situation[10] (see also Ref. 9). Extrapolation of the corresponding estimates to the Coulomb-gap regime would give a $T^{-7/4}$ dependence. An increase in the magnetoresistance in the ''strong-scattering'' situation was also predicted in Ref. 10, but it obeyed a weaker $T^{-3/8}$ law in the Mott regime, while extrapolation of the corresponding results to the Coulomb-gap regime gave $T^{-3/4}$ behavior.

As can easily be seen, a combination of negative magnetoresistance with quadratic positive magnetoresistance (which is caused by contraction of the wave functions) for the laws indicated always leads to enhancement of the contribution of the negative magnetoresistance as the temperature is lowered, but the value of the resistance at the minimum does not decrease in the strong-scattering regime. These predictions are not consistent with our experimental data.

In Ref. 8, we proposed a simple semiphenomenological model (based on the original model arguments of Shklovskiĭ and Spivak), which presumes a dominant role for the scatterers with energies close to the energies of the sites responsible for conduction. The decrease in the number of such sites upon passage into the Coulomb-gap region then actually does lead to suppression of the negative magnetoresistance as the temperature decreases, in agreement with experiment. We note that the effect observed is thus interpreted as a feature of the mechanism of interference negative magnetoresistance.

On the other hand, the positive magnetoresistance associated with spin effects[4,6] increases with decreasing temperature, and, as was demonstrated by the numerical calculations in Ref. 11, can also lead to the suppression of interference negative magnetoresistance. This raises the question, in particular, of the possibility of experimentally distinguishing between the suppression of negative magnetoresistance caused by a combination of the latter with spin positive magnetoresistance, which is not part of the interference contribution, and the possible effects associated with the interference contribution proper.

We note that the temperature range studied in Ref. 8 was not broad enough for a detailed investigation of the suppression of negative magnetoresistance. In particular, crossover on the temperature dependence of the resistance and the suppression of negative magnetoresistance were observed only for samples far from the metal–insulator transition, while samples close to that transition did not display such behavior and remained in the Mott conduction regime. This naturally raised the question of the behavior of such samples at lower temperatures, where their crossover point should be located, according to calculations.

In the present work we performed detailed investigations of variable-range hopping (VRH) conduction for doped CdTe crystals in the ultralow temperature range 0.03–1 K. The fairly broad temperature range of the investigations permitted the observation of various VRH conduction regimes (passage from Mott's law to conduction via Coulomb-gap states), even for samples that are close to the metal–insulator transition. The suppression of negative magnetoresistance was observed for all the samples as the temperature was lowered and upon passage to the Coulomb-gap conduction regime, attesting to the fundamental character of the effect observed. Features characteristic of the mechanisms described in Refs. 4 and 6 (a segment with saturation of the magnetoresistance) were discovered for the sample closest to the metal–insulator transition. At the same time, the magnetoresistance for the other samples (with a larger degree of compensation) did not exhibit distinct features. A comparison of the data obtained with the theory reveals that the suppression of negative magnetoresistance cannot be explained on the basis of spin mechanisms.

2. EXPERIMENT

The CdTe crystals were grown from a melt by horizontally directed crystallization and were doped in the melt by a fine source of Cl donor impurity atoms (the ionization energy of an isolated Cl donor atom at a Te site is close to the hydrogenic value and amounts to 14 meV). The density of electrically active donors was evaluated from Hall measurements on uncompensated samples: $N_d = 5\times10^{17}$ cm^{-3}. This value exceeds the critical value for the metal–insulator transition in CdTe, $n_{c0} = 1.5\times10^{17}$ cm^{-3}, and the uncompensated samples consequently had metallic conductivity. The samples investigated were compensated (the degree of com-

TABLE I. Principal parameters of the samples: $n_{300\,K}$ — the electron density at 300 K obtained from Hall measurements; T_0 and T_1 — parameters obtained from the temperature dependences of the resistance [see Eqs. (1) and (2)]; T_c – the experimentally determined crossover temperature; a — the localization radius.

No.	$n_{300\,K}$, cm^{-3}	T_0, K	T_1, K	T_c, K	a, Å
1	9×10^{16}	2×10^4	43	0.4	90
2	1.2×10^{17}	205	3.5	0.15	180

pensation was $\sim$0.2–0.5) and had an electron density close to 10^{17} cm^{-3} (the exact values of the electron density at room temperature are given in Table I). In the case of compensation the critical density is somewhat greater than n_{c0}, and the compensated samples with an electron density close to 10^{17} cm^{-3} are insulators.

The measurements at temperatures below 1 K were performed in the cryogenic system in the Freie Universität Berlin by R. Rentzsch and in the University of London by P. Fozooni. The data on the low-temperature behavior of the CdTe crystals were partially published in Ref. 12. We first focus on the results of the investigation of the temperature dependence of the resistance in the absence of a magnetic field. As we know, VRH conduction can be described by different laws, depending on the temperature and the parameters of the material. At high temperatures, at which the VRH density of states $g(\varepsilon_F)=$const in the characteristic energy range, Mott's law holds:

$$\rho(T)=\rho_0(T)\exp(T_0/T)^{1/4}, \quad T_0=\beta_0/g(\varepsilon_f)a^3. \tag{1}$$

At lower temperatures, at which the characteristic hopping takes place within the Coulomb gap, $g(\varepsilon)=g_0\varepsilon^2$, and

$$\rho(T)=\rho_1(T)\exp(T_1/T)^{1/2}, \quad T_1=\beta_1 e^2/\kappa a. \tag{2}$$

Here $\beta_0=21$ and $\beta_1=2.8$ are numerical coefficients, and κ is the dielectric constant. We also note the different temperature dependences of the pre-exponential factors for these two laws: $\rho_0(T)\propto T^{1/4}$, $\rho_1(T)\propto T^{1/2}$. As was shown in Refs. 7, 13, and 14, the influence of the pre-exponential factors is very significant for the samples that are near the metal–insulator transition. On the basis of data obtained for several CdTe samples in the temperature range 0.6–3 K, we previously

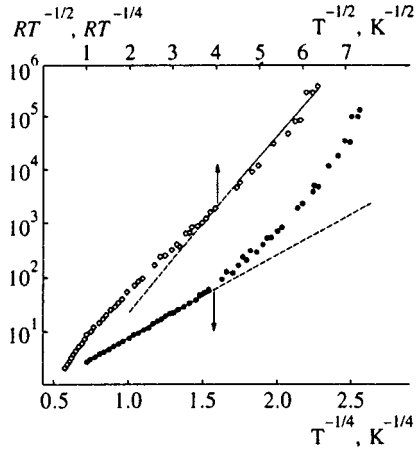

FIG. 1. Temperature dependence of the resistance of sample No. 2 plotted in the coordinates corresponding to Eqs. (1) and (2).

reported[7] the observation of crossover from Mott conduction to Éfros–Shklovskiĭ conduction for CdTe samples that were fairly far from the metal–insulator transition. At the same time, a less compensated sample exhibited pure Mott conduction in the temperature range indicated. Therefore, an investigation of that sample at lower temperatures would be of special interest.

Figure 1 presents the temperature dependence of the resistance of that sample (which is close to the metal-insulator transition) in $\log(RT^{-1/4})$ versus $T^{-1/4}$ and $\log(RT^{-1/2})$ versus $T^{-1/2}$ coordinates. As can be seen from Fig. 1, the conductivity of this sample is described well by law (1) in the high-temperature region $T>0.2$ K, and appreciable deviations toward higher resistances are observed as the temperature is lowered. At low temperatures the log $R(T)$ curves correspond to law (2), and as the temperature rises above $T_c=0.15$ K, deviations from this law are observed, at first toward higher resistances and then toward lower resistances. Thus, this sample, which is close to the transition, also exhibits crossover, but at significantly lower temperatures than in the more compensated samples.[7]

Let us examine the behavior of the magnetoresistance as a function of the temperature and the magnetic field. Figures

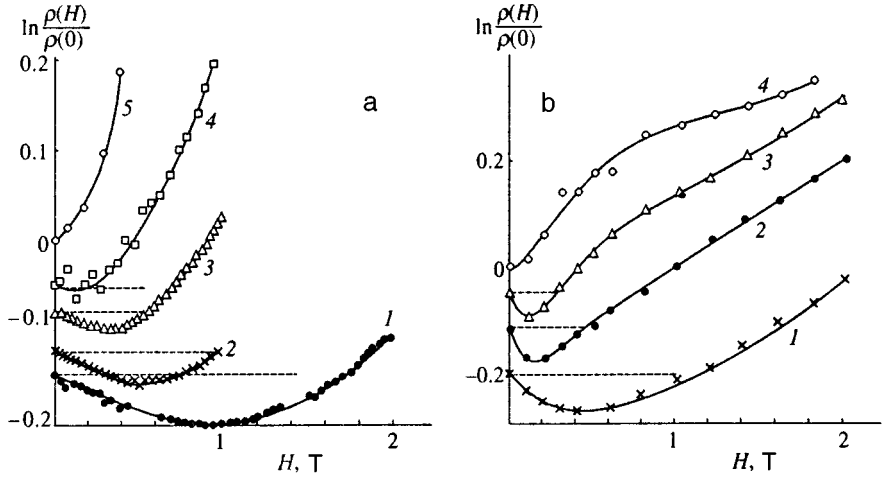

FIG. 2. Magnetoresistance of the two samples of CdTe(Cl) whose parameters are indicated in Table I for various temperatures: a — sample No. 1, b — sample No. 2. The numbers of the curves correspond to the following temperatures: a) *1* — 900, *2* — 600, *3* — 400, *4* — 300, *5* — 140 mK; b) *1* — 300, *2* — 150, *3* — 90, *4* — 60 mK.

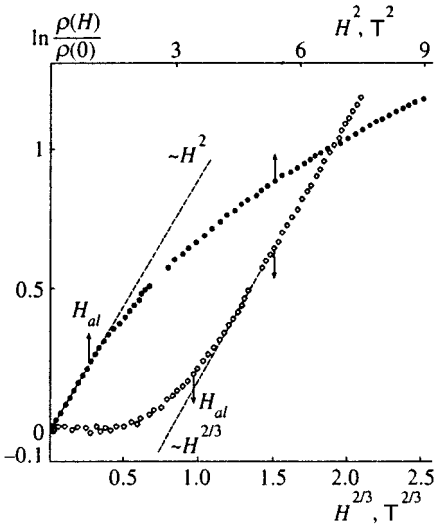

FIG. 3. Dependence of the magnetoresistance of sample No. 1 on H^2 and on $H^{2/3}$, $T=300$ mK. The arrows point to the values of H_{al} calculated for this temperature.

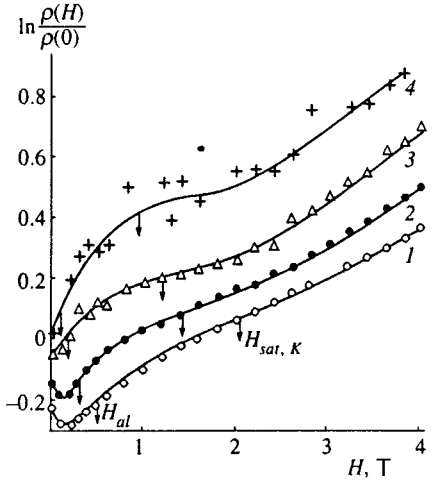

FIG. 4. Magnetoresistance of sample No. 2 at high magnetic fields and low temperatures: *1* — 36, *2* — 60, *3* — 90, *4* — 150 mK. The arrows point to the values of H_{al} and $H_{sat,K}$ calculated for the corresponding temperatures.

2a and 2b present the dependence of the resistance on the magnetic field H for samples Nos. 1 and 2 at various temperatures. At high temperatures (~ 1 K) the magnetoresistance is negative at fields from 0 to 2 T, and the maximum value of the negative magnetoresistance (at the minimum determined by the competition with orbital positive magnetoresistance) does not exceed 10%. At low temperatures the negative magnetoresistance at the minimum ($H_{\min}$) at first increases slightly with decreasing temperature, reaching a value of 10%, and then after a certain temperature, it decreases with decreasing temperature. In addition, the value of $H_{\min}$ shifts toward smaller values of H as the temperature is lowered. At temperatures less than T_c this drop in the contribution of the negative magnetoresistance becomes especially sharp, and at $T \approx T_c/2$ the latter vanishes completely (the point where the negative magnetoresistance vanishes is determined by how accurately it is measured at small H).

At high temperatures a segment with a quadratic dependence of $\rho(H)$ is observed in very weak magnetic fields and gives way to a linear law as H increases. However, at low temperatures linear behavior is observed for all the samples.

Let us now move on to the results for stronger magnetic fields. Figure 3 contains plots of $\rho(H)$ for sample No. 1 in $\ln(\rho(H)/\rho(0))$ versus H^2 coordinates and $\ln(\rho(H)/\rho(0))$ versus $H^{2/3}$ coordinates for $T=300$ mK. A well-defined region of quadratic positive magnetoresistance is observed. The analysis for other temperatures shows, however, that this region is displaced toward weaker magnetic fields as the temperature is lowered, while positive magnetoresistance, which depends weakly on H ($\propto H^{2/3}$), is observed at stronger fields (Fig. 3). This dependence is typical of the anomalous positive magnetoresistance associated with the contraction of the wave functions of the sites by the magnetic field in the presence of scatterers. As was shown in Ref. 2, the small distance between the scatterers in comparison with R ensures significant lowering of the ''magnetic barrier'' and leads to a significant decrease in the magnitude of the positive magnetoresistance (in comparison with the case of no scatterers). The dependence of the magnitude of this magnetoresistance on H gradually weakens as the magnetic field increases with passage from an H^2 law to an $H^{2/3}$ law.[1]

Sample No. 2, which is very close to the transition, displays a well-defined point of inflection between convex and concave portions of the curve at low temperatures (Fig. 4). Its position shifts toward weaker fields as the temperature is lowered, while the portion of the curve in its vicinity flattens, and at $T=36$ mK a well-defined region of saturation of the positive magnetoresistance is displayed at 0.3–2 T (we note that the magnitude of this magnetoresistance in the region of the feature increases with decreasing temperature and reaches 0.2 at $T=36$ mK). Below we shall interpret this feature as a manifestation of spin effects.[4,6]

3. DISCUSSION

Let us first analyze the temperature dependence of the quadratic positive magnetoresistance (which is associated with deformation of the wave functions of the sites in a magnetic field) that is observed for the samples under consideration at sufficiently low temperatures and not excessively strong magnetic fields. For VRH conduction at weak magnetic fields such that $\ln\rho(H)/\rho(0) \ll 1$ the dependence of the positive magnetoresistance on the magnetic field is described by the expression[2]

$$\ln \frac{\rho(H)}{\rho(0)} = \left(\frac{H}{B}\right)^2, \tag{3}$$

where

$$B^2 = \frac{\alpha c^2 \hbar^2}{r^3 a e^2} \tag{4}$$

is a parameter that depends on the temperature, the localization radius, and the conduction mechanism (law (1) or (2)):

$$B^2 = \frac{c^2 \hbar^2}{C_0 e^2 a^4} \left(\frac{T}{T_0}\right)^{3/4}, \tag{5}$$

$$B^2=\frac{c^2\hbar^2}{C_1e^2a^4}\left(\frac{T}{T_1}\right)^{3/2}. \tag{6}$$

Equation (5) corresponds to Mott's law, $C_0=0.0025$, and $\alpha\simeq 400$. Equation (6) corresponds to conduction via Coulomb-gap states, and two values are given for C_1: 0.0015 ($\alpha\sim 700$) and 0.0035 ($\alpha\sim 300$). We note that the value of B is of the order of the value of the saturation field for interference negative magnetoresistance $H_{\rm sat}\sim\Phi_0/r^{3/2}a^{1/2}$: $B=(\sqrt{\alpha}/\pi)H_{\rm sat}$.

The localization radii for the two hopping conduction regions (the Mott and Éfros–Shklovskiĭ regions) can be determined from the coefficient of the quadratic contribution. These values are presented in Table I. The negative magnetoresistance observed at weak magnetic fields, which is linear with respect to the magnetic field and small in absolute value, is caused by the interference of the contributions to the amplitude of the hopping probability from different trajectories, one of which includes a scatterer with a negative scattering amplitude μ (the energy of the corresponding center must be lower than the energy of the sites between which hopping occurs). Such interference obviously has a destructive character, the contribution of the corresponding configurations being enhanced owing to the logarithmic character of the configurational averaging. The suppression of such interference by a magnetic field also leads to negative magnetoresistance.[2] Within a semiqualitative phenomenological description the contribution of the magnetic field to the hopping probability can be expressed in the form

$$L(H,R)=N\mu^{5/2}a^{3/2}RH, \tag{7}$$

where R is the hopping distance and N is the number of scatterers. A more rigorous analysis, which takes into account, in particular, the energy dependences of the scattering processes, leads to a more complicated expression for the coefficient in the corresponding linear law.[8–10]

To compare the linear negative magnetoresistance with the positive magnetoresistance it is convenient to express the negative magnetoresistance in the form

$$\ln\frac{\rho(H)}{\rho(0)}=k\frac{H}{B}, \tag{8}$$

where B is defined by (5) and (6), while k depends on the details of the interference mechanism. In particular, in the strong-scattering limit the estimates in Ref. 10 give the value $k\simeq 1.3$ for k.

Summing the two contributions, we have

$$\ln\frac{\rho(H)}{\rho(0)}=-k\frac{H}{B}+\left(\frac{H}{B}\right)^2. \tag{9}$$

Hence we at once obtain

$$H_{\min}=\frac{k}{2}B,\quad \ln\frac{\rho(H)}{\rho(0)}\bigg|_{\min}=-\frac{k^2}{4}. \tag{10}$$

We note that, strictly speaking, the estimated data are valid only for $k<1$, since in the opposite case $H_{\min}$ is of the order of $H_{\rm sat}$, and thus the linear extrapolation for the negative magnetoresistance breaks down when $H\sim H_{\min}$.[10]

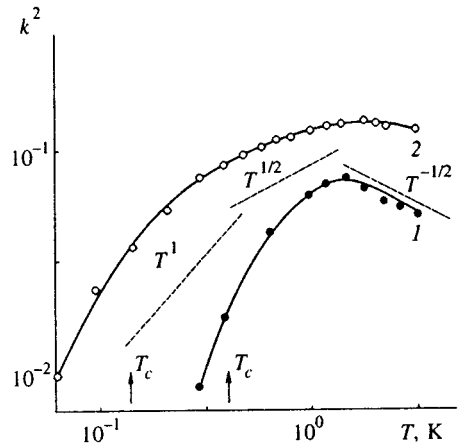

FIG. 5. Temperature dependence of the parameter k^2 determined from the magnetoresistance curves: $\ln\rho(H)/\rho(0)|_{\min}\sim k^2$; the dashed lines represent the power-law dependence of k^2 on the temperature predicted in Ref. 8 for various limiting cases; the arrows indicate the crossover temperatures.

