Onset of long-range ferromagnetic order in a system of ferromagnetic particles with giant magnetic moments

G. A. Takzeĭ*), L. P. Gun'ko, I. I. Sych, A. B. Surzhenko, S. V. Cherepov, and Yu. N. Troshchenkov

Institute of Magnetism, National Academy of Sciences of Ukraine and Ministry of Education of Ukraine, 252680 Kiev, Ukraine

I. Mirebeau

Laboratoire Léon Brillouin (CEA-CNRS), CEA-Saclay, 91191 Gif-sur-Yvette Cédex, France
(Submitted 23 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 1848–1858 (November 1998)

Magnetic, x-ray, and small-angle neutron scattering data obtained for the decomposed alloy $Cu_{64}Mn_9Al_{27}$ are used to show that the onset of long-range ferromagnetic order in a system of small superparamagnetic grains dissolved in a nonmagnetic matrix is attributable to cooperative ordering of their magnetic moments. © *1998 American Institute of Physics.* [S1063-7761(98)02211-2]

1. INTRODUCTION

In a model with localized magnetic moments, the onset of magnetically ordered states, an dferromagnetic states in particular, in metals and alloys of 3*d* transition metals is customarily linked to spin ordering of the atomic magnetic moments, whose magnitudes do not exceed a few Bohr magnetons μ_B (Ref. 1). It would be interesting to determine whether the onset of long-range ferromagnetic order is possible in systems of magnetic moments ranging from several hundred to several thousand μ_B in magnitude. To date, however, we have not had an altogether clear picture of this problem. For example, in a dense system of very small superparamagnetic iron grains of diameter $d \approx 30-50$ Å dissolved in an aluminum matrix, it is impossible to detect the onset of long-range ferromagnetic order at any temperature by neutron techniques.[2] On the other hand, drawing upon magnetic investigations, Kokorin *et al.*[3] have submitted qualitative arguments to suggest that such ordering is possible in decomposing alloys, in which a dense system of small ferromagnetic grains of diameter $d \approx 100$ Å is formed with magnetic moments $\mu \approx (10^3-10^4)\mu_B$.

In this paper, based on an investigation of critical neutron scattering in conjunction with magnetic and structural studies, the stated problem is solved for the decomposing alloy $Cu_{64}Mn_9Al_{27}$.

2. EXPERIMENTAL PROCEDURE

The alloy $Cu_{64}Mn_9Al_{27}$ was melted out in an induction furnace in a purified nitrogen atmosphere from raw components of at least 99.9% purity. X-ray fluorescence analysis showed that the chemical composition of the alloy did not deviate more than 0.3 at. % from the nominal for each component. A mutual inductance bridge was used to measure the real part χ' and imaginary part χ'' of the dynamic susceptibility at temperatures 4.2–300 K, and a vibrating-sample magnetometer was used to measure the static magnetization. X-ray examinations were performed by a photographic technique at room temperature in a rocking-sample x-ray chamber using Cu K_α radiation. All the experiments were carried out at the Institute of Magnetism in Kiev, Ukraine. Small-angle neutron scattering investigations at temperatures 10–300 K in the interval of scattering vectors $0.009027 \leqslant q \leqslant 0.07945$ Å were conducted on the PAXE facility at the Leon Brillouin Laboratory in Saclay, France.

3. EXPERIMENTAL RESULTS AND ANALYSIS

3.1. Magnetic and Structural Characteristics of the alloy $Cu_{64}Mn_9Al_{27}$

We first analyze the magnetic and structural characteristics of the investigated alloy under various heat-treatment schedules. According to x-ray data obtained in the present study, the alloy $Cu_{64}Mn_9Al_{27}$, heated to 1050 K and quenched in ice water, is a homogeneous solid solution with a crystal structure of the Cu_3Al type (lattice parameter $a = 2.986$ Å). The manganese atoms were distributed in a disordered pattern in the matrix of the alloy. Consequently, the presence of indirect Ruderman–Kittel–Kasuya–Yosida (RKKY) exchange between manganese atoms in the alloy can be expected to induce a spin-glass state of the kind found, for example, in classical spin glasses.[4]

Dynamic Susceptibility

It follows from Fig. 1a, which shows the temperature dependence of the real part χ' of the dynamic susceptibility of the alloy in the annealed state, a characteristic maximum is discerned at the freezing temperature $T_G = 40.2$ K, indicating that the alloy does in fact undergo transition from the paramagnetic to the spin-glass state as it cools. We note in passing that the paramagnetic Curie temperature $\theta_p = 4.5$ K in the Curie–Weiss law for the given alloy is very low (Fig. 1a), implying that the contributions of ferromagnetic and an-

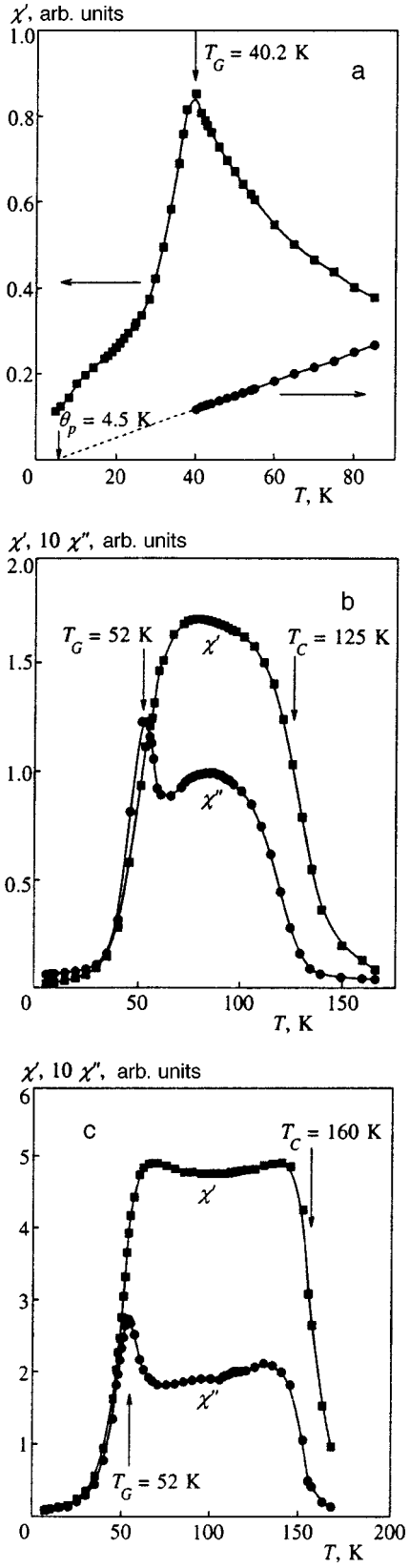


FIG. 1. Temperature dependence of the real (χ') and imaginary (χ'') parts of the dynamic susceptibility of the alloy $Cu_{64}Mn_9Al_{27}$ in the quenched state (a) and after annealing at $T_{an}=373$ K for 2 h (b) and 5 h (c). As a visual aid, the much smaller variable χ'' is plotted on a scale 10 times larger than χ' in each case (b, c). In Fig. 1a the paramagnetic Curie temperature θ_p in the Curie–Weiss law is obtained by extrapolating the temperature curve as $(\chi')^{-1}\rightarrow 0$.