Figure 5 presents the temperature dependence of $\ln(\rho(H)/\rho(0))|_{\min}\propto k^2$ for the two samples. Appreciable decreases in these parameters with decreasing temperature, which are especially sharp at the low temperatures $T<T_c$, are seen. We emphasize that the temperatures at which the sharp decrease in the negative magnetoresistance is observed are different for the two samples and correlate with the crossover temperature.

As we have already pointed out, the analysis of temperature dependence of interference negative magnetoresistance performed in Refs. 9 and 10 predicts a significant increase in the negative magnetoresistance with decreasing temperature in the Mott regime: according to a $T^{-7/8}$ law (weak scattering) or a $T^{-3/8}$ law (strong scattering[10]). In the regime of conduction via Coulomb-gap states the strong-scattering limit gives a $T^{-3/4}$ law (as follows from extrapolation of the results in Ref. 10 to the Coulomb-gap regime). In addition, in the strong-scattering limit the value of the magnetoresistance at the minimum does not vary with the temperature, and in the case of weak scattering it increases significantly with decreasing temperature. These results are not consistent with the experimental data presented here.

In Ref. 8, utilizing the original arguments of Shklovskiĭ and Spivak (which relate the interference component of the tunneling probability $L(H,R)$ to the scattering parameters, specifically the scattering amplitude μ and the density of scatterers N), we proposed a model, which takes into account the energy dependence of the scattering amplitude (its increase as the energy of the scatterers ε approaches the energies of the sites responsible for VRH). Under the assumption that the scattering potentials are small, it leads to the conclusion that the scatterers with energies in the vicinity of the characteristic energy band for VRH conduction play the decisive role (in fact, the proximity of the scattering on such centers to resonance permits compensation for the small values of the scattering potentials). Thus, a decrease in the number of such centers in the regime of conduction in the Coulomb gap leads to suppression of the negative magne-

toresistance upon passage to the corresponding VRH regime within such a model.

Figure 5 compares the corresponding theoretical predictions with the experimental data. Qualitative agreement is observed, although the observed decrease in the negative magnetoresistance with decreasing temperature is faster than the decrease predicted by modeling considerations.[8] We note, however, that both the semiphenomenological character of the model[8] and the model assumptions employed (particularly the assumption that the scattering potential V_0 is small) introduce a certain element of uncertainty and require a more detailed analysis of the situation. This is due particularly to the presence of spin effects, which lead to positive magnetoresistance and can thereby cause suppression of the negative magnetoresistance.

We therefore discuss the possible role of the spin effects. We begin with the effect associated with the influence of the spin degrees of freedom on the interference contribution proper that was considered in Ref. 4. In the three-site approximation,[8–10] under which there is only one scattering center, this effect reduces to suppression of the interference contribution for the configurations in which the electron spin in the (filled) scattering center is antiparallel to the spin of the electron participating in the hopping. Accordingly, alignment of the spins in the magnetic fields $H > H_{al}$, where

$$H_{al} \simeq T/g\mu_0,$$

μ_0 is the Bohr magnetic, and g is the g factor, leads to restoration of the corresponding contribution and thus to positive magnetoresistance. Inasmuch as the weight of the antiparallel configurations amounts to 50%, the positive magnetoresistance amounts to 50% of the total interference contribution ρ_{sat} for spinless electrons. Utilizing the expression for the filling functions of the spin sublevels in a magnetic field, we can represent the contribution to the magnetoresistance under discussion in the form

$$\frac{\rho_{sat}}{2}\tanh\left(\frac{H}{2H_{al}}\right).$$

Since $g=0.45$ for CdTe, the values of H_{al} for temperatures 0.036, 0.1, 0.3, and 1 K are estimated as ~ 0.1, ~ 0.3, ~ 1, and ~ 3 T, respectively. We note that, as can easily be seen, at an assigned temperature the position of this feature does not depend on the proximity of the sample to the metal–insulator transition [which is characterized, in particular, by ξ, where $\xi=(T_0/T)^{1/4}$ (the Mott regime) or $\xi=(T_1/T)^{1/2}$ (the Coulomb-gap regime)].

This spin mechanism of positive magnetoresistance is clearly displayed against a background of the (linear) orbital negative magnetoresistance caused by suppression of the interference contribution discussed above, as well as the (quadratic in weak fields) orbital positive magnetoresistance associated with contraction of the wave functions. In addition, the two former contributions are of an interference nature and have an upper bound ρ_{sat} (which does not exceed 40% of $\rho(H=0)$), while the orbital positive magnetoresistance undergoes an exponential increase at $H > B$. Furthermore, the orbital positive magnetoresistance predominates already at $H > H_{min}$. Therefore, the behavior of the magnetoresistance should depend on the relation between H_{al} and the value of $H_{min} < H_{sat} \sim B$. In fact, when $H > H_{min}$ and especially when $H > B$, the rapidly increasing orbital positive magnetoresistance clearly prevents the appearance of features associated with a spin contribution. Accordingly, when $H_{al} \gg H_{min}$, the spin factor can be neglected. It is not difficult to see that this inequality is satisfied in the high-temperature limit in the Mott regime (where $B \propto T^{3/8}$, while $H_{al} \propto T$).

On the other hand, in the limit $T \to 0$ we should expect $H_{al} \ll (H_{min}, B)$ because of the stronger temperature dependence of H_{al}. When $H \ll H_{min}$, the orbital positive magnetoresistance can obviously be neglected, and everything is determined by the competition between the spin positive magnetoresistance and the orbital negative magnetoresistance. In such a case the presence of spin positive magnetoresistance can be manifested as a ''jump'' $\Delta\rho(H)$ toward more positive values at the weaker fields $H < H_{al}$ followed by a decrease caused by the interference orbital magnetoresistance. Also, because the maximum value of the latter is twice the maximum spin positive magnetoresistance, reversal of the sign of the magnetoresistance as H increases is possible. (We note, however, that as the numerical calculations show, the corresponding nonmonotonic dependence can be observed only if H_{al} is at least an order of magnitude smaller than H_{min}. When H_{al} is large, the orbital positive magnetoresistance is significant, and the spin positive magnetoresistance is manifested only as a tendency to form a plateau.)

However, as the estimates show, the difference between H_{al} and H_{min} for the experimental situation under discussion is not large enough to correspond to one of the limiting cases just described. Therefore, when the behavior of the magnetoresistance is analyzed, the predominant contribution cannot be isolated. In particular, the competition between the spin positive magnetoresistance and the orbital negative magnetoresistance can lead to the mutual suppression of both. To evaluate the extent to which this circumstance can have a bearing on the observed suppression of the negative magnetoresistance as the temperature is lowered, let us consider the sum of the two contributions indicated in the region $H < \min(H_{al}, H_{sat})$, where they are both described by a linear law. This sum equals

$$\frac{\rho_{sat}}{2}\frac{H}{H_{al}} - k\frac{H}{B}, \tag{11}$$

or, if we have in mind a combination with quadratic orbital positive magnetoresistance, we can represent it in the form

$$-\tilde{k}\frac{H}{B}, \quad \text{where } \tilde{k} = k\left(1 - \frac{\rho_{sat}B}{kT}\right). \tag{12}$$

Since $H_{al} \propto T^{-1}$ and $B \propto T^{-3/4}$ in the Coulomb-gap regime (where suppression of the negative magnetoresistance is observed experimentally), it can easily be seen that suppression of the negative magnetoresistance by the spin contribution can occur if $T^{1/4}k/\rho_{sat} \to 0$ as $T \to 0$. We recall that in accordance with the predictions in Refs. 9 and 10, under weak-scattering conditions k increases rapidly with decreasing T, and the suppression of the negative magnetoresistance in the corresponding situation could not be attributed to the presence of a spin contribution. In the strong-scattering limit

$k=$ const and $\rho_{\rm sat}=$ const were predicted in Ref. 10, and in such a situation, in principle, the role of the spin positive magnetoresistance is in fact enhanced as the temperature is lowered.

However, we note that the difference between the temperature dependences of the two contributions is small in this case; we are dealing with the combination

$$k\left(1-\left(\frac{T^*}{T}\right)^{1/4}\right),$$

where T^* is a constant. It is clear that an appreciable decrease in the negative magnetoresistance with decreasing temperature occurs only over the temperature range in which the two contributions are nearly compensated ($T^*<T<T_m=T^*(4/3)^4\sim 3T^*$) and that because of this compensation the maximum attainable coefficient of the linear negative magnetoresistance (at $T=T_m$) amounts to no more than 1/3 of the value in the absence of a spin contribution. When the quadratic dependence of $\ln(\rho(H)/\rho(0))|_{\rm min}$ on the coefficient k in such a situation is taken into account, the value of the magnetoresistance at the minimum could not exceed 2%, which contradicts the experiment. On the other hand, as follows from the estimates, T^* (and thus the temperature T_m at which the decrease in the negative magnetoresistance begins) depends very sensitively on the parameters of the material ($T_m \propto T_1^3/a^8$), and this dependence does not correlate at all with the corresponding dependence of the crossover temperature. This is also at variance with the experiment.

Thus, the combination of orbital negative magnetoresistance with spin positive magnetoresistance is not capable of explaining the observed effect. We supplemented the foregoing arguments with numerical calculations that take into account either spin positive and orbital negative magnetoresistance or orbital positive magnetoresistance (see the Appendix).

Therefore, in our opinion, the observed effects suggest that a decrease in temperature in the Coulomb-gap regime leads to a decrease in the negative magnetoresistance, i.e., the coefficient k (which follows from the modeling considerations in Ref. 8). As follows from (12), the decrease in k is also accompanied by ''decompensation'' of the temperature dependences of the spin positive and negative magnetoresistance, and thus leads to sharper suppression of the negative magnetoresistance. Therefore, even a slow decrease in k can lead to rapid suppression of the negative magnetoresistance. For example, under the assumption of the dependence $k \propto T^{1/4}$ (which in the model of Ref. 8 corresponds to the experimental situation under consideration), the combination of orbital negative and spin positive magnetoresistance takes the form

$$\tilde{k} \propto T^{1/4}\left(1-\left(\frac{T^*}{T}\right)^{1/2}\right).$$

We now turn to the spin positive magnetoresistance described in Ref. 6, which is associated with the participation of doubly filled sites in hopping conduction. We recall that it is saturated in magnetic fields $H>H_{{\rm sat},K}$, at which the Zeeman splitting exceeds the width of the effective VRH energy band:

$$H_{{\rm sat},K}\simeq \frac{T\xi}{g\mu_0}, \tag{13}$$

and $H_{{\rm sat},K}$ is $\sim\xi\gg 1$ times greater than the characteristic field $H_{\rm al}$ describing the spin mechanism.[4] The value of the corresponding contribution in the saturation region is

$$\ln(\rho_K/\rho_0)\sim\ln(\rho_{{\rm sat},K}/\rho_0)=\gamma\xi, \tag{14}$$

where γ characterizes the relative fraction of the doubly filled sites.

Thus, it is clear that the manifestations of this mechanism are qualitatively similar to the spin interference mechanism considered above.[4] The primary quantitative difference lies in the larger values of the saturation field $H_{{\rm sat},K}$ (in comparison with $H_{\rm al}$), so that it can be observed only in the low-temperature limit; otherwise, this contribution is masked by the rapidly increasing orbital positive magnetoresistance. Another difference is the absence of a direct relationship between the value of $\rho_{{\rm sat},K}$ and the interference contribution responsible for the negative magnetoresistance. We note that with consideration of what has been stated this mechanism of positive magnetoresistance is associated with a larger number of independent parameters than is the mechanism in Ref. 4 and, thus, with greater arbitrariness in a comparison with experiment.

Finally, we also note that some uncertainty in the evaluation of the role of the spin mechanisms is attributable to the role of the intersite spin correlations, which are capable of suppressing the spin contribution described in Ref. 2.

We now compare the arguments presented above with experiment. The value of $H_{\rm al}$ for sample No. 1 is of the order of $H_{\rm min}$ for all the temperatures investigated. Therefore, the absence of plateau-like features is not surprising. At the same time, in our opinion, the abrupt decrease in $\ln(\rho(H)/\rho(0))|_{\rm min}$ with decreasing temperature suggests a decrease in k against the background of spin positive magnetoresistance.

On the other hand, at low temperatures sample No. 2 exhibits a well-defined feature in the range of fields 0.4–1 T, which might be associated with a spin contribution. However, as noted above, the behavior of the negative magnetoresistance in this sample, too, cannot be explained without postulating a decrease in k with decreasing temperature.

It can be stated for $T=36$ mK, at which the feature is most pronounced, that the beginning of the plateau-like segment corresponds to fields of ~ 0.4 T, which exceed the accurately determined value of $H_{\rm al}$ (~ 0.1 T). At the same time, the ''jump'' in the positive magnetoresistance ($\delta\rho/\rho_0\sim 0.2$) is appreciably greater than the maximum value of the negative magnetoresistance at the minimum (~ 0.07). From our standpoint, both of these findings attest to manifestations of the contribution described in Ref. 6 along with the contribution described in 4. The possible emergence of this contribution is associated with the comparatively small values of $H_{{\rm sat},K}$ for this sample at low temperatures.

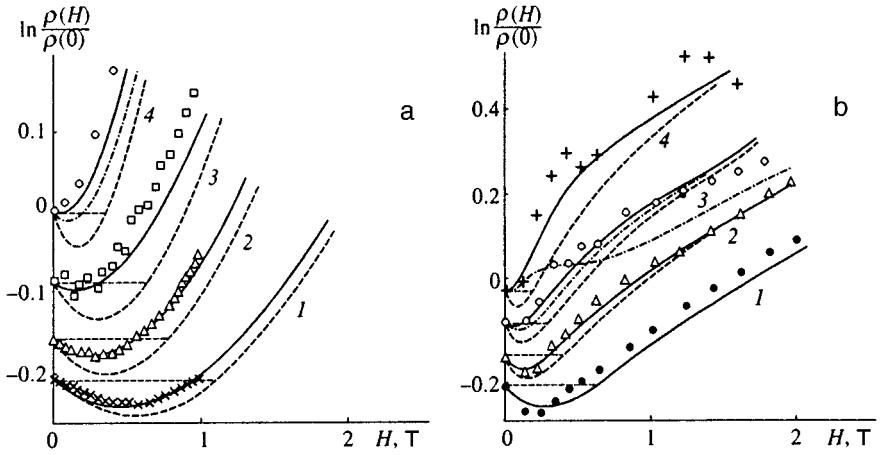

FIG. 6. Calculated dependences of the magnetoresistance for various temperatures constructed according to the procedure described in the Appendix for sample No. 1 (a) and for sample No. 2 (b); solid curves — calculation under the assumption that $k \propto T^{1/2}$, dashed curves — calculation under the assumption that $k = \text{const}$; a) *1* — 600, *2* — 400, *3* — 300, *4* — 150 mK, dot-dashed curve — calculation under the assumption that $k \propto T^{1/4}$; b) *1* — 150, *2* — 90, *3* — 60 mK, dot-dashed curve *4*— 36 mK, calculation under the assumption that the contribution described in Ref. 6 is neglected.

For a more convincing comparison of the theoretical arguments with experiment we performed some numerical calculations (whose details are presented in the Appendix). Figure 6 presents theoretical curves for the combination of interference negative, orbital positive, and spin positive magnetoresistance both with consideration of the decrease in k with decreasing temperature and for the strong-scattering situation. It is clear that the curves constructed under the assumption of $k = \text{const}$ do not in any way reproduce the observed behavior. At the same time, consideration of the decrease in k enables us to achieve fairly good agreement with experiment.

We note that better agreement is achieved for sample No. 2 when the spin contribution described in Ref. 6 is taken into account along with the spin contribution described in Ref. 2.

4. CONCLUSIONS

The investigation of low-temperature magnetoresistance performed in this work has disclosed the competition between various orbital and spin mechanisms.

In comparatively weak fields suppression of the negative magnetoresistance is observed as the temperature is lowered upon passage to hopping conduction via Coulomb-gap states.

Although magnetoresistance features which we associated with spin effects[2,5] were discovered for samples that are close to the metal–insulator transition, consideration of the latter does not allow us to quantitatively account for the observed suppression of the negative magnetoresistance. In addition, the very possibility of observing the indicated features results from the decrease in the negative magnetoresistance competing with them as the temperature is lowered.

At the same time, however, this decrease is capable of significantly enhancing spin positive magnetoresistance. As a result, even a comparatively weak seed decrease in the k, which characterizes the magnitude of the negative magnetoresistance, is capable of leading to reversal of the sign of the magnetoresistance even when the temperature is varied by only a factor of 2–3.

The arguments presented are confirmed by the numerical calculations.

Thus, these results, which were obtained for a broad range of temperatures, confirm the conclusion that we previously drew regarding the decrease in the interference contribution upon the transition to VRH conduction via Coulomb-gap states.

We thank D. V. Shamshurin for assisting with the low-temperature measurements.

This work was carried out with support from INTAS (Grant No. 93-1555), as well as a grant from the Russian Foundation for Fundamental Research (Project No. 95-02-04109-a).

APPENDIX A

To simulate the combination of spin positive, interference negative, and orbital positive magnetoresistance in sample No. 1, we used the following expression, which describes the sum of the corresponding interpolation formulas

$$\ln\frac{\rho(H)}{\rho_0} = \rho_{\text{sat}}\left(\left(\exp\left(\frac{g\mu H}{T}\right)+1\right)^{-1} - \frac{1}{2}\right) + \rho_{\text{sat}}\left(\exp\left(-\frac{k}{\rho_{\text{sat}}}\frac{H}{B}\right)-1\right) + \left(\left(\frac{H}{B}\right)^{-2} + c\left(\frac{H}{B}\right)^{-0.66}\right)^{-1}. \quad (15)$$

The value chosen for ρ_{sat} was 0.3 (which is smaller than the value for the strong-scattering limit). The value of the field B at $T = 300$ mK was set equal to 1.8 T. Although this expression takes into account the transition in the dependence of the orbital positive magnetoresistance from a quadratic dependence to a $H^{2/3}$ law in strong fields, the corresponding deviations for the sample under consideration are actually significant at fields significantly greater than $H_{\min}$. Since we were interested in the range of fields in the vicinity of $H_{\min}$, we used the value $c = 0.3$.

As for the temperature dependence $k(T)$, we discovered that the principal features of the observed behavior (complete suppression of the negative magnetoresistance at $T = 150$ mK from values of $\sim 10\%$ at $T = 600$ mK) are reproduced with $k \propto T^{1/4}$. However, the best agreement with ex-

periment is achieved for $k \propto T^{-1/2}$ and $k(T=300\ \text{mK})=0.7$. The calculated curves are presented in Fig. 6a along with experimental points for comparison. Clearly the agreement between them is at least satisfactory. In addition, for the suppression of the negative magnetoresistance proper (which can be observed in comparatively weak fields) the agreement can be regarded as good, while in stronger fields the temperature dependence of the slopes of the curves is somewhat stronger than predicted by the theory in Ref. 2.

For comparison, we also present the curves obtained for $k=\text{const}$. It is clear that these curves do not describe the observed behavior of the negative magnetoresistance at all. At the same time, for $T=150$ mK we also present the calculated curve corresponding to a $k \propto T^{1/4}$ dependence.

For sample No. 2, which is closer to the metal–insulator transition, in the field dependence of the orbital positive magnetoresistance we took into account the transition to a weaker dependence ($\propto H^{2/3}$), as well as the possible emergence of the spin contribution described in Ref. 6. For simplicity, we used the simplified interpolation equation proposed in Ref. 11 for this contribution. The sum of all the contributions discussed takes the form

$$\ln\frac{\rho(H)}{\rho_0}=A\rho_{\text{sat}}\left(\left(\exp\left(\frac{g\mu H}{T}\right)+1\right)^{-1}-\frac{1}{2}\right)+\rho_{\text{sat}}\left(\exp\left(-\frac{k}{\rho_{\text{sat}}}\frac{H}{B}\right)-1\right)+\left(\left(\frac{H}{B}\right)^{-2}+\left(\frac{H}{B}\right)^{-0.66}(c+CB^{-0.66})\right)^{-1}+\rho_{\text{sat},K}\frac{H^2}{H^2+H_{\text{sat},K}^2}. \tag{16}$$

Optimum agreement with experiment is achieved when $\rho_{\text{sat}}=0.4$, $c=0.5$, and $B(T=90\ \text{mK})=1.2$ T. We note, however, that in the strong fields corresponding to the $H^{2/3}$ law for the orbital positive magnetoresistance this sample exhibits a weaker temperature dependence of the corresponding slopes than that following from the theory in Ref. 2. It is conceivable that such behavior results from passage to a still weaker field dependence of the positive magnetoresistance as H increases.[1] To take this circumstance into account, we also introduced the fitting parameter C in (16), selecting its value at 2.5.

Thus, we employed the assumptions $k \propto T^{1/2}$ and $k(T=150\ \text{mK})=1.2$ for k. We stress, however, that the principal features of the observed behavior of sample No. 1 (the suppression of the negative magnetoresistance and the appearance of a positive magnetoresistance feature) are also reproduced for $k \propto T^{1/4}$, although neither features is as well-defined as in the case of $k \propto T^{1/2}$.

The calculated curves are presented in Fig. 6b along with experimental points for comparison. It is clear that the agreement for the sample under consideration can be regarded as satisfactory, and that it is better at weak fields, where suppression of the negative magnetoresistance is also manifested. At the same time, the curves constructed under the assumption that $k=\text{const}$ do not permit a description of the experiment at all. For 60 mK we also present the calculated curve for $k \propto T^{1/4}$. At the same time, for 36 mK the curve obtained without consideration of the spin contribution described in Ref. 6 is presented. It is clear that in the latter case it is not possible to achieve a faithful description of the observed positive magnetoresistance feature.

[1] B. I. Shklovskiĭ, Fiz. Tekh. Poluprovodn. **17**, 2055 (1983) [Sov. Phys. Semicond. **17**, 1311 (1983)].
[2] B. I. Shklovskii and B. Z. Spivak, in *Hopping Transport in Solids*, M. Pollak and B. Shklovskii (eds.), Elsevier, Amsterdam (1991), p. 271.
[3] M. E. Raikh, J. Czingon, Qiu-yi Ye, F. Koch *et al.*, Phys. Rev. B **45**, 6015 (1992).
[4] H. L. Zhao, B. Z. Spivak, M. P. Gelfand, and S. Feng, Phys. Rev. B **44**, 10 760 (1991).
[5] R. Rentzsch *et al.*, in *Hopping and Related Phenomena*, Proceedings of the 5th International Conference, Glasgow, Scotland (1993), p. 214.
[6] A. Kurobe and H. J. Kamamura, Phys. Soc. Jpn. **51**, 1904 (1982).
[7] N. V. Agrinskaya and V. I. Kozub, Zh. Éksp. Teor. Fiz. **106**, 848 (1994) [JETP **79**, 466 (1994)].
[8] N. V. Agrinskaya, V. I. Kozub, and D. V. Shamshur, Zh. Éksp. Teor. Fiz. **107**, 2063 (1995) [JETP **80**, 1142 (1995)].
[9] W. Schirmacher, Phys. Rev. B **45**, 2461 (1990).
[10] M. E. Raikh and G. F. Wessels, Phys. Rev. B **47**, 15 609 (1993).
[11] A. Frydman and Z. Ovadyahu, Solid State Commun. **94**, 745 (1995).
[12] R. Rentzsch, N. V. Agrinskaya, P. Fozooni, A. N. Ionov, M. J. Lea, B. Sandow, and D. V. Shamshur, in *Hopping and Related Phenomena*, Proceedings of the 6th International Conference, Jerusalem, Israel (1995), p. 187.
[13] R. Mansfield, in *Hopping Transport in Solids*, M. Pollak and B. Shklovskii (eds.), Elsevier, Amsterdam (1991), p. 349.
[14] T. G. Castner, in *Hopping Transport in Solids*, M. Pollak and B. Shklovskii (eds.), Elsevier, Amsterdam (1991), p. 1.

Translated by P. Shelnitz

Features of charge density waves in quasi-one-dimensional conductors at low temperatures

S. N. Artemenko*)

Institute of Radio Engineering and Electronics, Russian Academy of Sciences, 103907 Moscow, Russia
(Submitted 30 September 1996)
Zh. Éksp. Teor. Fiz. **111**, 1494–1512 (April 1997)

The reported study of charge density waves at low temperatures is based on a microscopic theory. One feature of a charge density wave at low temperatures is a large shift in the chemical potential near its defects, such as solitons, dislocations, and pinning centers, which leads to a higher conductivity of material along chains when the charge density wave is immobile, and the wave dynamics is controlled by this shift. Equations describing the dynamics of a charge density wave have been derived and used to estimate the velocity of 2π-solitons along conducting chains. The resulting soliton mobility has proved to be low and makes a small contribution to the conductivity. The large shift in the chemical potential near strong pinning centers can lead to a considerable increase in linear conductivity along conducting chains. [S1063-7761(97)02704-2]

1. INTRODUCTION

It is known[1] that below the Peierls transition temperature a charge or spin density wave is formed in a quasi-one-dimensional conductor; as a result, the conductor transits to a semiconducting (or semimetallic, such as $NbSe_3$) state. Transport properties of materials with charge and spin density waves are quite similar, and almost all statements in this paper concerning charge density waves also apply to spin density waves. In the Peierls state, single-electron excitations (electrons and holes) coexist with a deformable electronic crystal, which is called a charge density wave. If electric field E applied to a sample is higher than the threshold value E_T, which is a function of impurity concentration and temperature, the charge density wave is driven across the crystal and contributes to the electric current; as a result, the conductivity increases with electric field by several orders of magnitude. When $E<E_T$, the charge density wave cannot move as a whole, and the conductivity of the quasi-one-dimensional conductor is controlled by single-electron excitations. The material is then similar to an ordinary semiconductor, the only difference being that deformation of the charge density wave (for example, due to electric field) alters the concentration of electrons and holes, so perturbations of this wave alter the conductivity at a field strength below the threshold value. The change in the conductivity due to a charge density wave may be caused not only by its motion as a whole, but also by motion of its nonlinear excitations or defects of the electronic crystal, such as solitons and dislocations.

Presently the theory yields a fairly adequate description of conductors with charge density waves at relatively high temperatures $T>T_P/3$,[2] when the effects mentioned above are quite reproducible and described in terms of a charge density wave as an elastic medium whose interaction with impurities can be described in terms of the theory of weak (collective) pinning.[3] At lower temperatures, understanding of properties of charge density waves is not so clear. The activation energy of conductivity along chains at field strengths below the threshold value drops with temperature, whereas the activation energy of the transverse conductivity is constant,[4,5] and this behavior is often ascribed to the contribution of mobile defects in the charge density wave, such as 2π-phase solitons. Besides, at low temperatures the threshold field is considerably higher, the activation energy of nonlinear conductivity is a function of electric field, and the spread of sample parameters is notably wider, which indicates a greater role of defects at low temperatures. At very low temperatures (for example, below 20 K for TaS_3), there is a maximum in the low-frequency dielectric constant,[6–8] a greater role is played by metastable states and low-energy excitations of charge density waves, and new effects manifesting in the electric and thermodynamic parameters can be ascribed to glass-like properties of charge density waves.[9]

Models based on the existence of metastable states with different coordinate dependencies of the phase near pinning centers and transitions between such states have been suggested.[10–12] An interpretation of the low-temperature maximum in the dielectric constant as a manifestation of a relaxation mode due to impurity pinning has also been proposed.[13] The papers quoted above, however, did not take into account a possibility of a large shift in the chemical potential from the midgap position due to deformations of the charge density wave at low temperatures.

The example of soliton domain walls in a commensurate charge density wave[14] demonstrated that this shift should lead to a nonmonotonic temperature dependence of the screening range of nonuniform perturbations in the charge density wave, and a model was proposed for interpretation of the dielectric constant at low temperatures[15] based on this temperature dependence with charge density wave deformations caused by periodically distributed pinning centers. A large shift in the chemical potential should also occur[16,17] in 2π-solitons and around pinning centers, and this shift can generate metallic islands, where the chemical potential is in the continuum of real states either above or below the Peierls gap. As a result, the structure and typical dimensions of phase perturbations strongly depend on the temperature, and

 1063-7761/97/040823-10$10.00

single-electron excitations cannot be ignored even at very low temperatures. This, in particular, results in temperature dependence of the energy of both phase solitons and amplitude solitons.[18]

Thus, we need kinetic equations for the charge density wave and single-electron excitations that take account of the strong deformation of the wave around defects and its nonlinear perturbations in order to describe correctly the contribution of charge density waves to the low-temperature conductivity and, in particular, to test whether 2π-solitons (vacancies or interstitial defects of the electronic crystal) and other defects of the charge density wave can act as charge carriers. This paper presents such equations, which are different from the standard semiclassical equations since they take into consideration large differences between the charge density wave phases at neighboring chains and large deviations of the chemical potential from the midgap position. Thus nonlinear effects in the Coulomb screening of electric field generated by charge density due to deformation of a charge density wave is taken into account in the self-consistent field approximation. The resulting equations will be applied to estimates of the 2π-soliton velocity and their contribution to the conductivity.

In our calculations, the electron charge, Planck's and Boltzmann's constants are equated to unity.

2. BASIC EQUATIONS

In deriving the equations of motion for the charge density wave and expressions for the current density, we use the equations for Green's functions integrated with respect to the momentum component along the conducting chains. Similar equations were derived[20,14] using the Keldysh technique for nonequilibrium processes[19] in the case when all parameters vary little over a distance of the order of the Fermi wavelength $2\pi/p_F$ along the chains and are continuous functions in the direction perpendicular to the chains. Such equations can be easily generalized to the case of large differences between phases of the charge density waves at neighboring chains, which will be needed in studying perturbations localized at one or several chains. To this end, one should transfer to Wannier's site representation with respect to chain numbers, similarly to the equations for a layered semiconductor in Ref. 21, without using an expansion in terms of small gradients in the transverse direction.[14] Then we use for simplicity the tight-binding approximation for electrons at the chain, i.e., we consider the nearest neighbor interaction in which the energy as a function of the transverse momentum is expressed as

$$\epsilon_\perp = 2t_\perp(\cos ap_y + \cos ap_z), \quad t_\perp \ll \Delta.$$

As a result, we have an equation for the Green's functions introduced by Keldysh, which are matrices with respect to time indices, the index identifying the sheets of the Fermi surface of the quasi-one-dimensional conductor at $+p_F$ and $-p_F$, and the chain number n:

$$\check{g} = \begin{pmatrix} \hat{g}^R & \hat{g}^K \\ 0 & \hat{g}^A \end{pmatrix},$$

where $\hat{g}^R$ and $\hat{g}^A$ are the retarded and advanced Green's functions, and $\hat{g}^K$ is the Green's function introduced by Keldysh, containing information about the electron distribution. The arrays of the functions $\check{g}$ satisfy the equation

$$\begin{aligned} & iv\frac{d\check{g}_{nm}}{dx} + t_\perp \sum_i (A_{nn+i}\check{g}_{n+im} - \check{g}_{nm+i}A_{m+im}) \\ & + i\left(\sigma_z \frac{d\check{g}_{nm}}{dt_1} + \frac{d\check{g}_{nm}}{dt_2}\sigma_z\right) \\ & + (i\sigma_y\Delta_n - \Phi_n\sigma_z)\check{g}_{nm} - \check{g}_{nm}(\sigma_y\Delta_m - \sigma_z\Phi_m) \\ & + \frac{i}{2}\nu_f[\sigma_z\check{g}_{nn}\sigma_z\check{g}_{nm} - \check{g}_{nm}\sigma_z\check{g}_{mm}\sigma_z] \\ & - \frac{i}{4}\nu_b[\sigma_x\check{g}_{nn}\sigma_x\check{g}_{nm} - \check{g}_{nm}\sigma_x\check{g}_{mm}\sigma_x + \sigma_y\check{g}_{nn}\sigma_y\check{g}_{nm} \\ & \quad - \check{g}_{nm}\sigma_y\check{g}_{mm}\sigma_y] = 0, \end{aligned} \tag{1}$$

where the products assume the convolution with respect to time and matrix product, x is the coordinate along the chains, the self-consistent electric potential is contained in the equations in the form of a chirally invariant combination with the phase

$$\Phi_n = \phi_n - \frac{v}{2}\frac{d\varphi_n}{dx} - \frac{1}{2}\frac{d\varphi_n}{dt},$$

ϕ_n is the matrix element of the electric potential in the basis of the Wannier functions of the n-th chain, Δ_n and φ_n are the amplitude and phase of the order parameter in the nth chain,

$$A_{nm} = \sigma_z \cos\frac{\varphi_n - \varphi_m}{2} + i \sin\frac{\varphi_n - \varphi_m}{2},$$

σ_k are Pauli's matrices, and summation in the term containing $t_\perp$ is performed over nearest neighbors. The last terms in Eq. (1) represent the integral of elastic collisions, ν_f and ν_b are the forward- and back-scattering rates, i.e., rate of scattering without and with transition between different sheets of the Fermi surface. Note that the equations for $\check{g}$ are expressed in the representation in which the phase of the order parameter is extracted.

The order parameter satisfies the self-consistency conditions[14,20]:

$$\begin{aligned} & i\left(1 + \frac{1}{\omega_Q^2}\frac{\partial^2}{\partial t^2}\right)(\sigma_y \cos\varphi_n + \sigma_x \sin\varphi_n)\Delta_n \\ & = \frac{\lambda}{2}\int (\hat{g}^K_{nn} - \sigma_z\hat{g}^K_{nn}\sigma_z)d\epsilon, \end{aligned} \tag{2}$$

where ω_Q is the frequency of phonons with the wave vector of the Peierls instability. Summation of the components of the matrix equation (2) yields an equation for the amplitude, and their difference yields the equation for the phase.

The charge density at the nth chain can be expressed as

$$\rho_n = \frac{2}{\pi v}\left(\frac{1}{8}\int \mathrm{Tr}(\sigma_z\hat{g}^K_{nn})d\epsilon - \Phi\right). \tag{3}$$

The density of current along the chain is expressed as

$$j_n(x)=\frac{1}{4\pi}\int \mathrm{Tr}\ \hat{g}^K_{nn}d\epsilon. \quad (4)$$

The current in the direction perpendicular to the chains due to transitions between the nth and $(n+1)$-th chains is determined by the equation

$$j_{n,n+1}(x)\propto t_\perp \int \mathrm{Tr}(A_{nn+1}\hat{g}^K_{n+1\,n}-\hat{g}^K_{nn+1}A_{n+1n})d\epsilon. \quad (5)$$

In a general case, it is very difficult to solve Eq. (1). We limit our analysis to the case of smooth phase perturbations, when the phase changes little over a distance of the order of the coherence length v/Δ and the frequencies are much smaller than Δ. We take into account only the perturbations of the gap caused by the coordinate dependence of the phase and ignore amplitude solitons.[22,23]

Two types of Green's function perturbations in a state with a nonuniform charge density wave can be distinguished. The first type are perturbations in the state of thermodynamic equilibrium due to equilibrium deformations of the charge density wave, such as phase solitons, dislocations, and pinning centers. Such perturbations are described by the equilibrium distribution function, which can be derived from the retarded and advanced Green's functions. The function $\hat{g}^K$ in this case is defined as

$$\hat{g}^K=(\hat{g}^R-\hat{g}^A)\tanh(\epsilon/2T).$$

The latter equation does not hold when the distribution function is nonequilibrium and dissipation takes place. In this case, one must solve an additional equation for $\hat{g}^K$, which describes, in particular, the quasiparticle distribution function. The function $\hat{g}^K$ contains an anomalous component describing the deviation of the distribution function from equilibrium and defining the current and friction coefficient of the charge density wave.[14] In the next section we will consider perturbations in equilibrium states.

3. EQUATION FOR A NONUNIFORM CHARGE DENSITY WAVE IN THE STATE OF THERMODYNAMIC EQUILIBRIUM

We apply to Eq. (1) the perturbation theory in the parameter $t_\perp$ describing coupling between chains and in the electric field component along the chains $-d\phi/dx$.

In the case of a uniform state and zero coupling between chains, one can derive from Eq. (1) convenient expressions for the retarded and advanced Green's functions in the form of implicit solutions:

$$g_{nn}=g_n\sigma_z+f_n i\sigma_y\equiv\frac{\epsilon_n}{\xi_n}\sigma_z+\frac{\tilde{\Delta}_n}{\xi_n}i\sigma_y, \quad (6)$$

where

$$\epsilon_n=\epsilon-\Phi_n+i\nu_+g_n^{R(A)}/2,\quad \nu_+=\nu_f+\nu_b,$$

$$\tilde{\Delta}_n=\Delta_n-i\nu_f f_n^{R(A)}/2,\quad \xi_n^{R(A)}=\pm\sqrt{\epsilon_n^2-\tilde{\Delta}_n^2}.$$

The density of states of electrons is determined by the diagonal components of the Green's functions through the expression

$$N(\epsilon)=(g^R-g^A)/2. \quad (7)$$

Equations (6) and (7) indicate that both deformation of a charge density wave and electric potential lead to bending of energy bands, and the density of states as a function of coordinates is determined by the energy shift $\Phi_n(x)$. Since the electrochemical potential (i.e., the sum of the electric and chemical potentials) should be constant in equilibrium and the phase gradient equal to the local change in the charge density wave vector, then

$$\Phi_n(x)=-\mu_n(x),$$

where μ_n is the local shift in the chemical potential from the middle of the gap. (Here we ignore the asymmetry between electrons and holes, which usually occurs in Peierls conductors[24] and leads to a small shift of μ from the midgap position in an undistorted conductor with charge density waves.)