tiferromagnetic exchange interactions to the total exchange energy of the alloy are approximately equal. This property is very typical of dilute classical spin glasses.[4]

It should be emphasized that the magnetic characteristics of the investigated alloy are altered considerably by isothermal annealing at a temperature $T_{an}=373$ K for various durations t_{an}. After a 20-min anneal the alloy is still a spin glass with a freezing point $T_G=52$ K. After a longer anneal, however, decomposition causes the alloy to acquire the properties of ferromagnets at the Curie temperature T_C and then at lower temperatures T_G to undergo transition to a reentrant spin-glass state (Figs. 1b and 1c). We note that T_C increases considerably as the anneal time is increased. The temperature T_G remains essentially constant in this case.

The experiments reported here show that long-range ferromagnetic order sets in after the alloy is annealed with isothermal keeping times $T_{an}\geqslant 1.4$ h. All the experimental results described below refer to the alloy $Cu_{64}Mn_9Al_{27}$ annealed for five hours.

Diffuse X-Ray Scattering

We have used a diffuse x-ray scattering technique to investigate the processes involved in anneal-induced decomposition of the investigated alloy. The corresponding results are shown in Fig. 2, in which clearly two pairs of satellites are observed near the Bragg (110) reflection (Fig. 2a), and one pair is observed near the (200) reflection (Fig. 2b). This diffuse scattering pattern is typical of isomorphically decomposing solid solutions, whose matrix acquires an ensemble of coherent, equiaxed grains of the precipitated phase. Functioning as dilatation centers, these grains are disposed uniformly throughout the anisotropic elastic matrix. The number of satellites as a function of the Miller indices of the Bragg reflections indicates that the precipitated grains are distributed in the matrix of the alloy more or less regularly along the crystallographic $\langle 100\rangle$ direction.[5]

Using, for example, the diffuse reflection pattern near the Bragg (200) reflection (Fig. 2b), we can estimate the average distance D between the centers of the anneal-precipitated grains. According to Ref. 6, D corresponds to a periodic variation of the lattice constant of the matrix and can be estimated from the expression

$$D=a\frac{h\tan\theta}{(h^2+k^2+l^2)\,\delta\theta_1}, \tag{1}$$

where a is the lattice constant of the matrix, h, k, l are the Miller indices, θ is the Bragg angle (hkl) of the reflection, and $\delta\theta_1$ is the angular separation of the satellite centers. We can also estimate the average diameter d of the precipitated grains. According to Ref. 5,

$$d=\frac{\lambda}{2\,\delta\theta_2\cos\theta}, \tag{2}$$

where λ is the x-ray wavelength, and $\delta\theta_2$ is the angular width of the pair of satellites in reciprocal space. Calculations of D and d are summarized in Table I, in which the volume fraction and concentration N of precipitated grains

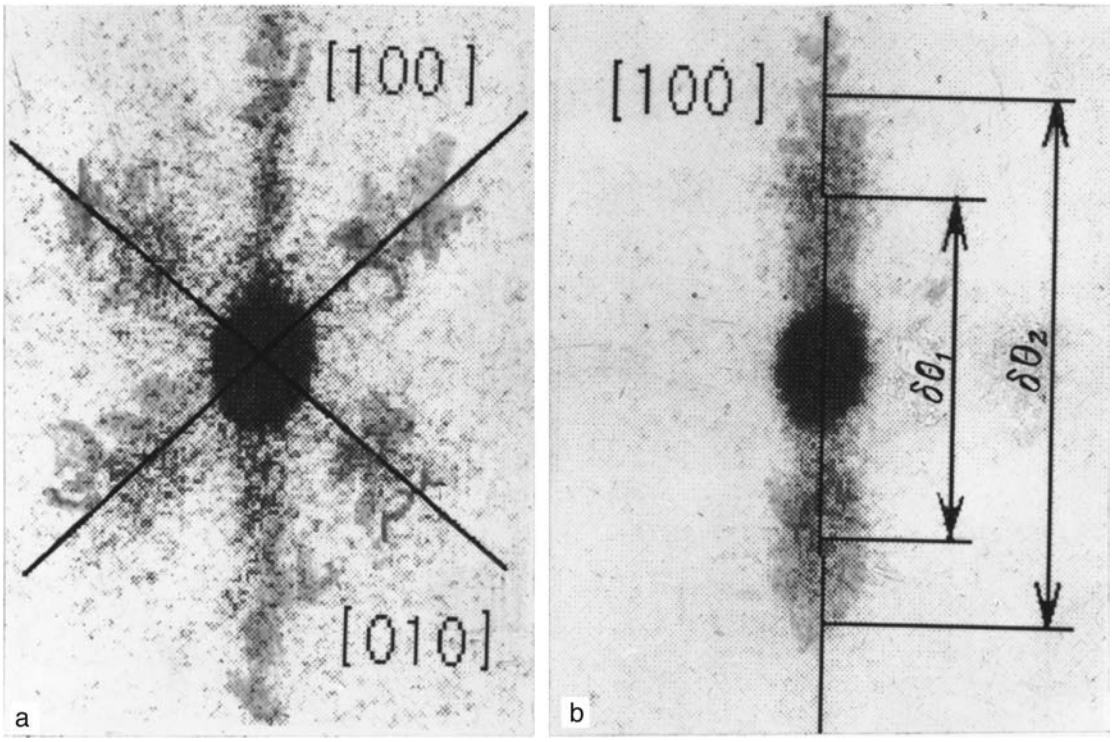


FIG. 2. Diffuse x-ray scattering pattern for the decomposing alloy $Cu_{64}Mn_9Al_{27}$ ($T_{an}=373$ K, $t_{an}=5$ h) near the Bragg (110) (a) and (200) (b) reflections.

are also calculated. Assuming that the grains formed in decomposition of the alloy are of stoichiometric composition Cu_2MnAl (Heusler alloy) and knowing the lattice constant $a=5.971$ Å, the number of Mn atoms per unit cell $n=4$, and the magnetic moment of the manganese atom $\mu=4\mu_B$ (Ref. 7) for the newly formed phase from x-ray data, we can readily find the average magnetic moment of a precipitated grain (see Table I).

Static magnetization

Investigations of the magnetization M at room temperature in magnetic fields up to 20 kOe reveal typical superparamagnetic behavior of the investigated alloy (Fig. 3). The anisotropy constant of the bulk Cu_2MnAl alloy at room temperature is indeed $K=10^3$ erg/cm^3 (Ref. 8). Hence, for the magnetic energy of an isolated grain of average diameter $d=30$ Å we obtain $KV/k_B=0.16\,\mathrm{K}\ll 300\,\mathrm{K}$, so that the condition for superparamagnetic behavior, $KV\ll k_BT$ (Ref. 1), is easily satisfied in the given situation. In the above estimates V is the grain volume.

TABLE I. Structural and Magnetic Characteristics of the Decomposed Alloy $Cu_{64}Mn_9Al_{27}$ ($T_{an}=373$ K, $t_{an}=5$ h).

Parameter	X-ray data	Magnetic data
Average grain diameter d, Å	30 ± 3	32 ± 3
Average distance between grain centers D, Å	47 ± 3	51
Average magnetic moment of grain μ, μ_B	1100 ± 320	1140 ± 150
Concentration of grains N, 10^{-8} cm^{-3}	8.84	7.34
Volume fraction of grains	0.136	0.126

The field dependence of the static magnetization M of an ensemble of noninteracting superparamagnetic grains must satisfy the Langevin equation when $H_A\ll H$, where H_A is the anisotropy field of an isolated grain; in the weak and strong magnetic field limits this equation can be written in the form[1]

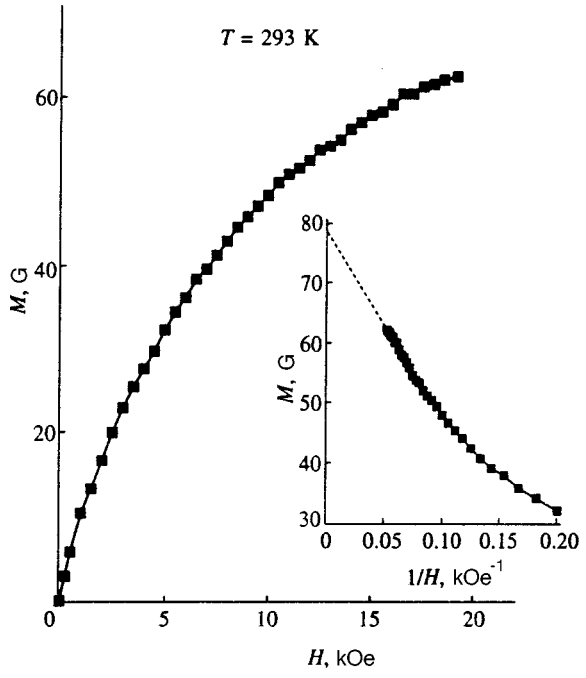


FIG. 3. Static magnetization M of the decomposed alloy $Cu_{64}Mn_9Al_{27}$ ($T_{an}=373$ K, $t_{an}=5$ h) at $T=293$ K.

$$M=\frac{N\mu^2 H}{2k_BT} \quad \text{for} \quad \frac{\mu H}{k_BT}\ll 1, \tag{3}$$

$$M=N\mu\left(1-\frac{k_BT}{\mu H}\right) \quad \text{for} \quad \frac{\mu H}{k_BT}\gg 1. \tag{4}$$

Assuming that the magnetic moment of the Mn atom in the grains is equal to the magnetic moment of the Mn atom in the bulk Cu_2MnAl compound, corresponding to the spontaneous magnetization $M_s=500$ G (Ref. 7), we can readily calculate the magnetic moment μ, diameter d, and concentration N of the grains. The results (see Table I) are in very good agreement with the values determined from x-ray data.

Summarizing this section, we conclude that the alloy aged at the temperature $T_{an}=373$ K for the time $t_{an}=5$ h is characterized by the onset of ferromagnetic ordering at $T_C=160$ K and transition to the reentrant spin-glass state at temperatures $T\leqslant T_G=51$ K, as is clearly evident from the dynamic susceptibility data (Fig. 1c). At room temperature $T=300\,\text{K}>T_C$ the alloy is a typical superparamagnet, where the magnetic moments, diameters, concentration, and spacing of the precipitated grains, calculated from magnetic and x-ray data, are in good agreement.

We have conducted small-angle neutron scattering investigations to analyze the nature of the transition from the superparamagnetic to the ferromagnetic state. This work is discussed in detail below.

3.2. Small-Angle Neutron Scattering

Figure 4 shows the temperature dependence of the intensity of small-angle neutron scattering for the alloy $Cu_{64}Mn_9Al_{27}$ after annealing for $t_{an}=5$ h and various neutron scattering vectors (the neutron wavelength is $\lambda=8$ Å). The data reveal that for small scattering vectors a critical neutron scattering peak is distinctly visible at $T_C=160$ K. Below T_C the intensity decreases and then begins to increase again; a second, low-temperature peak is observed for extremely small scattering vectors, which is very typical of reentrant spin glasses (see, e.g., the small-angle neutron scattering data for $Fe_{1-x}Al_x$ in Ref. 9 or for $(Fe_xNi_{1-x})_{80}P_{20}$ in Ref. 10). The intensity I of this scattering at temperatures below T_C exhibits a rather complex dependence on T and q. In the present study, however, we are mainly interested in the processes by which long-range ferromagnetic order is established in the decomposed alloy $Cu_{64}Mn_9Al_{27}$ in the vicinity of T_C.

Since long-range ferromagnetic order is established in the investigated alloy only in the decomposed state, as shown above, it is reasonable to assume that the carriers of magnetism in the given situation are not the magnetic moments of individual Mn atoms, but the magnetic moments of precipitated grains of the Cu_2MnAl phase with effective magnetic moment $\mu\approx 10^3\mu_B$.