According to Eqs. (6) and (7), in the limit $\nu\to 0$, the density of states has a square-root singularity at the gap edge. This singularity is blurred out owing to interaction between chains and scattering, which is considered to be moderate in this case, $\nu\ll\Delta$ (otherwise we would have a gapless Peierls state). The expressions for $\hat{g}^R$ and $\hat{g}^A$ with due account of scattering can be easily derived for energies $|\epsilon-\Delta|\ll\Delta$, where the diagonal g and off-diagonal f components of $\hat{g}^R$ and $\hat{g}^A$ matrices are similar, and the equations for $g^{R(A)}$ reduce to cubic equations. Their solutions are

$$g^{R(A)}=\frac{1}{4}\left(\frac{\Delta}{\nu}\right)^{1/3}\{i[(4\eta)^{1/3}-(1+\eta/2+\sqrt{1+\eta})^{1/3}-(1+\eta/2-\sqrt{1+\eta})^{1/3}]\pm\sqrt{3}[(1+\eta/2+\sqrt{1+\eta})^{1/3}-(1+\eta/2-\sqrt{1+\eta})^{1/3}]\}, \quad (8)$$

where $\eta=2(\epsilon-\Delta)^3/27\Delta\nu^2$, $\nu\equiv\nu_f+\nu_b$. For large η, when $\epsilon-\Delta\gg(\Delta\nu^2)^{1/3}$, the solutions given by Eq. (8) drop as $1/\sqrt{\epsilon-\Delta}$, and near the edge $(\eta\approx-1)$, when $|\epsilon-E_G|\ll(\Delta\nu)^{1/3}$, where the gap halfwidth $E_G=\Delta-3(\Delta\nu^2)^{1/3}$ is remormalized because of the spread of the density of states, Eq. (8) reduces to

$$g^{R(A)}=-i\left(\frac{\Delta}{2\nu}\right)^{1/3}\pm\frac{\sqrt{3}}{\nu^{2/3}}\left(\frac{\Delta}{2}\right)^{1/6}\sqrt{\epsilon-E_G}. \quad (9)$$

It follows from this equation that the density of states drops to zero at the gap edge, in accordance with numerical calculations.[25]

Now let us calculate corrections of the form $g_{x,nn}\sigma_x$ to the solution of Eq. (6) for $\hat{g}^R$ and $\hat{g}^A$. After substituting these corrections into the self-consistency condition (2), we obtain an equation for the phase. The interaction between chains described by the parameter $t_\perp/\Delta$ contributes only in the second order of the perturbation theory:

$$g_{x,nn}=-\sum_i\frac{it_\perp^2\Delta\ \sin(\varphi_n-\varphi_{n+i})}{\xi_n\xi_{n+i}(\xi_n+\xi_{n+i})}-\frac{i\hbar v\Delta}{2\xi_n^3}\frac{d\mu}{dx}. \quad (10)$$

The gradients of $\Phi_n(x)$ and $\Delta(x)$ also contribute to $g_{x,nn}$ and f_{nn}, which determine the perturbations of the phase and amplitude of the charge density wave, respec-

tively. By writing equations for the components of matrices in Eq. (1) in the explicit form and neglecting interaction between chains, we obtain identical equations for $g^{R(A)}_{x,nn}$ in the form

$$v^2\frac{d^2g_x}{dx^2}+4\tilde{\xi}^{\,2}g_x=-2iv\left(f\frac{d\Phi}{dx}+g\frac{d\Delta}{dx}\right), \tag{11}$$

where the energy is shifted by $\Phi_n(x)$,

$$\tilde{\xi}^{\,2}=\epsilon^2-\Delta^2+6i\nu\Delta f^R.$$

The expression for the correction which determines the perturbation $\Delta(x)$ has the form

$$\delta f=-\frac{ivg}{2(\epsilon g-\Delta f)}\frac{dg_x}{dx}, \tag{12}$$

where g and f are the advanced and retarded Green's functions in the zeroth approximation with respect to gradients and $t_\perp$.

If $\tilde{\xi}$ on the left-hand side of Eq. (11) is sufficiently large, the second coordinate derivative can be omitted, i.e., we can use a local approximation corresponding to the semiclassical approximation, which yields standard equations for the phase used, for example, in the problem of phase solitons.[26,27] It will be demonstrated below that this approximation applies if the chemical potential in bent energy bands does not overlap with bands of allowed states or penetrates little into these bands. The solution of Eq. (11) has the form

$$g_x=-\int_{-\infty}^{\infty}dx_1\frac{f\Phi'+g\Delta'}{2\tilde{\xi}}\exp\frac{2i\tilde{\xi}|x-x_1|}{v}, \tag{13}$$

where primes denote coordinate derivatives. Substitution of Eqs. (10), (13), and (12) into the self-consistency condition (2) yields equations for the phase and amplitude of a charge density wave. But integration with respect to energy in Eq. (2) is quite difficult in a general case, since g and f are complex functions of the factors leading to the spread in the density of states.

Therefore we consider a limiting case of an ideal quasi-one-dimensional conductor, which applies when typical quasiparticle energies (i.e., temperature T or $\mu-E_G$ for $|\mu|>\Delta$) are larger than both $(\Delta\nu^2)^{1/3}$ and the spread of the one-dimensional density of states due to the incomplete nesting of the sheets of the Fermi surface shifted by the charge density wave vector. In this case, the phase is determined by the equation

$$\begin{aligned}J\sum_i\sin(\varphi_{n+i}-\varphi_n)&=\frac{i\Delta}{4v}\int_{-\infty}^{\infty}d\epsilon\int_{-\infty}^{\infty}dx_1[\Delta\Phi'(x_1)+\epsilon\Delta'(x_1)]\\&\times\left(\frac{\exp(2i\xi^R|x-x_1|/v)}{(\xi^R)^2}-\frac{\exp(2i\xi^A|x-x_1|/v)}{(\xi^A)^2}\right)\\&\times\tanh\frac{\epsilon-\Phi(x)}{2T},\end{aligned} \tag{14}$$

where $\xi^{R(A)}=\pm\sqrt{(\epsilon\pm i0)^2-\Delta^2}$, and the symbol $i0$ indicates the side on which the integration contour in ϵ bypasses the singularities.

The integration on the right-hand side of Eq. (14) yields

$$\begin{aligned}\mathrm{Re}\sum_{n=0}^{\infty}\int_{-\infty}^{\infty}&\frac{2\pi\Delta_0T\{\Delta_0\Phi'(x_1)+[\Phi(x)+iT_n]\Delta'(x_1)\}}{v\tilde{\zeta}_n^{\,2}}\\&\times\exp\left(-\frac{2\tilde{\zeta}_n|x-x_1|}{v}\right)dx_1,\end{aligned} \tag{15}$$

where

$$\tilde{\zeta}_n=\sqrt{\Delta_0^2+[T_n+i\Phi(x)]^2},$$

$T_n=(2n+1)\pi T$. The coefficient J describing the interaction between chains is determined, in a general case, by integrating Eq. (10) and its expression is rather cumbersome.

Consider the expression for the case $\Delta\gg T$, which is realized practically at all temperatures below the fluctuation region and also for the case when a chemical-potential shift comparable to the gap width occurs only at one chain:

$$\begin{aligned}J=\frac{t_\perp^2}{v}&\frac{\Delta^2}{|\Phi|\sqrt{\Delta^2-\Phi^2/4}}\\&\times\left[2\arcsin\frac{|\Phi|}{2\Delta}-\theta(|\Phi|-\Delta)\right.\\&\left.\times\left(\frac{\pi}{2}-\arcsin\frac{2\Delta^2-\Phi^2}{|\Phi|\Delta}\right)\right].\end{aligned}$$

Note that the function $J(\Phi)$ is rather flat: $J(0)=t_\perp^2/v$, $J(0)/J(\Delta)\approx0.8$; therefore the model with the constant $J=J(0)$ yields qualitatively correct results.

In order to make the problem a closed one, Eqs. (14) and (15) must be supplemented with the equation describing the deviation of the charge density wave amplitude from its unperturbed value Δ_0:

$$\begin{aligned}\Delta-\Delta_0=\mathrm{Re}\sum_{n=0}^{\infty}2\pi v\Delta T\left(\frac{1}{\tilde{\zeta}_n}-\frac{1}{\zeta_n}\right)-\mathrm{Re}\sum_{n=0}^{\infty}\frac{v}{2}\int_{-\infty}^{\infty}dx_1\\\times\frac{\pi T[\mu(x)+iT_n][\Delta_0\mu''(x_1)+(\mu(x)+iT_n)\Delta''(x_1)]}{\tilde{\zeta}_n^{\,4}}\\\times\exp\left(-\frac{2\tilde{\zeta}_n|x-x_1|}{v}\right),\end{aligned} \tag{16}$$

where $\zeta_n=\tilde{\zeta}(\Phi=0)$. Note that the first term on the right-hand side of Eq. (16) contains the unperturbed value of the amplitude and should be linearized with respect to $\Delta-\Delta_0$.

If the integrand in Eq. (14) changes little over the distance of about $v/\sqrt{\Delta^2-\mu^2}$, the local approximation applies, and the equation can be reduced to the standard semiclassical equation for the phase[26,27] with an additional force proportional to the gradient of Δ. Let us prove that the local approximation holds when the shifted chemical potential is either within the band gap or slightly penetrates into the band of allowed states, i.e., metallic islands are small. To do this, we add and subtract $\tanh(\epsilon/2T)$ to $\tanh[(\epsilon-\Phi(x))/2T]$ on

the right-hand side of Eq. (14) and calculate the integrals with the factor $\tanh(\epsilon/2T)$ and the difference between the two hyperbolic tangents separately.

The main contribution to the integral of the term with $\tanh(\epsilon/2T)$ with respect to ϵ calculated using residues at the tangent poles in the complex plane comes from the region $\epsilon \sim i\Delta$. Therefore the typical range of the dropping exponent in the integral with respect to x_1 is small and approximately equals v/Δ, and since we consider perturbations almost constant over the coherence length, the functions Φ' and Δ' can be taken out of the integrand. After calculating the integrals, we find on the right-hand side of Eq. (14) the semiclassical result $N_s\Phi'$, where $N_s = 1 - \sqrt{2\pi\Delta/T}\,\exp(-\Delta/T) \sim 1$ for $\Delta > T$.

One can easily prove that the remaining integral with respect to ϵ, containing the difference $\tanh[(\epsilon - \Phi(x))/2T] - \tanh(\epsilon/2T)$ is proportional to the vanishingly small exponential $\exp[(\Delta - |\Phi|)/T]$ if $\Delta - |\Phi(x)| \gg T$, and it can be omitted. Thus, the local approximation applies to the regions outside metallic islands.

Now let us assume that the point with the coordinate x is within a small metallic island with the length l_m and show that the latter integral responsible for the nonlocal contribution to the equation is small at sufficiently small l_m. If $l_m \ll v/\max\{T, \sqrt{\Delta(|\Phi| - \Delta)}\}$, the typical range of the exponential is notably larger than the island dimension, and x in the exponential can be ignored. Since the functions $\Phi(x)$ and $\Delta(x)$ describing a soliton are odd, the integral with respect to x_1 contains only an odd function and vanishes. Thus, the nonlocal contribution is also small in the case of a sufficiently small metallic island.

In a general case of an arbitrary perturbation, the semiclassical approximation does not apply, and the full system of integro differential equations (14)–(16) must be solved.

4. TIME-DEPENDENT EQUATIONS

If electric field is applied to a conductor with charge density waves and electric current flows across it, the phase equation contains terms due to deviation from equilibrium and a quasiparticle current is generated. In order to calculate the quasiparticle distribution, one must solve the kinetic equation for $\hat{g}^K$, which can be reduced to the semiclassical kinetic equations for the distribution function in the case of smooth perturbations.[14] We are, however, interested in perturbations which are gradual along chains, but vary rapidly in directions perpendicular to the chains. Such perturbations occur, for example, in solitons and near pinning centers. In this section, we generalize the previously developed approach[14] to such perturbations.

Our analysis is limited to the case of low frequencies, when the typical times of changes in the electric field, phase of the charge density wave, etc. are larger than the energy and momentum relaxation times. In this case, the energy distribution of quasiparticles is described by a Fermi function with a chemical potential μ, which is a function of coordinates and time and in the nonequilibrium case, is, generally speaking, different from Φ. In calculating μ, we use the Poisson equation with the charge density on the nth chain defined by Eq. (3) with

$$\mathrm{Tr}(\sigma_z \hat{g}^K_{nn}) = 2N(\epsilon)\,\tanh[(\epsilon - \mu)/2T],$$

where the density of states $N(\epsilon)$ is defined by Eq. (7) with Green's functions calculated in the previous section:

$$\rho_n = -\frac{\kappa^2}{4\pi}\left[\frac{v}{2}\frac{d\varphi_n}{dx} + f(\mu_n)\right] + \frac{\varepsilon_\Delta}{4\pi}\frac{d^2\Phi}{dx^2}. \tag{17}$$

Here $1/\kappa$ is the screening radius in the metallic state, ϵ_Δ is the dielectric constant component due to the Peierls gap owing to the corrections to $g^{R(A)}$ (the derivation is given in Ref. 14), $f(\mu)$ is the contribution of single-electron excitations to the charge density, analogous to similar to the contribution of electrons and holes in conventional semiconductors:

$$f(\mu) = \int_\Delta^\infty d\epsilon N(\epsilon)[n_F(\epsilon - \mu) - n_F(\epsilon + \mu)], \tag{18}$$

n_F is the Fermi distribution function. The explicit expression for $f(\mu)$ depends on the broadening of the density of states near the gap edge, although the approximate shape of $f(\mu)$ is universal. If the density of states broadening in $g^{R(A)}$ for $\Delta - |\mu| \gg T$ is neglected,

$$f(\mu) = N_Q \sinh(\mu/T), \quad N_Q = \sqrt{2\pi\Delta/T}\,\exp(-\Delta/T),$$

and for $|\mu| - \Delta \gg T$

$$f(\mu) = \sqrt{\mu^2 - \Delta^2}.$$

Since all the functions of the coordinate along the chain are smooth, perturbations of the function $\hat{g}^K$ generated by an electric field aligned with the chains are semiclassical and largely controlled by perturbations in the quasiparticle distribution function, whereas perturbations of the functions $\hat{g}^{R(A)}$ can be neglected. The perturbation of the distribution function can be calculated by the equation[14]

$$n_z = \frac{vV_n' + (\nu_b/2)G_-\dot{\varphi}_n}{\nu_{\mathrm{eff}}}\frac{dn_F(\epsilon - \mu_n)}{d\epsilon}, \tag{19}$$

$$V_n = \Phi_n - \mu_n, \tag{20}$$

where

$$\nu_{\mathrm{eff}} = \nu_b G_-/2 + i\Delta F_+/G_-, \quad G_- = g^R - g^A,$$
$$F_+ = f^R + f^A,$$

and V_n is the electrochemical potential, which vanishes in the equilibrium state because the field-induced and diffusion components of the current cancel each other. A dot over a function denotes a time derivative.

The perturbation of the distribution function defined by Eq. (19) describes the longitudinal current and adds to the phase equation the term corresponding to friction of a moving charge density wave. The corresponding nonequilibrium corrections to the Green's function components have the form

$$\mathrm{Tr}\, g^K = G_- n_z, \quad g_x^{(a)} = -F_+ n_z.$$

Substituting these expressions into Eq. (4), we calculate the current density along the chains:

$$j_l=\frac{\sigma_{Nl}\nu_b}{2v}\left(\frac{\partial\varphi_n}{\partial t}+\int_{-\infty}^{\infty}G_-n_z d\epsilon\right)-\frac{\varepsilon_\Delta}{4\pi}\frac{d^2\Phi}{dxdt}$$

$$=\frac{\sigma_{Nl}\nu_b}{2v}(1-b)\frac{\partial\varphi_n}{\partial t}-\sigma_l V_n'-\frac{\varepsilon_\Delta}{4\pi}\frac{d^2\Phi}{dxdt}, \qquad (21)$$

where σ_{Nl} is the chain conductivity in the normal state (i.e., when $\Delta=0$). The expression for the quasiparticle conductivity σ_l, as well as the parameter b describing the effect of quasiparticles on the current due to charge density waves, depends on the density of states broadening and shift of the chemical potential. They are determined by the contributions of the first and second terms in the distribution function described by Eq. (19) to the integral in Eq. (21). Calculations of these parameters in the limit of pure material and $|\mu|<\Delta$ are given in Ref. 14. We do not give the explicit expression for b, because at low temperatures this parameter contributes little to the charge-density-wave current along the chains (although it largely controls the contribution of charge density waves to the Hall effect and thermal conductivity[28,29]). In this paper, a calculation of σ_l is given in the limit of $T\ll\Delta$ only for an arbitrary shift of the chemical potential:

$$\sigma_l=4\sigma_{Nl}\frac{\nu_b T}{\nu\Delta}\left[\ln\left(2\cosh\frac{E_G-|\mu|}{2T}\right)-\frac{E_G-|\mu|}{2T}\right]. \qquad (22)$$

In deriving this equation, the density of states broadening near the gap edge was neglected; in the opposite limit of large broadening due to a high scattering rate, the factor 4 in Eq. (9) should be replaced by 3.