In this regard we discuss certain aspects of the critical neutron scattering pattern. According to Ref. 2, the neutron scattering cross section $S(q)$ for interacting magnetic grains, being proportional to the small-angle neutron scattering intensity, can be written as the product of two functions, corresponding to the in-grain and intergranular correlations:

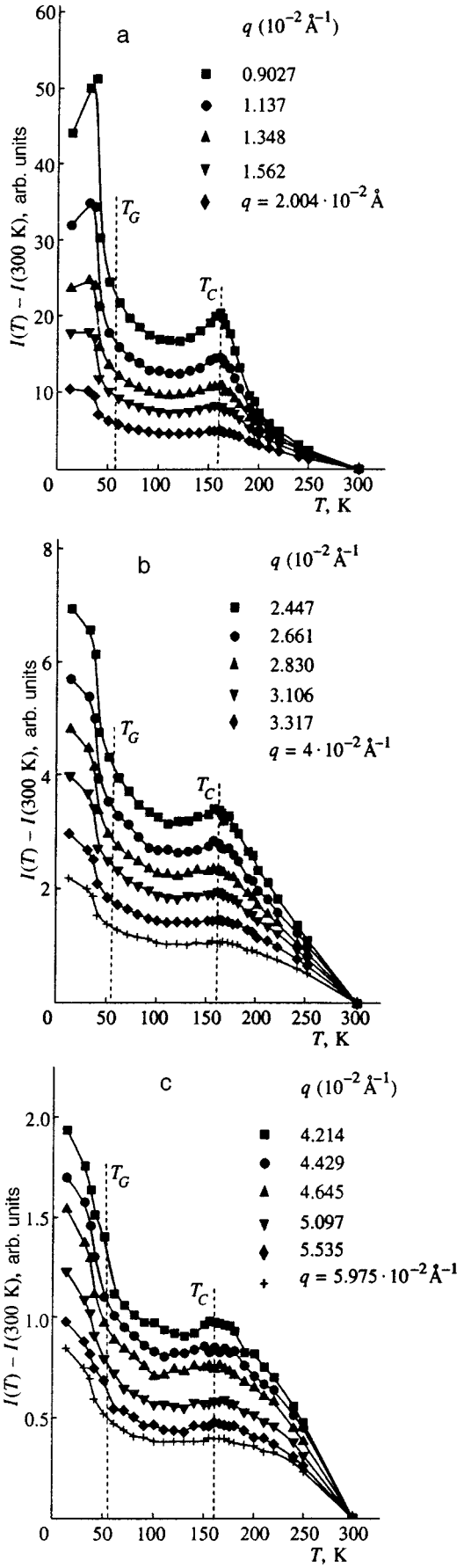


FIG. 4. Temperature dependence of the small-angle neutron scattering intensity I for the alloy $Cu_{64}Mn_9Al_{27}$ ($T_{an}=373$ K, $t_{an}=5$ h) at various values of the scattering vector q (Å^{-1}).

$$S(q) \propto \mu^2 F^2(q, R_1) I(q, R_2),$$

where μ is the magnetic moment of the grain, $F(q, R_1)$ is its magnetic form factor, and $I(q, R_2)$ is a function associated with intergranular cooperative fluctuations of the magnetic moment. Here we have neglected critical in-grain magnetization fluctuations, because they occur only at much higher temperatures around $T_C^* \approx 700$ K, as in the bulk Cu_2MnAl alloy, well above the observed $T_C = 160$ K.

The magnetic form factor $F(q, R_1)$, which usually differs very little from the chemical form factor obtained from x-ray studies (see, e.g., Ref. 2), includes the characteristic correlation length R_1. At a lower temperature R_1 should increase slightly to a value close to the grain diameter $d = 30$ Å without acquiring anomalies at T_C. On the other hand, the intergranular function $I(q, R_2)$ can be expressed in terms of the Ornstein-Zernike function (see below) and depends on the intergranular correlation length R_2, which diverges as $T \to T_C^+$, according to Ref. 11 and on the assumption that long-range ferromagnetic order is established in the system of superparamagnetic grains. We discuss this issue a little farther along in the article, turning our attention for now to the temperature dependence of small-angle neutron scattering (Fig. 4).

Obviously, when the magnetic moments of the ensemble of precipitated grains in the alloy are ferromagnetically ordered, but the distribution of the intergranular distances and the size distribution of the grains are unknown, the correlation length R_2 of ferromagnetic critical fluctuations near T_C must satisfy the condition $R_2 \geqslant \xi = D - d$, where $\xi \approx 20$ Å is the distance between the surfaces of the precipitated grains along the line joining their centers. In other words, the critical neutron scattering peak in the temperature curves of small-angle neutron scattering intensity near $T_C = 160$ K should be observed only for neutrons with wave vectors $q = \xi^{-1} < 0.05$ Å^{-1}. It is evident from Fig. 4 that this in indeed the situation. For $q > 0.04214$ Å^{-1}, which corresponds to
$\xi < 23.7$ Å, the sharp anomaly in critical neutron scattering in the vicinity of T_C will essentially vanish. The most likely occurrence will be a diffuse anomaly, which occurs because, in reality, the decomposition of the alloy results in the formation of a system of precipitated grains characterized not by the average grain diameters d determined above, or the intergranular distances D, but by certain distributions of these quantities. In our opinion, this consideration accounts for the very weak anomalies in the temperature curves of the small-angle neutron scattering intensity near T_C for neutron scattering vectors 0.04429 Å$^{-1} < q < 0.7945$ Å^{-1}.