Calculating the term due to the nonequilibrium contribution to the quasiparticle distribution function (19) for the phase equation and adding it to the phase equation derived in the previous section, we obtain the equation of motion for the phase. Here we write this equation, for simplicity, in the semiclassical approximation, whose applicability to metallic islands is restricted, as was mentioned above, to small island dimensions:

$$\frac{1}{2v}\frac{m^*}{m}\frac{\partial^2\varphi_n}{\partial t^2}+\gamma\frac{\partial\varphi_n}{\partial t}-\frac{v}{2}\frac{\partial^2\varphi_n}{\partial x^2}$$

$$+J\sum_i\sin(\varphi_n-\varphi_{n+i})=E_n, \qquad (23)$$

where m is the electron mass, $m^*=(1+4\Delta^2/\lambda\omega_Q^2)m$ is the ''effective mass of the charge density wave'' due to the term with the time derivative in Eq. (2) (recall that large effective mass is a feature distinguishing charge density waves from spin density waves, whose effective mass should be equal to the band electron mass[1]). The coefficient of friction is defined as

$$\gamma\frac{\partial\varphi_n}{\partial t}=\frac{i\Delta}{2v}\int F_+ n_z d\epsilon. \qquad (24)$$

Like the quasiparticle conductivity, γ is a function of the density of states broadening. This parameter was calculated in Ref. 14 for the case of zero broadening and $\Delta-|\mu|\gg T$, and it was demonstrated that most of the friction is due to quasiparticles with energies near the gap edge, so the integral in the expression for γ had to be cut off near the density of states singularity. Here we analyze the opposite limit $T\gg(\Delta\nu^2)^{1/3}$, when the main contribution to the integral in Eq. (24) is due to quasiparticles near the broadened edge of the density of states:

$$\gamma=\frac{3\nu_b\Delta^{1/3}T}{(2\nu)^{4/3}}\left[\ln\left(2\cosh\frac{E_G-|\mu|}{2T}\right)-\frac{E_G-|\mu|}{2T}\right]. \qquad (25)$$

Note that in the lowest-order approximation the right-hand side of Eq. (23) contains the electric field $E=-d\phi/dx$, but not the gradient of the electrochemical potential, which determines the quasiparticle current in Eq. (21). The contribution proportional to the chemical potential gradient appears if one takes into account quasiparticle corrections, which are small at low temperatures.

The stability study of Eq. (23) with the interaction between neighboring chains indicates that the homogeneous solution with equal phases on all chains is unstable against small phase perturbations. In the stable solution, the phase differences between neighboring chains is π. This conclusion corresponds to the well-known fact that in the tight-binding approximation the opposite sheets of the Fermi surface coincide after a shift by a wave vector corresponding to the period doubling in the direction perpendicular to the chains. Therefore the charge density wave generated after the Peierls transition also has a double period in the perpendicular direction. In what follows, we define φ_n as a phase deviation from the stable solution, in which neighboring chains are in antiphase. The equations for such deviations differ from Eq. (23) by the sign of J.

Unlike the response to a field aligned with the chains, the current density in the direction perpendicular to the chains cannot be described in the semiclassical approximation since we consider perturbations in which the phase difference between neighboring chains can be large. In this case, the expression for the current density must include not only quasiparticle distribution functions, but also corrections to the retarded and advanced Green's functions. In calculating Green's functions components that are off-diagonal in the chain indices, which are needed for determining the current between chains, one can neglect terms with derivatives and consider only the first order of the perturbation theory in the chain coupling parameter $t_\perp$.

Calculation of the combination of Green's functions in Eq. (5) for the current density among the chains yields

$$j_t=i\frac{\sigma_{Nt}\nu_+}{4d}\int_{-\infty}^{\infty}d\epsilon\left(\tanh\frac{\epsilon-V_n}{2T}-\tanh\frac{\epsilon-V_{n+1}}{2T}\right)$$

$$\times(F^{RA}+F^{AR}-F^{RR}-F^{AA}), \qquad (26)$$

where σ_{Nt} is the conductivity in the direction perpendicular to the chains in the normal state, and

$$F^{IJ}=\frac{g_n^I g_{n+1}^J-1+f_n^I f_{n+1}^J\cos(\varphi_n-\varphi_{n+1})}{\zeta_n^I+\zeta_{n+1}^J}.$$

Equation (26) indicates that the current between the nth and $(n+1)$-th chains is caused by the electrochemical potential difference

$$V_n-V_{n+1}=\Phi_n-\mu_n-\Phi_{n+1}+\mu_{n+1},$$

and depends on the chemical potential shifts on the chains since, according to Eq. (6), the components of the Green's functions g_n, f_n, f_{n+1}, and g_{n+1} as functions of energy are shifted by Φ_n and Φ_{n+1}, respectively. Note also that the conductivity σ_t is a function of the phase difference between neighboring chains, and this effect is similar to the tunneling current between two conductors with charge density waves.[30]

Calculation of integrals in Eq. (26) in a general case is cumbersome; therefore we will discuss only general features of the conductivity and its expressions in limiting cases. At low temperatures and small Φ_n and Φ_{n+1}, the conductivity is exponentially small:

$$\sigma_t \propto \sigma_{Nt}[1+\cos(\varphi_n-\varphi_{n+1})]\exp(-\Delta/T). \tag{27}$$

Note that under these conditions the longitudinal conductivity is also exponentially small:

$$\sigma_l \propto \sigma_{Nl}\ \exp(-\Delta/T).$$

If one of the chains has a large shift in the potential $\Phi_n \approx \mu_n$, which is usually the case near a pinning center or in a phase soliton,[16,17] the longitudinal and transverse conductivities behave differently. The conductivity along the chain described by Eq. (22) increases proportionally to $\exp(|\Phi|/T)$ owing to the high local density of quasiparticles, whereas in the transverse conductivity the large chemical potential shift affects only the pre-exponential factor in Eq. (27), which notably drops with Φ. The reason is that at small chemical potential differences between neighboring chains the conductivity is inversely proportional to the scattering rate, which is typical of semiconductors and leads to small denominators $\zeta_n^R+\zeta_{n+1}^A$ in Eq. (26). At large $\Phi_n-\Phi_{n+1} \gg \sqrt{\Delta T}$ the response is similar to the tunneling current between chains, and the denominators in the first two terms of Eq. (26) are no longer small, so the integration yields the approximate result

$$\sigma_t = O\left(\frac{\Delta \nu^2 T^{3/2}}{(\Delta \nu^2+\mu T^2)\mu^{3/2}}\right)\sigma_{Nt}\ \exp\left(-\frac{\Delta}{T}\right). \tag{28}$$

The equation of motion for the phase combined with the Poisson and continuity equations, Eq. (17) for the charge density, and Eqs. (21) and (26) for the current density determine the contribution of a moving charge density wave or its parts to the conductivity.

5. CONTRIBUTION OF PHASE SOLITONS TO CONDUCTIVITY

Both analytic[16] and numerical calculations[17] of phase solitons indicate that the difference between their structure and energy derived from such diverse models as the model of interaction between neighboring chains and the model of a single chain (which describes interaction among many chains in the self-consistent field approximation[10]) is purely quantitative. This means that essential features of solitons are independent of the specific lattice structure and energy spectrum of a quasi-one-dimensional conductor. In our estimates of the soliton velocity due to a voltage applied to a quasi-one-dimensional conductor given below, we will simplify the model and reduce it ultimately to the simplest model with one chain by neglecting in our equations the phase and potential perturbations on neighboring chains.

In this model, we will use the semiclassical equation (23) for solitons without metallic islands or with a small island. Since we seek a solution for a soliton moving along chains, all functions of time should have the form

$$\varphi(x,t) \equiv \varphi[x-x_s(t)];$$

therefore the time derivatives can be replaced with coordinate derivatives. Let us use dimensionless variables with energy, time, and length units being $t_\perp$, $1/t_\perp$, and $v/t_\perp$, respectively. Then the phase equation in the approximation linear with respect to the voltage and velocity takes the form

$$\frac{\partial \varphi_n}{\partial x}\left(\frac{m^*}{2m}\ddot{x}_s+\gamma\dot{x}_s\right)+\sum_i \sin(\varphi_n-\varphi_{n+i})=\frac{\partial(\mu_n+V_n)}{\partial x}, \tag{29}$$

where the dimensionless parameter γ is measured in units of $t_\perp/v$, and the electric potential is expressed in terms of the electrochemical and chemical potentials by means of Eq. (20).

The components of the Laplacian in the Poisson equation containing coordinates in the plane perpendicular to the chains is expressed in the discrete form, neglecting the terms with ϵ_Δ:

$$\zeta\sum_i (\phi_{n+i}-\phi_n)=\frac{1}{2}\frac{d\varphi_n}{dx}+f(\mu_n), \tag{30}$$

where the dimensionless parameter $\zeta=\hbar v/8e^2 \sim 10^{-2}$.

In solving this problem, we also need the continuity equation, which, after substituting into it the expressions for the charge and current density, takes the form

$$\frac{\partial}{\partial x}\left(\frac{\sigma_l}{\sigma_{Nl}}\frac{\partial V_n}{\partial x}\right)+a\sum_i \frac{\sigma_t}{\sigma_{Nt}}(V_{n+i}-V_n)+2\dot{x}_s\nu_b$$
$$\times\frac{\partial}{\partial x}\left[f(\mu)+b\frac{\partial\varphi_n}{\partial x}\right]=0, \tag{31}$$

where

$$a=(\sigma_{Nt}/\sigma_{Nl})(v^2/t_\perp d^2)\sim 1.$$

In order to calculate x_s, we multiply Eq. (29) by $\partial\varphi_n/\partial x$, sum over n, and integrate with respect to x from $-\infty$ to ∞. As a result, some of the terms can be expressed as integrals of function derivatives, whose perturbations vanish at infinity; therefore the nonzero contribution to the integrals is made only by several terms, namely

$$\left(\frac{m^*}{2m}\ddot{x}_s+\gamma\dot{x}_s\right)\sum_n \int_{-\infty}^{\infty}\left(\frac{\partial\varphi_n}{\partial x}\right)^2 dx=\sum_n \int_{-\infty}^{\infty}\frac{\partial V_n}{\partial x}\frac{\partial\varphi_n}{\partial x}dx. \tag{32}$$

In the first order of the perturbation theory, unperturbed solutions for a static soliton can be substituted into the integrals in Eq. (32).

In order to derive x_s from Eq. (32), one must derive the function $V_n(x)$ from the continuity equation (31). Since within most of the soliton the shift in the chemical potential on the central chain is large,[16,17], the conductivity σ_t, ac-

cording to Eq. (28), is negligible, and a solution of Eq. (31) can be derived by rejecting the second term in Eq. (32).

As a result, we find a solution for the gradient of the electrochemical potential:

$$\frac{\partial V_n}{\partial x} = -\frac{j_{\rm ext}}{\sigma_l} + \dot{x}_s \frac{2\nu_b \sigma_{Nl}}{\sigma_l}\left(f(\mu_n) + b\frac{\partial \varphi_n}{\partial x}\right), \tag{33}$$

where the integration constant is derived from the condition that far from the soliton the current is determined by single-electron excitations, $j_{\rm ext}=\sigma_{0l}E(x=\infty)$ is the current density and $\sigma_{0l}=\sigma_l(x=\infty)$ is the conductivity far from the soliton. After substituting Eq. (33) into Eq. (32) for x_s, the first term yields the force which gives rise to a soliton movement, and the second term yields the additional friction connected with a charge redistribution caused by the soliton movement. The nature of this contribution is similar to that of the thermally activated quasiparticle contribution to the damping of a charge density wave[31,32] moving as a whole in the presence of weak pinning centers at relatively high temperatures.

Thus, owing to the smallness of current between the chains at a large shift in the chemical potential from the midgap position, the gradient of the electrochemical potential described by Eq. (33) is determined by the phase and potential on the central chain. This provides justification for using the model with a single chain in analyzing the soliton mobility and its contribution to the conductivity. Therefore, taking into account that the phase and potential perturbations in the soliton drop rapidly with the distance from the central chain, we neglect the contributions of all the chains except the central one. Substituting Eq. (33) into Eq. (32), we obtain the equation of motion for the soliton:

$$M\ddot{x}_s + 2\nu_b\sigma_{Nl}\Gamma\dot{x} = Aj_{\rm ext}/\sigma_{0l}, \tag{34}$$

where

$$M = (m^*/2m)\int (\varphi')^2 dx, \quad A = \int \varphi'(\sigma_{0l}/\sigma_l)\,dx,$$

$$\Gamma = \int \varphi'[(f + b\varphi'/\sigma_l + \gamma\varphi'/2\nu_b\sigma_{Nl})]dx.$$

Let us average Eq. (33) over the coordinate x and substitute into it the soliton velocity derived from Eq. (34) in the limit of low frequencies, when the term with the effective mass can be rejected. Assuming that the linear density of noninteracting solitons on the chain is n_s, we obtain the relation between the average field and current, which determines the conductivity of a crystal with solitons:

$$\overline{E} = \frac{j_{\rm ext}}{\sigma_{0l}}\left[1 - n_s\left(\int \frac{\delta\sigma}{\sigma_l}dx + \frac{A}{\Gamma}\int \frac{f + b\varphi'}{\sigma_l}dx\right)\right], \tag{35}$$

where $\delta\sigma = \sigma_l - \sigma_{0l}$. The expression in the parentheses describes the decrease in resistivity due to solitons, the second term being the contribution due to the solitons' drift, and the first term is related to the local increase in the conductivity due to the screening of the soliton charge and chemical potential shift.

In calculating solitons' contribution to the linear conductivity, one can ignore the effect of the applied field on the soliton shape and substitute solutions for an immobile soliton into all the integrals in Eqs. (34) and (35). In the single-chain model and in the case when there is no metallic island at the middle of the soliton (or the island is small), the phase equations and Poisson equation describing a static soliton can be derived from Eq. (29) and (30) by omitting the variables describing neighboring chains. If the number of neighboring chains is four, we have

$$\frac{d\mu}{dx} = -4\sin\varphi, \quad \frac{1}{2}\frac{d\varphi}{dx} = -4\zeta\phi - TN_Q \sinh\frac{\mu}{T}. \tag{36}$$

Solutions to Eq. (36) have a simple form in two limiting cases: at $T=0$, when the effect of quasiparticles is negligible,

$$\varphi = 4\tan^{-1}[\exp(\sqrt{32\zeta}x)],$$

and at higher temperatures, when $N_Q \gg \zeta$ and the quasineutrality condition holds:

$$\varphi = 2\cot^{-1}[\sqrt{1 + 2/N_Q T^2}\sinh(8N_Q x)].$$

Substituting these expressions into Eq. (35), we obtain the resistivity of a sample with noninteracting solitons in the limiting cases mentioned above and when the temperature is sufficiently low: $N_Q T^2 \ll 1$. In dimensional units, the resistivity is

$$\rho = \frac{1}{\sigma_{0l}}(1 - n_s) \times \begin{cases} a_1/\sqrt{N_Q} + b_1 T/t_\perp & \text{for } N_Q \ll \zeta, \\ (a_2/\sqrt{\zeta} + b_2 T/t_\perp)\ln[t_\perp/T\sqrt{\zeta}] & \text{for } N_Q \gg \zeta, \end{cases} \tag{37}$$

where the factors a and b are of the order of unity, and their exact values depend on the lattice structure, specifically, the number of nearest neighbors. Equation (37) clearly indicates that in both limiting cases the terms with the factors a describing soliton's contribution to the conductivity due to the chemical potential shift are notably larger than the contribution of the soliton drift described by the terms with the factors b. The soliton mobility is low because the force acting on a soliton when current flows across the sample is low owing to the screening of the electric field in the central region of the soliton, where the chemical potential shift is large and the local quasiparticle concentration is higher. Therefore the main contribution to the force driving the soliton is largely generated at large distances from the center, where the chemical potential shift is small and the applied electric field is not screened by quasiparticles, whereas the entire region of the phase perturbations in the soliton contributes to the factor Γ responsible for the soliton friction. Note that the main contribution to the friction is due to the first term in Γ, which describes the dissipation caused by the charge redistribution at the soliton center during its motion.

Note that pinning impurities can also make a contribution to the linear conductivity similar to the soliton's contribution to the conductivity owing to the chemical potential shift, and unrelated to its drift. The chemical potential is also shifted around a strong pinning center,[16,17] and the drop in this shift with the distance from the center is similar to that

in a soliton. As it is well known, the pinning center can be described by adding to the first of Eqs. (36) the term $v_Q\delta(x-x')\sin(Qx+\varphi)$, where the dimensionless potential $v_Q=2V_Q\Delta/\lambda t_\perp v$ is expressed through the Fourier component of the impurity potential V_Q corresponding to the wave vector Q of the charge density wave.[14] It follows from the solution of Eqs. (36) with δ-functions that the pinning center generates a local phase perturbation with a considerable shift of the chemical potential, which decays at a distance of the order of the soliton length. For example, in the limit $T\to 0$ the phase φ_i and chemical potential μ_i near an impurity are determined by the relationship

$$2|\mu_i|=\sqrt{8/\zeta}\,\sin(\varphi_i/2)=v_Q\,\sin(Qx_i+\varphi_i).$$

Hence it follows that pinning leads to a large shift in μ_i comparable to Δ if the impurity potential V is of the order of the Fermi energy, i.e., is comparable to characteristic atomic energies. Note that the formal solution in the model with the δ-function yields a potential jump at the impurity, but it does not invalidate this model because it means a rapid drop in the potential over a short distance of the order of the Thomas–Fermi screening length.

Note also that an investigation based on a three-dimensional model that takes into account many chains shows up the relatively slow drop of the perturbations, described by a power law, at large distances from the impurity, where the chemical potential shift is small, and this small decrease may result in a collective pinning.

The large shift in the chemical potential at the pinning center produces a contribution to the resistivity similar to that of the terms with the factors a in Eq. (37), but the soliton concentration n_s should be replaced by the linear concentration of pinning centers n_p. Since the soliton concentration at low temperatures $n_s\propto\exp(-E_s/T)$, it drops rapidly with decreasing temperature since it is controlled by thermal activation and the soliton energy increases with decreasing temperature.[16,17] Therefore the effect of pinning centers on the conductivity should be stronger at low temperatures than that of solitons.

Thus, our calculations contradict the concept used previously by many authors, which holds that the soliton drift can account for the conductivity along the chains exceeding the thermally activated conductivity with the activation energy close to Δ. One can see that the contribution of the soliton drift is very small. The increase in the conductivity along the chains due to the chemical potential perturbation around a pinning center is considerably larger.

Our calculations also indicate that this increase in the longitudinal conductivity should not be accompanied by the increase in the transverse conductivity, as is observed in experiments. The dimension of the region with enhanced conductivity is determined by the soliton length l_s, which increases with decreasing temperature and is a macroscopic parameter, whereas in the perpendicular directions the perturbation drops over a distance comparable to the interchain distance. Therefore regions with enhanced conductivity overlap in the longitudinal direction at relatively small concentrations of the pinning centers $n_p\sim 1/l_s$, which leads to a notable increase in the longitudinal conductivity over the thermally activated conductivity with a constant activation energy. The solution of the conductivity problem in the case of a relatively high pinning center concentration is very complicated and deserves special treatment. One can expect that in this case a large role should be played by metastable states with different spatial phase distributions around pinning centers.

6. CONCLUSIONS

It is well known that the effects due to long-range Coulomb interaction and its screening determine static and dynamic properties of nonuniform perturbations of charge density waves at high temperatures. For example, linear screening by single-electron excitations determines their typical range (rigidity of charge density waves)[20,33] and their velocity.[31,32] It follows from calculations given above that Coulomb effects should be also included in the description of low-temperature dynamics of charge density waves, when there are few electron–hole excitations in a uniform charge density wave.

In regions of nonuniform perturbations of the charge density wave, the chemical potential shifts notably from its equilibrium position near the middle of the Peierls gap, producing a considerable effect on the structure and mobility of nonlinear perturbations of the charge density wave.

We have derived equations which apply to the dynamics of perturbations with large variations of the phase and potential between neighboring chains, such as 2π-solitons or pinning centers. In such regions, the single-electron conductivity along the chains increases notably, whereas the conductivity between the chains does not change. The calculation of the soliton velocity due to applied electric field indicates that the velocity of solitons is too small, and their drift contributes little to the low-temperature conductivity. The differences between the observed temperature dependence of the linear conductivity and that expected from conventional electron–hole conductivity can be ascribed to the chemical potential shift and increase in the quasiparticle concentration around pinning centers.

The author is grateful to F. Gleĭsberg, P. Monceau, V. Ya. Pokrovskii, S. V. Zaĭtsev-Zotov and W. Wonneberger for stimulating discussions. Special thanks for the hospitality to the Centre de Récherches sur les Très Basses Températures (Grenoble, France), where part of this work was done under the auspices of an exchange program between the Landau Institute and the École Normale Supérieur. This work was supported by the Russian Fund for Fundamental Research, Grant No. 95-02-05392, and Physics of Solid-State Nanostructures Program, Grant No. 1-018.

*)E-mail: art@mail.cplire.ru.

[1] G. Grüner, *Density Waves in Solids*, Addison-Wesley, Reading (1994).
[2] *Charge Density Waves in Solids*, L. Gor'kov and G. Grüner (eds.), Elsevier Science Publications, Amsterdam (1989).
[3] P. A. Lee and T. M. Rice, Phys. Rev. B **19**, 3970 (1979).
[4] T. Takoshima, M. Ido, K. Tsutsumi, T. Sambongi, S. Honma, K. Yamaya, and Y. Abe, Solid State Commun. **35**, 911 (1980).
[5] S. K. Zhilinskiĭ, M. E. Itkis, F. Ya. Nad', and V. B. Preobrazhenskiĭ, Zh. Éksp. Teor. Fiz. **85**, 362 (1983) [Sov. Phys. JETP **58**, 211 (1983)].

[6] Jie Yang and N. P. Ong, Phys. Rev. B **44**, 1991 (1991).
[7] G. Kriza, Y. Kim, A. Beleznay, and G. Mihaly, Solid State Commun. **79**, 811 (1991).
[8] F. Ya. Nad' and P. Monceau, Solid State Commun. **87**, 13 (1993); Synthetic Metals **70**, 1255 (1995).
[9] J. C. Lasjaunias, K. Biljakovic, and P. Monceau, Phys. Rev. B **53**, 7699 (1996).
[10] A. I. Larkin, Zh. Éksp. Teor. Fiz. **105**, 1793 (1994) [JETP **78**, 971 (1994)].
[11] A. Larkin and S. Brazovskii, Solid State Commun. **93**, 275 (1995).
[12] Yu. N. Ovchinnikov, K. Biljakovic, and P. Monceau, Europhys. Lett. **34**, 645 (1996).
[13] W. Wonneberger, Solid State Commun. **97**, 891 (1996).
[14] S. N. Artemenko and A. F. Volkov, in *Charge Density Waves in Solids*, L. Gor'kov and G. Grüner (eds.), Elsevier Science Publications, Amsterdam (1989), Ch. 9.
[15] A. F. Volkov, Physics Lett. A **182**, 433 (1993).
[16] S. N. Artemenko and F. Glyaisberg, JETP Lett. **61**, 779 (1995).
[17] S. N. Artemenko and F. Gleisberg, Phys. Rev. Lett. **75**, 497 (1995).
[18] S. N. Artemenko, JETP Lett. **63**, 43 (1996).
[19] L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)].
[20] S. N. Artemenko and A. F. Volkov, Zh. Éksp. Teor. Fiz. **81**, 1872 (1981) [Sov. Phys. JETP **54**, 992 (1981)].
[21] S. N. Artemenko, Zh. Éksp. Teor. Fiz. **79**, 162 (1980) [Sov. Phys. JETP **52**, 81 (1980)].
[22] S. A. Brazovskiĭ, JETP Lett. **28**, (1978) [*sic*]; Zh. Éksp. Teor. Fiz. **78**, 678 (1980) [*sic*].
[23] W. P. Su, J. R. Schrieffer, and A. J. Heeger, Phys. Rev. Lett. **42**, 1698 (1979).
[24] S. N. Artemenko, V. Ya. Pokrovskiĭ, and S. V. Zaĭtsev-Zotov, Zh. Éksp. Teor. Fiz. **110**, 1069 (1996) [JETP **83**, 590 (1996)].
[25] R. H. McKenzie and J. W. Wilkins, Phys. Rev. Lett. **69**, 1085 (1992).
[26] B. Horovitz, J. A. Krumhansl, and E. Domany, Phys. Rev. Lett. **38**, 778 (1977).
[27] S. A. Brazovskiĭ and S. I. Matveenko, Zh. Éksp. Teor. Fiz. **99**, 887 (1991) [Sov. Phys. JETP **72**, 492 (1991)].
[28] S. N. Artemenko and A. N. Kruglov, Fiz. Tverd. Tela **26**, 239 (1984) [*sic*].
[29] S. N. Artemenko, Synthetic Metals **36**, 381 (1990).
[30] S. N. Artemenko and A. F. Volkov, Zh. Éksp. Teor. Fiz. **91**, 1536 (1984) [Sov. Phys. JETP **64**, 906 (1984)].
[31] L. Sneddon, Phys. Rev. B **29**, 719 (1984).
[32] P. B. Littlewood, Phys. Rev. B **36**, 480 (1987).
[33] Y. Kurihara, J. Phys. Soc. Jap. **49**, 852 (1980).

Translation provided by the Russian Editorial office.

Electrical conductivity in the metal–polymer–metal system: the role of boundary conditions

V. M. Kornilov and A. N. Lachinov

Institute of Physics of Molecules and Crystals, Russian Academy of Sciences, 450065 Ufa, Russia
(Submitted 30 November 1995; resubmitted 18 October 1996)
Zh. Éksp. Teor. Fiz. **111**, 1513–1529 (April 1997)

We study a new type of charge instability in electroactive polymers, an instability caused by a change in the boundary condition at the metal–polymer interface. The change in the boundary conditions is achieved by melting one of the electrodes in a measuring cell of the metal–polymer–metal sandwich type. Charge instability manifests itself in the form of an insulator–metal phase transition, which emerges when one of the electrodes is melted. We list the results of numerous studies of the role of artifacts. Finally, we propose a model that explains several features of the phenomenon. © *1997 American Institute of Physics.* [S1063-7761(97)02804-7]

1. INTRODUCTION

The existence of high-conductivity states in undoped polymers, discovered at the end of the 1980s,[1–3] still remains a remarkable but unexplained phenomenon. What is remarkable is the extremely low value of the threshold perturbations that initiate an insulator–metal phase transition in such polymers.[4] In particular, the pressure threshold is about 10^4 Pa (Ref. 5) and the field threshold is 10^3 V/cm (Ref. 6). These quantities are so small that they certainly have no effect on the energy band structure. It is known,[7] for instance, that the metallic state in electroactive polymers of the polyacetylene type is reached at pressures of about 10^{10} Pa, which is considerably higher than the threshold value. Here the transition is attributed not to a change in the intermolecular bonds, for the pressure is not sufficiently high, but to an increase in the strength of the intermolecular interaction.

It can be assumed that in a metal–polymer–metal system such perturbations affect the energy band structure of the polymer indirectly, leading to energetically favored variations in the polymer sample. This may occur, for instance, because of variations in the space or surface charge density. Generation of the high-conductivity state of a polymer sample under rapid ionization at the peak of the thermally stimulated current was discussed in Ref. 8. There it was also noted that exciting the high-conductivity state requires attaining a certain critical concentration of the uncompensated charge in the sample's bulk. Such concentration was achieved by selecting an appropriate rate for heating (or cooling) the sample.

The effect of trapping states on the threshold characteristics of the formation of the high-conductivity state in polymer films was studied in Refs. 8 and 9. It was established that by increasing the trap concentration the high-conductivity state can be reached in samples that are 100 μm thick, rather than 0.1–1 μm thick as reported earlier.

In thin (up to 1 μm) samples the surface plays an important role, since the concentration of defects and, consequently, of traps on it is high. If we bear in mind that this sample thickness is comparable to the depth of penetration of the space charge generated by the bending of the energy bands at the surface, we may assume that the surface states have a strong effect on the conditions in which the high-conductivity state is formed. A brief report on this effect can be found in Ref. 10.

In view of this, the goal of the present work is to study the role of surface states at the metal–polymer interface in the process of generation of the high-conductivity state in the polymer field.

The idea behind the experiment is as follows. When a metal and a polymer are in contact, the bands near the surface (the interface) bend as a result of the evening out of the Fermi levels and the redistribution of charges. The potential energy $V_b(z_0)$ of the carriers at a distance z from the phase boundary determines the magnitude of this band bending:[11]

$$V_b(z)=\begin{cases} V_0-\dfrac{2\pi e^2 N(z-z_0)^2}{\varepsilon} & \text{if } 0<z<z_0, \\ V_0 & \text{if } z>z_0, \end{cases} \tag{1}$$

where z_0 is the depth of penetration of the polymer's bulk by the surface charge N. According to various estimates, z_0 may vary by a factor of ten depending on the type of polymer and the estimation method.[12,13] Note that the average value $z_0 \sim 3\ \mu$m (Ref. 13) is comparable to the thickness of the films employed earlier in studies of switching to the high-conductivity state.[4–6]

Using the expression for $V_b(z_0)$, we can determine the surface charge density

$$N=\frac{\varepsilon W_0}{2\pi(ez_0)^2}, \tag{2}$$

where $W_0=\varphi_p-\varphi_m$, with φ_p and φ_m the work functions of the polymer and the metal, respectively. Thus, N can be changed by varying the ratio of the work functions of the polymer and the metal. Some metals, e.g., indium, are known[14,15] to change their work function very rapidly in the course of melting. Hence, if we use such a metal for one of the electrodes, at the transition temperature the charge density at the polymer–metal phase boundary may change dras-

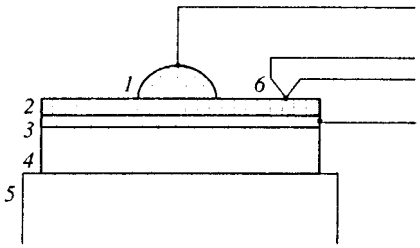

FIG. 1. Schematic of the measuring cell: *1* is the upper electrode, *2* is the polymer film, *3* is the lower electrode, *4* is the glass support plate, *5* is the heater, and *6* is the thermocouple.

tically, which, as expected by analogy with thermally stimulated switching,[8] leads to the appearance of the high-conductivity state in the polymer sample.

2. SAMPLES AND METHODS

We studied polymers of the polyphthalidylidenearylene type, in which earlier Zolotukhin *et al.*[16] observed phenomena associated with generation of high-conductivity states. The results are given for a typical representative of this class of polymers, polyphthalidylidenebiphenylilene (PPB).

PPB is easily soluble in chloroform and cyclohexanone. This allows fabrication of polymer samples of different thicknesses (from fractions of a micrometer[17] to several dozen micrometers). To fabricate uniform films we used the method of centrifuging a solution of PPB in cyclohexanone.

The polymer films needed for electrophysical measurements were fabricated on a support made of polished glass with a flat metal electrode (we called this the lower electrode), Optically polished microscope glass of the *Teget* brand or cover glass of the *Ilmglas* brand was used. The experimental measuring cell was of the ''sandwich'' type (Fig. 1). The lower electrode was manufactured by vacuum thermal sputtering of two metal layers: a $\sim$200–500 Å inner layer of chromium or vanadium for better adhesion to the glass, and a $\sim$1000 Å contact layer of copper, gold, aluminum, calcium, etc. to ensure good conductivity and the possibility of attaching leads.

Various metals, such as indium, gallium, and Wood's alloy, were used for the electrode (the upper electrode) on the surface of the polymer film.

The block diagram of the electric circuit is depicted in Fig. 2. The measuring cell was placed in a heater that allowed heating the cell to 250–300 °C. The voltage across the sample was supplied by a B5-12 power source. The ballast resistor was needed to limit the current flowing through the sample, with the resistance chosen in such a way that $R_s \gg R_r$ and $R_s > R_b/100$. Here the first inequality follows from the condition that the reference resistor introduces the smallest possible disturbance in the measuring circuit, while the second reflects the accuracy of the measuring devices.

The signal on the reference resistor was measured by feeding it to the Y terminal of an N307 chart recorder, and scanning along in the X coordinate was done via the signal from a differential copper–constantan thermocouple, with one junction in contact with the sample and the other inside a vessel with melting ice.

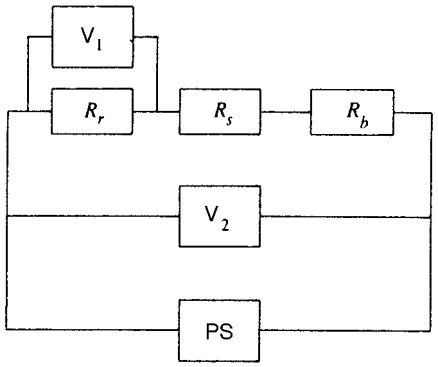

FIG. 2. Block diagram of the electric circuit used for measuring the temperature dependence and the current–voltage characteristics: R_s is the sample, R_r is the reference resistor, R_b is the ballast resistor, V_1 and V_2 are voltmeters, and PS is the power source.

3. RESULTS OF MEASUREMENTS

3.a. The effect of temperature on the electrical conductivity of the metal–polymer–metal system

Figure 3 depicts the temperature dependence of the current flowing through a polymer film. Here curve *1* corresponds to the case where Wood's alloy is used for the upper electrode, curve *2* to the case of indium, and *3* to the case of gallium. A characteristic feature of all the curves is a sharp increase in conductivity near the melting points of the upper electrode material. When the temperature is lowered the low conductivity of the metal 1–polymer–metal 2 system is restored.

Let us discuss in greater detail the case when Wood's alloy was used for the upper electrode. At room temperature the polymer sample has a specific conductance $\sim 10^{-14}(\Omega\text{cm})^{-1}$ (the sample resistance was $\sim$10 GΩ). As the temperature rises to 40–45 °C, current fluctuations begin to appear in the circuit, with the fluctuations corresponding to a drop in potential on the ballast resistor amounting to 1–10% of the applied voltage (Fig. 4). In the process the resistance of the system consisting of the lower electrode, the

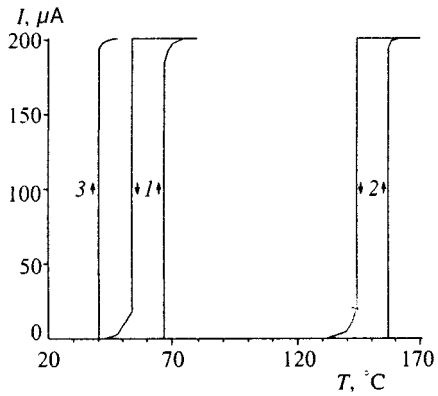

FIG. 3. The temperature dependence of the current flowing through the sample for different metals acting as the upper electrode: *1*—Wood's alloy, *2*—indium, and *3*—gallium. The voltage supplied by the power source was 10 V, the arrows indicate the direction in which the temperature changed, the cell heating rate was 5 °C min^{-1}, and the polymer layer was 2 μm thick.

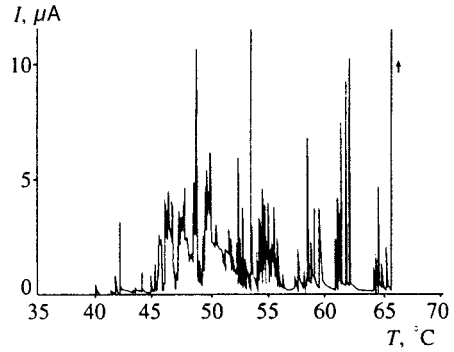

FIG. 4. Current fluctuations in the pre-transition region (the sample temperature is lower than the melting point of the upper electrode, and the arrow indicates the sample's transition to the high-conductivity state). The voltage supplied by the power source was 10 V, the cell heating rate was 5 °C min^{-1}, and the polymer layer was 2-μm thick.

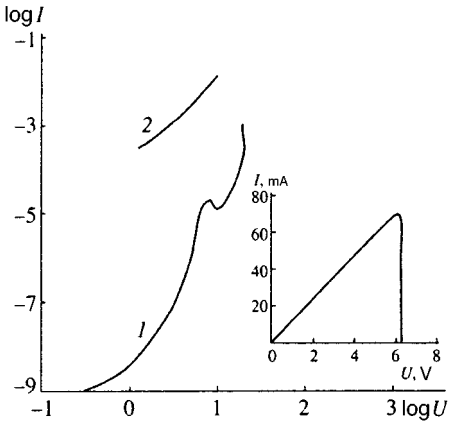

FIG. 6. Current–voltage characteristics of the polymer sample at different temperatures: *1*—at $T=20$ °C, and *2*—at $T=60$ °C. The inset depicts the current–voltage characteristic for temperatures higher than the melting point of the upper electrode. Wood's alloy was used for a low-melting metal.

polymer, and the upper electrode drops to 1–5 MΩ. Current fluctuations exist up to the melting point of the upper electrode.

At the melting point of the upper electrode (68 °C for Wood's alloy) there is a sudden jump in the current in the circuit. The voltage drop on the ballast resistor in this case becomes roughly equal to the applied voltage. Here the resistance of the system consisting of the lower electrode, the polymer, and the upper electrode amounts to 0.1–5 Ω.

The temperature dependence of the conductivity of this system exhibits metallic behavior, since the sample's resistance grows when the temperature of the measuring cell is increased to 250 °C (Fig. 5). Here the calculated temperature resistance coefficient is 3×10^{-3} K^{-1}, which within the experimental error coincides with the temperature resistance coefficient of the electrodes. The temperature dependence of the polymer sample proper is masked in this case by the properties of the electrodes but, we must note, is not of an activation nature.

A similar pattern is observed when indium is used for the upper electrode (curve *2* in Fig. 3). As the temperature rises to 125–130 °C, current fluctuations in the measuring circuit increase, with the resistance dropping to 1–10 MΩ. When the melting point for indium (156.2 °C) is reached, a sharp jump in current is registered in the measuring circuit, with a corresponding drop in the sample's resistance to 0.1–5 Ω. A further increase in the temperature of the measuring cell to 220 °C results in only a slight increase in the sample's resistance.