As mentioned above, the presence of the critical small-angle neutron scattering peak at T_C for small q is dictated by the divergence of the correlation length R_2 in the function $I(q, R_2)$. Experiment shows that at temperatures $T > T_C$ the function $I(q, R_2)$ is very closely given by

$$I(q, R_2) = I_0 + \frac{A}{q^2 + (1/R_2)^2}, \tag{5}$$

where I_0 is the background intensity, which generally incorporates the contributions of nuclear and magnetic scattering

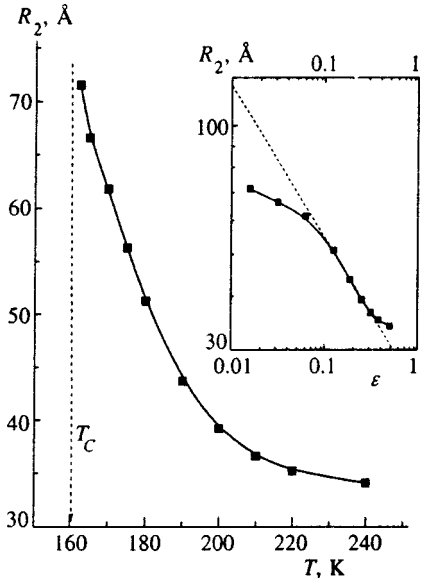


FIG. 5. Temperature dependence of the correlation length R_2 of ferromagnetic critical fluctuations, determined from neutron data in the paramagnetic temperature range for the decomposed alloy $Cu_{64}Mn_9Al_{27}$ ($T_{an} = 373$ K, $t_{an} = 5$ h). Inset: R_2 vs. the normalized temperature $\varepsilon = T/T_C - 1$ on a log–log scale.

by the superparamagnetic grains. The second term represents the well-known Ornstein–Zernike function.[12] Here A is the amplitude, which depends weakly on the temperature. Least-squares processing of the small-angle neutron scattering data according to Eq. (5) shows that the background I_0 is all but independent of the temperature and, hence, is mainly attributable to nuclear scattering. The foregoing procedure has also been used to determine the temperature dependence of the correlation length R_2 (Fig. 5). It follows from the figure that $R_2(T)$ does in fact decrease with the temperature, beginning at $R_2 \approx 34$ Å (comparable to $\xi = 20$ Å) and 240 K, to $R_2 \approx 71$ Å $> D = 50$ Å near T_C.

We have attempted to describe the temperature dependence of R_2 using the relation

$$R_2 \propto \varepsilon^{-\nu}, \tag{6}$$

where $\varepsilon = (T/T_C - 1)$ is the normalized temperature and ν is the critical index of the correlation length. It is evident from the inset to Fig. 5, which shows the corresponding results, that relation (6) is satisfied over a narrow range of ε. Accordingly, the critical index of the correlation length R_2 is $\nu = 0.35 \pm 0.1$, or half the value $\nu = 0.7$ for Heisenberg ferromagnets and some reentrant spin glasses.[13]

The discrepancy can occur for several reasons. The first is purely methodological. For a more rigorous analysis of the experimental data, values of the background intensity I_0 measured at temperatures above 300 K must be used in Eq. (5). However, such measurements are difficult, because the structural and magnetic state of the alloy change very rapidly above room temperature. On the other hand, estimates show that when the background I_0 is included in Eq. (5) at temperatures 160–300 K, R_2 can be determined to within 10–15%. This means that the values obtained for R_2 in the vicinity of T_C are not far from the truth. Then again, relation

(6) follows from a similarity theory[11] developed for ordinary ferromagnets, whereby the distance D between magnetism carriers (spins) must obey the inequality $D \ll R_2$. In the given situation, however, in the vicinity of T_C we have $R_2 \approx 1.5D$ (Fig. 5), which casts doubt on the validity of (6) for processing the experimental results in small-angle neutron scattering.

Summarizing this section, a study of small-angle neutron scattering shows that long-range ferromagnetic order is established in the decomposing alloy as a result of ferromagnetic ordering in the system of macroscopic magnetic moments ($\mu \approx 1100\mu_B$) formed by grains of the stoichiometric Cu_2MnAl phase with average diameters $d = 30$ Å and an average spacing between the grains $D = 50$ Å.

4. DISCUSSION OF THE RESULTS

Here we discuss the possible reasons for the emergence of long-range ferromagnetic order during the cooling of a system of superparamagnetic grains with giant magnetic moments. Estimates in the present study show that, given the parameters found for the system of precipitated particles (see Table I) intergranular dipole interaction yields too low a Curie temperature $T_C \approx (6-90)$ K in comparison with the experimental value. Owing to the quasiregular disposition of the precipitated grains in the matrix of the alloy, indirect RKKY exchange between grains can be expected to play a major role in the evolution of long-range ferromagnetic order in the given situation, as encountered in multilayer magnetic structures of the Co/Cu/Co or Fe/Ag/Fe type.[14] In the latter case the alternating magnetic and nonmagnetic layers have approximately the same thicknesses as the quantities d and ξ found in our work. Once again we must emphasize that the quasiregular disposition of precipitated grains in the $Cu_{64}Mn_9Al_{27}$ matrix is what makes the onset of long-range ferromagnetic order possible. This conclusion is corroborated by the transition from the quasiregular distribution of precipitated grains to a disordered distribution as a result of coalescence processes in the late stages of decomposition of the alloy. The alloy then loses the properties of ferromagnets and acquires those of spin glasses.[5]

It is important to note that only 35% of the manganese atoms participate in the formation of the stoichiometric ferromagnetic Cu_2MnAl phase as a result of decomposition of the alloy investigated here. All other Mn atoms are still dissolved in the nonmagnetic matrix. As emphasized in Ref. 15, the ''lost spins'' can also play a significant part in the formation of long-range ferromagnetic order in heterogeneous magnetic systems.

The above estimates show that near T_C, the ferromagnetic correlation length only slightly exceeds the intergranular spacing. This suggests strong magnetic inhomogeneity of the alloy and the extraordinary character of the processes by which long-range ferromagnetic order is established in it. The dynamic susceptibility data indicate the same result. Indeed, it follows from Fig. 1c that χ'' does not exhibit a sharp anomaly at the Curie temperature $T_C = 160$ K. In contrast, a sharp anomaly of χ'' at T_C is very characteristic of the majority of ordinary ferromagnets and reentrant spin glasses, in which ferromagnetic order evolves in the spin system (see, e.g., Ref. 16). This suggests indirectly that in the given situation, long-range ferromagnetic order is indeed established in the system of precipitated grains of the new phase and not in the spin system.

The authors are grateful to L. Noirez and J. Teixeira for their collaboration in the small-angle neutron scattering measurements at the PAXE spectrometer. One author (Mirebeau) would like to thank M. Hennion for many profitable discussions. Another author (Takzeĭ) is indebted to the Léon Brillouin Laboratory for financial support at the time of the neutron experiments.

*)E-mail: gtakz@guukr.frinet.kiev.ua

[1] S. V. Vonsovskiĭ, *Magnetism*, Vols. 1 and 2, Nauka, Moscow (1979) [transl. of previous edition: Wiley, New York (1974)].
[2] C. Bellouard, I. Mirebeau, and M. Hennion, Phys. Rev. B **53**, 5570 (1996).
[3] V. V. Kokorin and I. A. Osipenko, JETP Lett. **29**, 610 (1979); V. V. Kokorin and I. A. Osipenko, Phys. Status Solidi A **74**, K147 (1982).
[4] K. Binder and A. P. Young, Rev. Mod. Phys. **58**, 801 (1986).
[5] V. V. Kokorin, *Martensitic Transformations in Inhomogeneous Solid Solutions* [in Russian], Naukova Dumka, Kiev (1987).
[6] V. Daniel and H. Lipson, Proc. R. Soc. London, Ser. A **182**, 378 (1943).
[7] R. Bozorth, *Ferromagnetism*, Van Nostrand, New York (1951).
[8] K. Aouagi and M. Sugihara, J. Phys. Soc. Jpn. **16**, 1027 (1961).
[9] H. R. Child, J. Appl. Phys. **52**, 1732 (1981); K. Motoya, S. M. Shapiro, and Y. Muraoka, Phys. Rev. B **28**, 6183 (1983).
[10] M. B. Salamon, A. P. Murani, J. L. Tholence, and J. L. Walter, Phys. Rev. B **33**, 7837 (1986).
[11] H. E. Stanley, *Introduction to Phase Transitions and Critical Phenomena*, Clarendon Press, Oxford (1971).
[12] J. M. Ziman, *Models of Disorder: The Theoretical Physics of Homogeneously Disordered Systems*, Cambridge Univ. Press, Cambridge (1979).
[13] M. Hennion and I. Mirebeau, J. Magn. Magn. Mater. **140–144**, 1565 (1995).
[14] M. Hennion and J. F. Cochran, Adv. Phys. **42**, 523 (1993).
[15] J. C. Slonczewski, J. Appl. Phys. **73**, 5957 (1993).
[16] G. A. Takzeĭ, A. M. Kostyshin, Yu. P. Grebenyuk, and I. I. Sych, Zh. Eksp. Teor. Fiz. **89**, 2181 (1985) [Sov. Phys. JETP **62**, 1259 (1985)].

Translated by James S. Wood

Nonequilibrium paramagnetic susceptibility of gallium impurity centers in lead telluride

A. N. Vasil'ev,*) T. N. Voloshok, and S. V. Kuvshinnikov

M. V. Lomonosov Moscow State University 119899 Moscow, Russia

B. A. Volkov

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 17 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1859–1867 (November 1998)

The temperature dependences of the resistance and magnetic susceptibility are studied in gallium-doped lead telluride, which is characterized by a delayed photoconductivity effect, under various illumination conditions. After a sample is illuminated at low temperatures, the magnetic susceptibility is diamagnetic in the region of metallic delayed conductivity (for $T<50$ K). In the region of thermodynamic equilibrium ($T>70$ K), where conductivity is activational, the magnetic susceptibility is likewise diamagnetic and essentially equals the low-temperature value. A paramagnetic susceptibility peak is observed in the transitional region ($T\sim 50-70$ K), where the conductivity is of a nonequilibrium character but the carriers are still nondegenerate. This peak increases in magnitude with the rate of measurements in the indicated temperature range. In addition, a paramagnetic variation of the susceptibility following the Curie law is observed with uncontrollable (weak) illumination from the cryostat cap at low temperatures ($T<25$ K). The interpretation of the observed dependences is based on notions of variable valence of gallium in lead telluride, while the appearance of a paramagnetic susceptibility peak is attributed to the presence of shallow localized levels of gallium in a trivalent state. © *1998 American Institute of Physics.* [S1063-7761(98)02311-7]

1. Lead telluride doped with group-III elements (Ga, In, ...) is one of the semiconductor materials that at low temperatures exhibit unusual behavior reminiscent of the behavior of III–V semiconductors with *DX* centers. This analogy, in the sense of the physical picture of the phenomenon, will become clear from what follows, if one takes into account the polyvalence of the impurities producing *DX* centers (for example, carbon with a valence of 2 and 4). This includes the delayed monopolar photoconductivity and the variable electrical activity of the corresponding impurities. In the phenomenon of delayed photoconductivity, the conductivity of such semiconductors at low temperatures increases rapidly under illumination, while the low-resistance state arising in the process remains for a long time after the illumination is switched off. The delayed photoconductivity effect in lead telluride is due to the specific properties of gallium (and other group-III elements) as an impurity substituting for lead in this material. The electrical activity of gallium impurity in lead telluride is such that the group-III element Ga, replacing the group-IV element Pb, is a donor, if the chemical potential of the carriers is sufficiently low. The donor action of Ga can be understood simply by assuming that a gallium atom incorporated in the crystal matrix is in a trivalent state. Such a state corresponds to a s^0p^3 configuration of the outer electronic shells of the impurity. The polyvalent behavior of Ga (and group-III elements in general) is well-known in chemistry.[2] In various compounds these elements possess valences from 1 to 3, which corresponds to the electronic configurations of a free atom (or atom in a cubic environment) s^2p^1,s^1p^2, and s^0p^3, respectively. In this sense the donor action of Ga in IV–VI compounds is easily understood: trivalent Ga substituting for a divalent metal furnishes the bands with one extra electron.

The electrical activity of group-III impurities in IV–VI cubic compounds (NaCl structure) can be represented in band language as follows. It is well known[2] that in the tight-binding approximation the electronic spectrum of these compounds is formed by the *s* and *p* orbitals of the metal and chalcogen: the *s* orbitals form two deep completely filled bands, while the *p* orbitals form the actual valence and conduction bands. There are six such bands, according to the number of *p* electrons in the diatomic unit cell. The six *p* electrons (two from the metal and four from the chalcogen) completely fill the three bottom valence bands. The three top bands are empty and form the conduction band. The extrema of these bands are all located at the *L* points of the Brillouin zone of a fcc lattice, where a narrow gap is formed between the valence and conduction bands. It can be shown that at the extrema of the valence bands the Bloch wave functions are constructed from the *p* orbitals of the chalcogen, while the Bloch functions at the extrema of the conduction band are constructed from the *p* orbitals of the metal. This is because the *p* states of the chalcogen lie lower in energy than the *p* states of the metal. In such a situation, substitution of a group-III impurity in a trivalent state for the metal gives rise to an excess electron in the bands; this is what causes the donor action of such an impurity. If this impurity were to replace a metal in an univalent state, then it would be an acceptor.