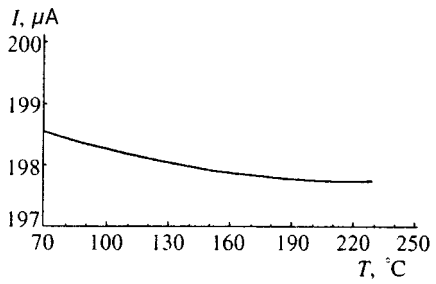

FIG. 5. The temperature dependence of the current flowing through the sample when the sample is heated to a temperature higher than the melting point of the upper electrode. The voltage supplied by the power source was 10 V, the cell heating rate was 5 °C min^{-1}, and the polymer layer was 2-μm thick.

Finally, for gallium as the material of the upper electrode (curve *3* in Fig. 3) the sample goes into a low-resistance state as the temperature of the measuring cell rises to the melting point of gallium (29.75 °C).

When the temperature of the measuring cell is lowered, all processes involving changes in conductivity are reversed. When the solidification point of the upper electrode is reached, the current in the measuring circuit suddenly drops (this is depicted by the arrows in Fig. 3). The observed hysteresis phenomenon can probably be explained by the inertia of the measuring cell. When the temperature is lowered still further, current fluctuations become evident, and at room temperature the system returns to its initial insulator state.

3.b. The effect of temperature on the current–voltage characteristics of the metal–polymer–metal system

Studies of the current–voltage characteristics of the film sample at different temperatures have revealed the existence of three types of such characteristics. The first type (curve *1* in Fig. 6) is observed at the initial (room) temperature, when the sample is in the insulator state. This type corresponds to an exponential dependence typical of charge transfer through polymers and is often explained by Frenkel–Paul processes.[18]

The second type is represented by a power function with an exponent close to two (curve *2* in Fig. 6). Such a current–voltage characteristic is observed near the transition point when the temperature is raised or lowered. Note that upon reversal of the direction of temperature variation after the transition has occurred, current–voltage characteristics of this type are observed even at the temperature at which the measurements were started. Restoration of an exponential current–voltage characteristic is achieved after a certain time has elapsed (roughly one hour).

Usually the power-like nature of the current–voltage characteristic is explained by the presence of a space charge

in the system, which limits charge transfer.[19] The slow restoration of the initial type of the current–voltage characteristic can probably be explained by the long recombination times of the space charge.

The third type (the inset in Fig. 6) appears at temperatures above the electrode melting point and after the polymer has reached the high-conductivity state. The characteristic has a linear ohmic shape.

3.c. Analysis of the role of artifacts

The sharp drop in the sample resistance at the melting point of upper electrode can be explained in different ways. First, there can be direct contact between the electrodes due to through holes, inhomogeneities, and impurities in the polymer film. Second, the resistance can drop because of electrical breakdown of the polymer film at points where the electric field is locally inhomogeneous. Third, molten metal can diffuse through the polymer film, forming conducting metallic bridges in the process. To check these assumptions, we conducted additional experiments.

1. Through holes and inhomogeneities. We used polymer films of different thicknesses in our experiments. The thickness of a polymer film was varied by successive deposition of polymer layers from a solution of a fixed concentration. It was assumed that each successive polymer layer reduces the number of defects in the previous layer and diminishes the contribution of the defects to the switching effect in temperature. After a layer was deposited, the samples were studied to see whether the effect was still present. It was found that the effect reproduces itself equally successfully up to a maximum number of layers equal to eight. The overall thickness of the polymer layer in this case was no less than 5 μm.

To resolve the problem of the relative importance of the phase transition in the metal and the aggregation state of the metal we set up an experiment with a liquid metal electrode. When the upper electrode was liquid mercury deposited on the polymer film, no switching to the high-conductivity state was observed. We also note that in the liquid state such metals as Wood's alloy, indium, gallium, and mercury do not wet the polymer film.

2. Electrical breakdown. It was found that neither the type of material selected for the electrodes nor the polarity of the applied voltage affects the transition to the high-conductivity state.

An attempt was made to determine the minimum voltage of the power supply at which a switching effect still occurs, and it was found that this value is certainly lower than 1 mV. For different film thicknesses this corresponds to an electric field strength of 10 to 100 V/cm. All these values are too small to initiate electrical breakdown in the polymer film. According to Ref. 20, the breakdown voltage for polymers is no lower than 10^5V/cm.

Studying the temperature dependence of the contact potential difference in a metal–polymer–metal system, we found that the dependence exhibits singularities in the form of potential fluctuations at temperatures close to the transition temperature, while at the melting point the contact potential difference drops sharply (Fig. 7).

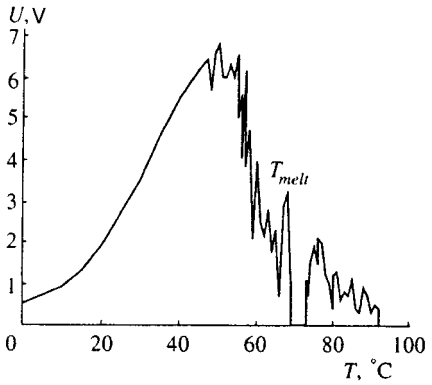

FIG. 7. The temperature dependence of the contact potential difference in the metal–polymer–metal sandwich. Fluctuations near the melting point of the upper electrode (Wood's alloy) are clearly visible.

3. Metallic bridges in the polymer matrix. In analyzing this assumption we discovered that the switching effect depends on the magnitude of the current flowing through the sample. The experimental conditions are favorable to diffusion, so that an increase in current should facilitate further formation of metallic bridges. In our case, however, the effect was just the opposite.

As the current flowing through the sample was increased, at temperatures above the melting point of one of the electrodes the sample resistance was found to drastically increase, with the sample transforming entirely into the low-conductivity state. The inset in Fig. 6 depicts the current–voltage characteristic of a polymer film in the conducting state. The current linearly increases with voltage up to a threshold value I_{th}, after which the conductivity suddenly drops, with the sample transforming into a state with intermediate conductivity. This is reflected by an N-shaped section with negative incremental resistance appearing on the current–voltage characteristic at exactly such voltages. The magnitude of I_{th} increases as the film becomes thinner (up to 2–3 Å).

4. Results of electron-microscope analysis of polymer films. The assumptions concerning the effect of defects, electrical breakdown, and metallic bridges on the discussed phenomena were checked by direct observation of polymer films in a transmission electron microscope.[21] It was found that a polymer film obtained as a result of the formation process is durable and homogeneous. After repeated cyclic transitions of the polymer films into the conducting state and back, no through holes or traces of electrical breakdown were observed, irrespective of how the sample was subjected to external perturbations.

However, the question of thermal diffusion of metal atoms into the polymer and the effect of such atoms on the charge transfer in the polymer film are extremely important, all the more so as it has been suggested, based on indirect results of experiments[22] (the effect of a magnetic field on charge transfer in conducting channels in polymers at extremely low temperatures), that the diffusion of the atoms of the metal electrode into the polymer sample determines all the unusual properties of this system.

It must be noted at this point that diffusion of metal

atoms into a polymer can be achieved, but the conditions are extremely stringent, requiring that the sample be maintained for a long time (five to six hours) at a current that is as close as possible to the threshold value I_{th} (see the inset in Fig. 6) and at a temperature above the metal's melting point, i.e., in the presence of the high-conductivity phase in the sample. Here, due to thermal diffusion of the metal atoms at the points of current flow, there occurs an irreversible process in which the conductivity of the polymer stabilizes at a certain level. As a result, the electrophysical properties of the sample change. In particular, the initial low-conductivity state is not restored after the temperature is lowered, the effect of ''switch-off'' of the high-conductivity state by the critical current (see the inset in Fig. 6) is absent, and the fluctuation instability regions near the melting–crystallization points of the electrodes disappear.

The experiment, however, made it possible to produce electron micrographs of the regions of current flow in the polymer. As a result of studying polymer films subjected to the treatment mentioned earlier with a transmission electron microscope, we obtained micrographs, with typical ones depicted in Fig. 8a. The dark areas correspond to regions of current flow and diffusion of metal atoms into the polymer. If the sample is in the insulator state, no such decoration occurs, irrespective of the applied voltage and the aggregation state of the metal electrode (Fig. 8b).

The size of the conducting regions is 50–250 nm, which agrees well with the estimates given earlier in Refs. 22 and 23 and obtained from the results of indirect measurements involving liquid crystals and the use of the current-spreading method. It is also evident that these regions have an inner structure and consist of smaller entities.

Certainly, the switching to the high-conductivity state initiated by the melting of the upper electrode and the switching caused by an electric field, uniaxial pressure, or thermally stimulated currents have common features. Among these are the metallic temperature dependence of the resistance of a sample in the high-conductivity state, the high conductivity of a pure undoped polymer, the existence of a threshold value for the current flowing through the sample, and the occurence of states with the activation mechanism of conductivity.

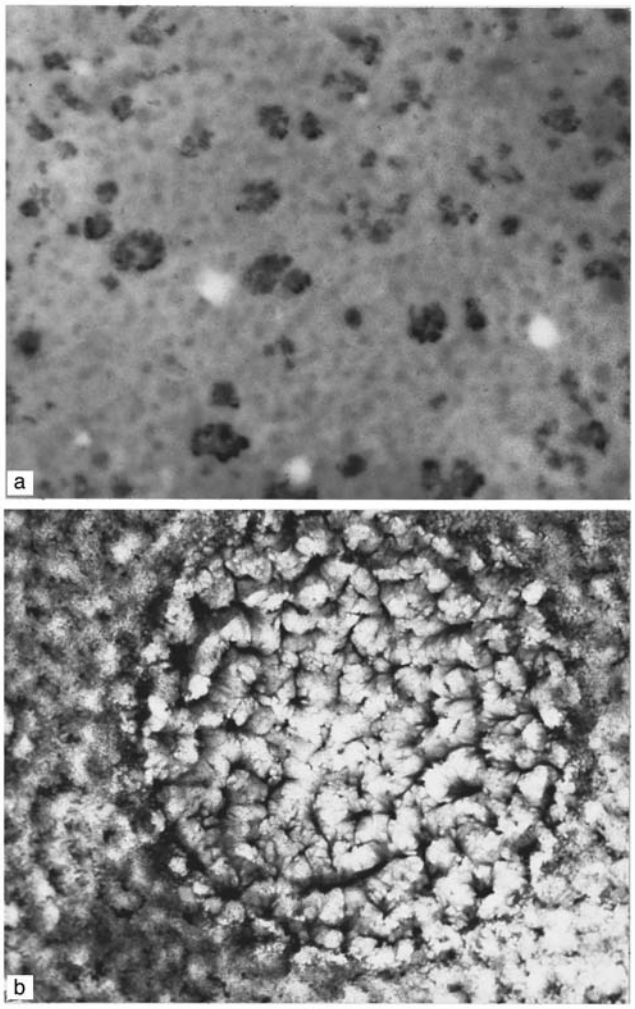

FIG. 8. Electron micrographs of the polymer film. Magnification is 2×10^4. (a) Electron microscope images of a diffusion-decorated electrically conducting film; the dark areas correspond to regions of current flow, and the size of these regions is 50–250 nm. (b) Electron microscope images of the sample in the insulator state. The ''cobblestone'' texture is caused by the support's microprofile.

4. DISCUSSION OF EXPERIMENTAL RESULTS AND CONCLUSIONS

Both the presence of regions of fluctuation excitation of currents in a polymer film at temperature below and above the temperature of the transition to liquid state of one of the electrodes and the fact that the temperature of the transition in the polymer film to the high-conductivity metallic state coincides with the melting point of the electrode are an indication that the surface–surface interaction in the metal–polymer junction has a strong effect on the overall conductivity of the sample.

Figure 9 schematically depicts a surface potential barrier and the related surface charge regions. This case corresponds to $\varphi_m < \varphi_p$ (φ_m is the work function of the metal and φ_p the work function of the polymer), since usually holes are the majority charge carriers in polymers. Under such conditions the electrons go from the metal to the polymer, with the result that a region of negative space charge is formed in the polymer.[24]

Let us examine a metal–polymer junction with allowance for the assumptions characteristic of a Schottky

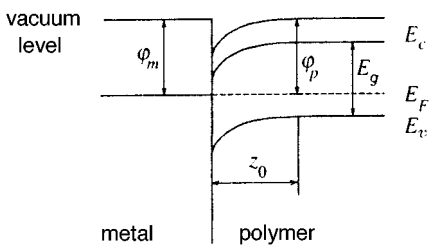

FIG. 9. The surface potential barrier at the metal-polymer interface: φ_p and φ_m are the work functions for the polymer and the metal, respectively; E_g is the polymer's energy gap, E_F is the Fermi level, E_c and E_v fix the bottom of the conduction band and the top of the valence band, and z_0 is the depth of penetration of the space charge.

barrier:[25] the junction between the polymer and the metal is uniform and sharp; there are no surface states; the polymer is uniform up to the interface, at which the band gap E_g changes abruptly; the one-electron approximation is valid, i.e., the electron affinity (χ) is defined as $\chi=\varphi-E_g$, and the contribution of surface dipoles and, accordingly, other effects influencing the absolute values of φ_m, φ_p, and χ do not change the differences of these quantities when the two materials come into contact; and the polymer parameters φ_p, χ, E_g, and z_0 are independent of the position of the Fermi level.

With these restrictions, the height of the potential barrier between the polymer and the metal, φ_b, is given by the following relationship:

$$\varphi_b=\varphi_m-\chi. \tag{3}$$

The thermodynamic equilibrium that sets in when contact between the metal and the polymer is established is related to the evening out of the Fermi levels E_F. The uncompensated charge in the polymer generates a space charge of density eN in a region of size z_0.

The parabolic potential $V_b(z_0)$ in (1) determines the size of the band bending for the carriers at a distance z from the interface. The magnitude of the surface charge, the sign of that charge, and the depth of penetration of the charge into the bulk of the polymer sample depend on the ratio of the work functions of the polymer and the metal, on the acceptor–donor properties of some groups of polymers, and on other parameters of the materials in contact.[26] In this connection we discuss the possible effect of variations in the metal work function on the charge state of the metal–polymer junction, variations that result from changes in the aggregation state of the metal.

Combining Eq. (2) and the results of Ref. 15, where it was shown that the metal work function is temperature-independent but has a singularity at the solid–liquid phase transition point, we find that

$$\begin{aligned} W_{01}&=\varphi_p-\varphi_{m1},\\ W_{02}&=\varphi_p-\varphi_{m2}, \end{aligned} \tag{4}$$

where W_{01} and φ_{m1} are the energy parameters of the metal electrode at temperatures below the phase transition point, and W_{02} and φ_{m2} are the same parameters at the phase transition point. If the corresponding parameters differ, the surface charge density changes by

$$\Delta N=N_1-N_2=\frac{\varepsilon W_{01}}{2\pi(ez_0)^2}-\frac{\varepsilon W_{02}}{2\pi(ez_0)^2}. \tag{5}$$

Combining this with (4), we get

$$\Delta N=N_1-N_2=\frac{\varepsilon\Delta\varphi_m}{2\pi(ez_0)^2}, \tag{6}$$

i.e., the space charge concentration near the surface must experience the largest jump, equal to $\varepsilon\Delta\varphi_m/2\pi(ez_0)^2$, at the phase transition point. Thus, the difference in the metal work functions near the melting point has the greatest effect on the variation of the surface charge density near the metal–polymer junction.

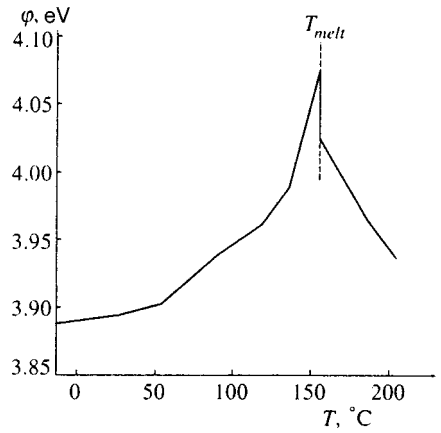

FIG. 10. The temperature dependence of the indium work function near the phase transition point.[15]

To determine the role of the parameter z_0, we use Eq. (2) to estimate z_0. The surface charge density was found by analyzing the capacitance–voltage characteristics for metal–polymer–metal systems and amounted to 10^{-7}–10^{-8}C/cm, so that z_0 was found to be 0.1–1 μm, respectively. Note that here z_0 coincides with the critical thickness of polymer films, $\sim$1–2 μm, above which usually no variations in electrical conductivity are observed. This fact can be interpreted as ''charge instability,'' an effect that emerges in polymer samples whose thickness is close to the depth of penetration of the surface charge induced at the metal–polymer interface. Bearing in mind that there are always two electrodes and, accordingly, two surfaces, the critical thickness can be doubled.

Thus, the entire volume of the polymer lies within the region affected by the field generated by the densest surface part of the space charge. On the other hand, a nonequilibrium charge permeates the sample, while in ''thicker'' samples there is a region outside z_0 in which the effect of the surface charge is less important. The nonequilibrium charge in this region is generated by the ''tails'' of the surface states, with the result that the concentration of this charge is low.

Let us examine the case where the metal (indium) is in contact with the polymer (PPB). Alchagirov *et al.*[15] measured the temperature dependence of the indium work function $\varphi_{\rm In}$ (Fig. 10). Their result shows that $\varphi_{\rm In}$ increases with temperature and at the melting point reaches its maximum value $\sim$4.1 eV. A further increase in temperature results in a sharp drop in the work function of the indium electrode.

The relative variation in $\varphi_{\rm In}$ at the melting point amounts to about 0.2 eV. According to Eq. (6), the relative variation of the surface charge density, $\Delta N/N$, is close to 80%. Obviously, only a fraction of this charge is redistributed in the polymer's bulk at temperatures close to the metal's phase transition point, and its effect depends on the relationship between such quantities as the charge production rate, the effective trapping cross section, the electron–hole recombination rates, and diffusion coefficients.