It is obvious that what state arises is determined by the

difference of the total energies of the corresponding electronic configurations of the impurity incorporated into the crystal matrix. Since the reaction $s^2p^1 - s^0p^3$ requires the excitation of two *s* electrons into *p* states, it is obvious that it can be energetically expedient only when the energy of the single-electron *p* states is sufficiently low. In a crystal the chemical potential in the actual bands plays the role of the energy of the *p* states. It is obvious that the lower the chemical potential, the more easily a trivalent state is formed. For some value of the chemical potential the energies of the configurations s^2p^1 and s^0p^3 can become equal to one another. This value of the chemical potential determines its pinning energy, for which impurities in both the uni- and trivalent states are present simultaneously in the crystal. The specific value of the pinning energy, or equivalently, the energy of the mixed state is determined by the properties of the semiconductor matrix (specifically, the work function) and the specific type of group-III impurity (Ga, In, Tl). For In in PbTe this level in the conduction band, while for Tl it lies in the valence band.[5–9] As follows from the analysis below, the pinning level for Ga in PbTe lies in the gap.

Thus far we have not considered the possibility of Ga being in a divalent state s^1p^2. In nature such a state exists only in metastable form. For example, $GaCl_2$ is unstable with respect to the reaction

$$2GaCl_2 \rightarrow GaCl + GaCl_3.$$

There are no fundamental reasons for believing that such a state cannot be realized in IV–VI compounds. To investigate this possibility it is necessary to measure the magnetic susceptibility, or more precisely, the *g* factor of an electron in a partially filled *s* shell of an impurity. In the s^1p^2 configuration this *g* factor is 2, while the *g* factor of itinerant carriers is of the order of 50. More importantly, magnetic measurements can elucidate the question of the existence of shallow levels associated with the impurity Ga in the trivalent state. The possibility of the appearance of such levels in the case of the present electronic mechanism leading to the electrical activity of group-III impurities was shown in Ref. 10. The point is that when two *s* electrons are transferred into a band, an additional attractive potential arises on the impurity because the screening of the atomic core by the *s* electrons vanishes. This potential has an atomic scale and acts only on electrons located at the bottom of the conduction band.[1)] The latter is due to the fact that, as noted above, the wave functions of only these electrons are different from zero at the metal atoms. Single filling of the shallow levels should result in an additional substantial contribution to the paramagnetic susceptibility, since, as already mentioned, the *g* factor of the electrons bound in these levels is extremely large ($g \sim 50$) because of the smallness of the gap in the band spectrum.

2. The electronic properties of PbTe(Ga) were studied by means of measurements of the resistance and magnetic susceptibility of bulk single crystals. The Ga concentration in the experimental samples was 0.3 at.%, which corresponds to the region of greatest photosensitivity for this compound. The electric spark method was used to cut out samples in the shape of rectangular parallelepipeds. The samples were chemically etched to remove the damaged surface layer. The resistance was measured by a potentiometric method. The magnetic susceptibility was measured by the Faraday method in the region where the magnetization depends linearly on the external field in the range 0.1–1 T. Both the resistance and susceptibility were measured with heating in the temperature range 4–150 K. A miniature incandescent lamp placed near the surface of the sample was used either for brief illumination of the crystal at liquid-helium temperature or for constant illumination of the sample with varying temperature.

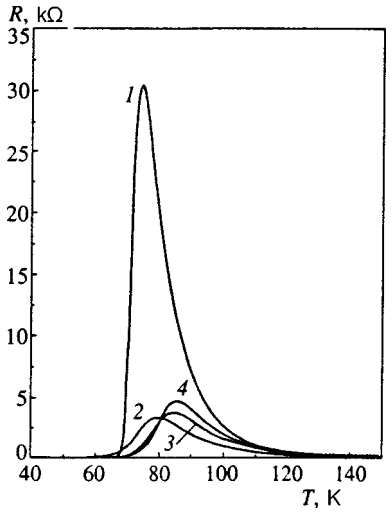


FIG. 1. Temperature dependence of the resistance of PbTe(Ga) obtained under various regimes of illumination by an incandescent lamp. Curve *1* — brief illumination at liquid-helium temperature, curves *2–4* — constant illumination with various (increasing) intensities.

The room-temperature resistivity of a *n*-PbTe(Ga) single crystal is $\rho \sim 10^{-1}\,\Omega \cdot$cm, which corresponds to an electron density of the order of $10^{17} - 10^{18}$ cm^{-3}. The resistivity in a cooled metal chamber screened from illumination increased rapidly, reaching values $\rho \geqslant 10^5\,\Omega \cdot$cm at low temperatures. The carrier activation energy was determined from the slope of the temperature dependence of the resistivity to be $E_a = 66$ meV, which differs considerably from the gap width ($E_g = 220$ meV) in PbTe. After brief illumination at liquid-helium temperatures the resistivity of the sample once again decreased to $\rho \sim 10^{-1}\,\Omega \cdot$cm, which is comparable to the high-temperature value. The low-temperature resistance did not change with time, except for very rapid restoration of a certain portion of the resistance immediately after the illumination was switched off.

Figure 1 shows the temperature dependence of the resistance of the sample obtained with various illumination regimes. Curve *1* corresponds to the following arrangement of the experiment. The sample at liquid-helium temperature was converted into a metallic state by means of illumination. Next, the illumination was switched off and the sample was heated for an hour up to room temperature in a dark chamber. Under these conditions the resistance at first changed little with increasing temperature. Then in the range $T \sim 50 - 60$ K the resistivity increased sharply (by three orders of magnitude) up to several hundreds of $\Omega \cdot$cm, after which the resistance once again decreased in the standard activational

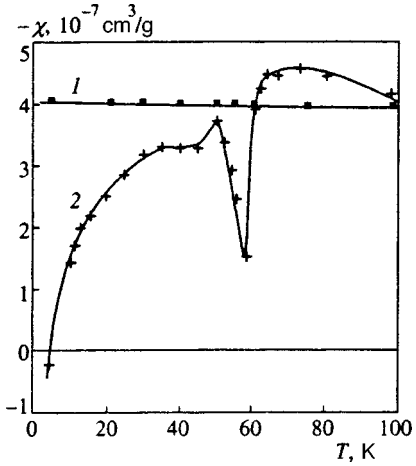


FIG. 2. Temperature dependence of the susceptibility of PdTe (curve *1*) and PbTe(Ga) (curve *2*) under conditions of uncontrollable weak illumination by the cryostat cap.