The recombination and relaxation of the nonequilibrium charge in the polymer can take different routes. Some of these are depicted in Fig. 11. By recombination we mean all

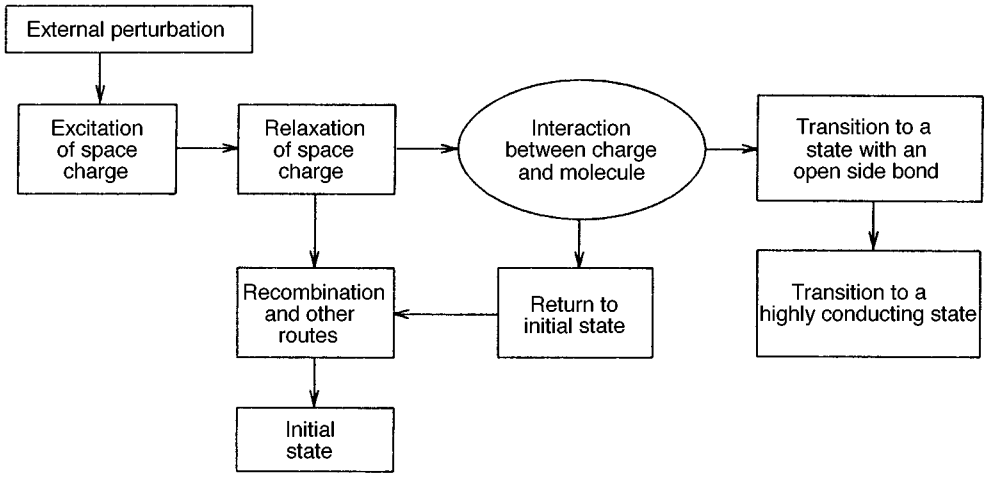

FIG. 11. The diagram illustrating the transition of the polymer film in a metal–polymer–metal sandwich to the high-conductivity state.

the traditional routes by which an equilibrium distribution sets in. But there is also the possibility of a nontraditional route, in which the charge interacting with a polymer molecule transforms the molecule's electron subsystem by opening the C–O bond in a side fragment. In this way the molecule becomes transformed into a new state. The likelihood of such a result of the interaction of a PPB molecule and an elementary charge has been mentioned earlier in the experimental work of Lachinov *et al.*[27] and the theoretical paper of Johansson *et al.*[28] Here the state was characterized as charged, with an uncompensated positive charge on a quaternary carbon atom in the skeleton part of the molecule and a negative charge on the CO_2 group in a side fragment (Fig. 12b). When the relative concentration of such states is so high that the wave functions of neighboring state overlap, a transition to the high-conductivity state probably takes place.

Experiments carried out to establish the mechanism of the interaction of free electrons and compounds that model the electronic state of a monomer chain (negative-ion mass spectrometry) have shown that the transition of the compound to a higher-energy charged state takes place at energies close to those of thermal electrons.[29] This means that the energy of an electron belonging to the cloud of the nonstationary space charge may be sufficient for initiating charged states in a macromolecule by transfering the molecule into another stable state. External excitation (in our case the jump in surface potential) generates this charge. When the concentration of the charge states is high, a polaron lattice with a corresponding half-filled metallic subband may be formed (at least in principle).

One consequence of this model is the dependence of the given phenomenon on the rate of variation of the sample temperature. When the rate is low, the excess charge near the surface has time to relax to an equilibrium distribution, and the critical concentration needed for the formation of a conducting state is not attained.

Figure 13 depicts the temperature dependence of the conductivity of a polymer film with indium as the material for one of the electrodes. The sample is placed in a thermostated cell, and the rate of temperature variation rate amounts to 0.1 °C min^{-1} (curve *1*). No transition to the high-

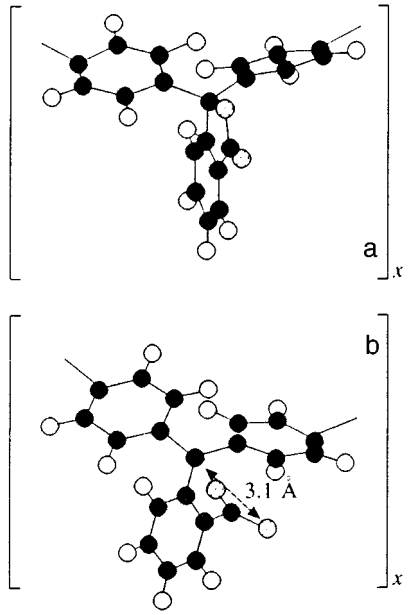

FIG. 12. The energy-stable states of a PPB molecule:[28] (a)—an electroneutral state, and (b)—a state with an open C–O bond in a side fragment.

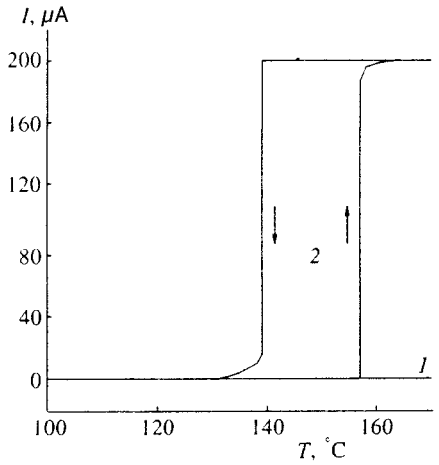

FIG. 13. The temperature dependence of the conductivity of the polymer film for different heating rates: *1* corresponds to 0.1 °C min^{-1}, and *2* corresponds to 5 °C min^{-1}. The arrows indicate the direction in which the temperature changes.

conductivity state occurs. The sample goes into the high-conductivity state when the heating rate is roughly to 5 °C min^{-1} (curve *2*).

This conclusion is indirectly corroborated by the fact that thin films of polymers belonging to this class exhibit thermally stimulated switching to the high-conductivity state,[30] an effect that occurs when the space charge concentration reaches a certain critical value because of thermal ionization of traps within the polymer's bulk at high enough heating rates. At low heating rates there is no switching to the high-conductivity state. The appearance of the high-conductivity state is related to the formation of a polaron band due to the mutual interaction of electron–phonon states caused by the exchange of virtual phonons.

We suggest interpreting the fluctuations in the current (Fig. 4) and the contact potential difference (Fig. 7) from the viewpoint of the phenomena of pre-melting and post-melting. Ubbelohde[31] was probably the first to suggest that there may be a region preceding melting, a pre-melting region. Frenkel[32] proposed a theory of cooperative heterophase fluctuations. By now, from the numerous studies of phase transitions of the melting type (solid–liquid), it is evident that all such transitions are accompanied by anomalies in the physical properties, which emerge long before the melting temperature is reached. These effects is observed in melting of insulators,[33] metals,[34,35] and semiconductors.[36]

At temperatures much lower than the melting point of the substance ($\Delta T \sim 20-150$ K), discrepancies in the behavior of the specific heat[37] and thermal diffusivity and conductivity[38] set in, pre-melting regions appear near impurities,[35] and there emerges a melted layer, which becomes thicker as the bulk melting point is approached.[39] As shown in Refs. 33, 34, and 36, in this temperature range temperature oscillations of the substance being melted are observed. Here the pre-melting process is of an exothermic fluctuation nature and consists of a set of thermal pulses with a repetition frequency ranging from 0.05 to 0.45 Hz. A comparison of these results with the data depicted in Fig. 4 shows that the registered current fluctuations in the metal–polymer–metal system can be explained by the existence of a pre-melting effect in the metal used as an electrode. According to current ideas on pre-melting, the defects on the metal surface — and even more so, the contact with polymer molecules — enhance this effect, broadening the temperature interval for pre-melting.[35]

Pre-melting effects strongly influence the energy structure of the metal–polymer junction and, in particular, the size of the surface charge. The fluctuations of this charge in the polymer are registered as current fluctuations. This conclusion is based on a well-known relationship,[40] in which the work function depends on the energy of the interaction of atoms and the ionization potential, i.e., on the parameters that vary in the pre-melting region. Thus, in the case of films whose thickness is of the order of the depth of penetration of the surface charge into the sample's volume, the fluctuations of the surface charge are actually fluctuations of the space charge. They ''follow'' the energy variations in the metal electrode and manifest themselves in the form of current fluctuations in the measuring circuit.

Let us examine the oscillations of the contact potential difference, which are observed in our experimental structure near the temperature of the phase transition of the metal electrode (Fig. 7). As the temperature rises, the contact potential difference increases according to an exponential law, which agrees with its nature. But when the temperature reaches 45–47 °C, this increase stops and low-frequency (0.01–0.005 Hz) fluctuations of the contact potential difference appear in the experimental structure. When the melting point of the metal is reached, the contact potential difference becomes minimal and low-frequency oscillations again occur in the system, with the oscillations registered in the 20–30 °C range above the melting point.

The curves in Figs. 3, 4, and 7 are reproducible provided that the preliminary preparation of the sample takes place under identical conditions: equal initial temperatures and equal times of exposure to the initial temperature, with the latter exceeding the time of charge relaxation in the metal–polymer–metal system. According to our estimates, this time must be at least six hours, which points to the relaxation nature of the phenomenon. If for some reason the polymer material was initially in the high-conductivity state, the dependence of the contact potential difference on the temperature in the phase transition region exhibits no special features.

Comparison of the results depicted in Figs. 3 (curve *1*), 4, and 7 shows that the temperatures and temperature intervals coincide for the pre-melting regions, the melting point of the metal, and the post-melting regions. The deviation of the temperature dependence of the contact potential difference from the exponential dependence and the presence of other special temperature features are probably a consequence of the excitation in the polymer film of a nonequilibrium space charge under considerable variations of the boundary conditions at the metal–polymer interface in temperature ranges corresponding to pre- and post-melting phenomena and near the melting point of the metal. This, in particular, does not contradict the fact (mentioned earlier by Kokorin and Chernenko[41]) of generation of a pulse of thermoelectric power in the melting of metals.

Summing up the results of the present work, we should mention the following. The proposed interpretation of the mechanism describing the transition to the high-conductivity state suggests that such transitions follow different routes. An external energy perturbation disrupts the equilibrium charge state in the surface layer of the polymer. The metal–polymer interface is convenient in that the large work function makes it possible to generate a high-density surface charge. The transition in the polymer occurs as a result of a specific interaction between the macromolecules and the cloud of nonequilibrium charge. Such interaction leads to the emergence of charged states in the macromolecules and in a certain sense is similar to self-doping. The subsequent overlap of the wave functions of the new states, when the concentration of these states reaches a critical value, leads to the polymer's transition to the high-conductivity state.

Many thanks go to V. S. Korsakov for his help in measuring the surface charge density. The present work was sup-

ported financially by the Russian Fund for Fundamental Research (Project No. 96-02-19208a).

[1] N. S. Enikolopyan, L. N. Grigorov, and S. G. Smirnova, JETP Lett. **49**, 371 (1989).
[2] V. M. Arkhangorodskiĭ, E. G. Guk, A. M. El'yashevich, A. N. Ionov, V. M. Tuchkevich, and I. S. Shlimak, Dokl. Akad. Nauk SSSR **309**, 634 (1989) [Sov. Phys. Doklady **34**, 1016 (1989)].
[3] A. N. Lachinov, A. Yu. Zherebov, and V. M. Kornilov, JETP Lett. **52**, 103 (1990).
[4] A. N. Lachinov, A. Yu. Zherebov, and V. M. Kornilov, Zh. Éksp. Teor. Fiz. **102**, 187 (1992). [Sov. Phys. JETP **75**, 99 (1992)].
[5] A. N. Lachinov, A. Yu. Zherebov, and V. M. Kornilov, Synth. Met. **44**, 111 (1991).
[6] A. N. Lachinov and A. Yu. Zherebov, Synth. Met. **44**, 99 (1991).
[7] A. Brillante, M. Hanfland, K. Syassen, and J. Hocker, Physica B **139**/**140**, 533 (1986).
[8] A. N. Lachinov, A. Yu. Zherebov, and M. G. Zolotukhin, Synth. Met. **59**, 377 (1993).
[9] A. Yu. Zherebov and A. N. Lachinov, Synth. Met. **46**, 181 (1992).
[10] V. M. Kornilov and A. N. Lachinov, JETP Lett. **61**, 921 (1995).
[11] T. Bechstedt and R. Enderlein, *Semiconductor Surfaces and Interfaces. Their Atomic and Electronic Structures*, Academie-Verlag, Berlin (1988).
[12] S. A. Nenakhov, G. I. Shcherbina, A. E. Chalykh, and V. M. Muller, Poverkhnost' No. 3, 77 (1994).
[13] A. G. Lipson, E. V. Kuznetsova, D. M. Sakov, and Yu. P. Toporov, Poverkhnost' No. 12, 74 (1992).
[14] B. B. Alchagirov, Kh. B. Khokonov, and R. Kh. Arkhestov, Dokl. Akad. Nauk SSSR **326**, 121 (1992).
[15] B. B. Alchagirov, Kh. Kh. Kalazhokov, and Kh. B. Khokonov, Izv. Akad. Nauk SSSR, Ser. Fiz. **55**, 2463 (1991).
[16] M. G. Zolotukhin, A. A. Panasenko, V. S. Sultanova *et al.*, Macromol. Chem. **186**, 1747 (1985).
[17] J. R. Rasmusson, Th. Kugler, R. Erlandsson *et al.*, Synth. Met. **76**, 195 (1996).
[18] G. A. N. Connell, D. L. Camphausen, and W. Paul, Philos. Mag. **26**, 541 (1972).
[19] K. Kao and W. Hwang, *Electrical Transport in Solids*, Pergamon Press, Oxford (1981).
[20] G. I. Skanavi, *Physics of Insulators* [in Russian], Scientific and Technical Press, Moscow (1958).
[21] V. M. Kornilov and A. N. Lachinov, Synth. Met. **53**, 71 (1992).
[22] A. M. El'yashevich, A. N. Ionov, M. M. Rivkin, and V. M. Tuchkevich, Fiz. Tverd. Tela (Leningrad) **34**, 457 (1992) [Sov. Phys. Solid State **34**, 1850 (1992)].
[23] O. A. Skaldin, A. Yu. Zherebov, A. N. Lachinov, A. N. Chuvyrov, and V. A. Delev, JETP Lett. **51**, 159 (1990).
[24] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed., Wiley, New York (1981).
[25] W. Schottky, Z. Phys. B **113**, 367 (1939).
[26] M. Logdlund, P. Dannetun, C. Fredriksson *et al.*, Synth. Met. **67**, 141 (1994).
[27] A. N. Lachinov, M. G. Zolotukhin, A. Yu. Zherebov, S. N. Salazkin, A. N. Chuvyrov, and I. L. Valeeva, JETP Lett. **44**, 349 (1986).
[28] N. Johansson, A. Lachinov, S. Stafstrom *et al.*, Synth. Met. **67**, 319 (1994).
[29] B. G. Zykov, Yu. V. Vasil'ev, V. S. Fal'ko, A. N. Lachinov, V. I. Khvostenko, and N. G. Gileva, JETP Lett. **64**, 439 (1996).
[30] A. N. Lachinov and A. Yu. Zherebov, Synth. Met. **55**/**1**, 530 (1993).
[31] A. R. Ubbelohde, Trans. Faraday Society **34**, 29 (1938).
[32] Ya. I. Frenkel, *Kinetic Theory of Liquids*, Dover, New York (1955).
[33] L. A. Bityutskaya and E. S. Mashkina, Pis'ma Zh. Tekh. Fiz. **21**, 85 (1995) [Tech. Phys. Lett. **21**, 763 (1995)].
[34] L. A. Bityutskaya and E. S. Mashkina, Pis'ma Zh. Tekh. Fiz. **21**, 90 (1995) [Tech. Phys. Lett. **21**, 1032 (1995)].
[35] E. A. Stern and Ke Zhang, Phys. Rev. Lett. **60**, 1872 (1988).
[36] L. A. Bityutskaya and E. S. Mashkina, Pis'ma Zh. Tekh. Fiz. **21**, 8 (1995) [Tech. Phys. Lett. **21**, 828 (1995)].
[37] E. B. Amitin, Yu. F. Minenkov, O. A. Nabutovskaya, and I. E. Paukov, Zh. Éksp. Teor. Fiz. **89**, 2092 (1985) [Sov. Phys. JETP **62**, 1207 (1985)].
[38] V. E. Zinov'ev, V. F. Polev, S. G. Taluts, and P. V. Gel'd, Fiz. Tverd. Tela (Leningrad) **28**, 2914 (1986) [Sov. Phys. Solid State **28**, 1639 (1986)].
[39] E. Theis and K. Horn, Phys. Rev. B **51**, 7157 (1995).
[40] E. M. Savitskiı, I. V. Burov, and L. N. Litvak, Dokl. Akad. Nauk SSSR **218**, 818 (1974).
[41] V. V. Kokorin and V. A. Chernenko, Fiz. Met. Metalloved. **61**, 1224 (1986).

Translated by Eugene Yankovsky

Announcement: New Publications from TVP scientific publishers

Zh. Éksp. Teor. Fiz. **111**, 1530 (April 1997)

[S1063-7761(97)02904-1]

V. I. Vozyakov, V. S. Kapitonov, and V. N. Popov, *Collective Effects in Quantum Crystallizing Systems [in Russian]*, (Series on Progress in Theoretical Physics, Vol. 1), viii +200 pp. Print run: 300. Price: 66000 rubles.

Functional (continuous) integration provides a universal approach to the study of various kinds of physical systems. In quantum statistical mechanics it does not so much compete with the operator approach, as supplement it and make it more intuitive, especially in studies of collective perturbations and phase transitions and in renormalizing perturbation theory. In this book the applications of this method to phase transitions and collective perturbations in solids are discussed taking the crystal structure into account. Interest in this topic is currently stimulated by the search for the key mechanism of high temperature superconductivity.

This book, written by leading experts in statistical physics and condensed matter physics, is intended for scientists and students in the corresponding specialties.

From the Table of Contents: *Ch*. 1. Functional integrals in statistical physics. *Ch*. 2 Coulomb crystals. *Ch*. 3. Quantum crystallization in Bose systems. *Ch*. 4. Quantum crystallization in Fermi systems. *Ch*. 5. The funnctional approach to the theory of the electronic properties of adsorbed films with a regular structure. *Ch*. 6. Spin systems. *Ch*. 7. The Hubbard model and high temperature superconductivity.

A. Ya. Khinchin, *Selected Papers on Probability Theory [in Russian]*, Vol. 1, compiled by B. V. Gnedenko and edited by A. M. Zubkin, xi+552 pages. Print run: 350. Price: 180500 rubles.

The principal aim of this book is to reflect the remarkable contributions of A. Ya. Khinchin to twentieth century mathematics and, especially, to probability theory and statistical physics. The influence of his creativity on the modern theory of probability and his role as one of the leading founders of this theory have, by now, been widely recognized. In the meantime, many of his brilliant results and innovative ideas have been inaccessible to mathematicians and theoretical physicists because of the absence of translations or of old journals and books in many libraries. Now this lack has been corrected by this thorough and almost exhaustive collection of his work on probability theory. Published in 1923–1956 in Italian, Russian, German, and French, these papers have been carefully collected and prepared for publication by his student, B. V. Gnedenko. Many of them are of both historical and methodological interest. This collection of papers is being published for the first time.

Place orders at one of the following addresses: 117966, Moscow GSP-1, ul. Vavilova, 42 (tel.: 095 135 2380, FAX: 095 113 5125)

117418, Moscow, ul. Krasikova, 32 19216 (tel.: 095 332 4435, FAX: 095 125 2131)