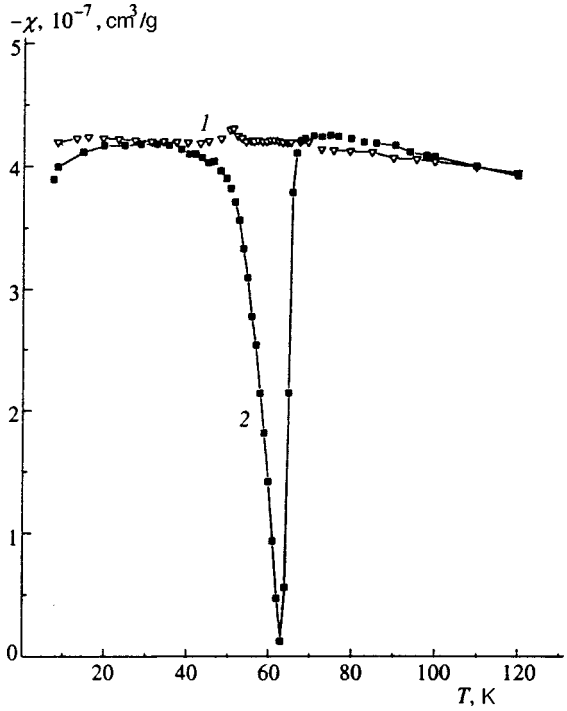


FIG. 3. Temperature dependence of the magnetic susceptibility of PbTe(Ga) obtained under various regimes of illumination by an incandescent lamp. Curve *1* — constant illumination by an incandescent lamp, *2* — brief illumination at liquid-helium temperature.

manner with activation energy E_a. The low-temperature (metastable) part of this curve depended on the heating rate, especially strongly on the low-temperature shoulder of the resistance peak. The high-temperature (activational) part of the curve did not depend on the heating rate. The curves *2–4* represent the results of measurements of the resistance of the sample with constant illumination of various intensities. One can see that under these conditions the resistance peak decreases considerably and shifts to high temperatures as the illumination intensity increases.

As with the resistance measurements, the magnetic susceptibility likewise was measured in two regimes, i.e., either with constant illumination at all temperatures or after brief illumination at liquid-helium temperature. Moreover, the measurements of the susceptibility of PbTe(Ga) were performed under conditions of uncontrollable illumination (the method used to perform the magnetic measurements did not permit isolating the sample completely from illumination by the cryostat cap). The calibration measurements were performed on a sample of undoped PbTe. Figure 2 (curve *1*) shows the results of the calibration measurements. They were completely independent of the illumination of the sample and are identical to the tabulated values of the susceptibility.[11] Curve *2* in this figure describes measurements performed on the PbTe(Ga) sample under uncontrollable illumination conditions. One can see that under conditions when illumination comes only from the cryostat cap the temperature variation of the susceptibility at low temperatures (<25 K) is close to the Curie law ($\chi = C/T$). Note, however, that the Curie constant C measured in these experiments varies from one experiment to another. The paramagnetic peak at $T \sim 60$ K observed in the PbTe(Ga) sample is of greatest interest. This peak grows as the temperature varies more rapidly.

Figure 3 shows the temperature dependences of the magnetic susceptibility of PbTe(Ga) that were obtained in various regimes of illumination by an incandescent lamp. The curve *1* represents the results of measurements of the susceptibility in a regime of continuous illumination by an incandescent lamp. In the presence of constant illumination the diamagnetic susceptibility of PbTe(Ga) decreases slightly with increasing temperature, though a small increase of the diamagnetic susceptibility is observed at ~ 50 K. The curve *2* shows the results of susceptibility measurements after brief preliminary illumination of the sample at liquid-helium temperature. The next heating was conducted in the absence of illumination (except for illumination from the cap). Similarly to measurements performed under conditions of uncontrollable illumination, here a paramagnetic peak, whose position and magnitude depend on the rate of heating, is likewise observed at $T \sim 60$ K. Comparing the measurements of the resistance (Fig. 1) and magnetic susceptbility (Figs. 2 and 3) of the PbTe (Ga) sample shows that the paramagnetic peak at $T \sim 60$ K lies on the low-temperature shoulder of the resistance peak.

3. The following conclusions can be drawn from the data obtained.

a) In the region of thermodynamic equilibrium, where the conductivity is of an activational character (and at higher temperatures), the magnetic susceptibility is diamagnetic and decreases only slightly with increasing temperature. Its value is the same as that of the undoped sample.

b) In the low-temperature region ($T < 50$ K) with metallic delayed photoconductivity, where the photoexcited itinerant electons satisfy quasi-Fermi statistics (the itinerant carriers are degenerate), the magnetic susceptibility is likewise diamagnetic and essentially equal to the high-temperature value.

c) A sharp paramagnetic peak is observed in samples in Ga only in the transitional region ($T \sim 50-70$ K), where the

conductivity is strongly nonequilibrium but the carriers are still nondegenerate. This peak becomes more pronounced as the measurements are performed more rapidly in this temperature range.

d) Moreover, additional growth of paramagnetism following the Curie law is observed in the low-temperature region $T<25$ K, but only if the illumination from the cap is weak, when the system of photoexcited carriers is nondegenerate.

The first and second assertions agree with the theory of magnetic susceptibility of IV–VI semiconductors with a quasirelativistic band spectrum. The expression for the magnetic susceptibility calculated by the linear-response method is[12]

$$\chi=-\frac{\alpha}{6\pi^2}\frac{v}{c}\int_0^{\Lambda}\{n(-\varepsilon_p,\mu,T)-n(\varepsilon_p,\mu,T)\}\frac{v\,dp}{\varepsilon_p}, \quad (1)$$

where α is the fine-structure constant, v is the matrix element of the interband velocity, c is the speed of light, $\varepsilon_p=\sqrt{\Delta^2+v^2p^2}$ is the quasirelativistic electronic spectrum, Δ is the half-width of the forbidden band, $n(\varepsilon_p)$ and $n(-\varepsilon_p)$ are the Fermi distribution functions of the electrons in the conduction and valence bands, respectively, μ and T are the chemical potential and temperature of the electrons, and Λ is the cutoff parameter for the momentum p ($\Lambda v\gg\Delta$). It follows from Eq. (1) that in the range of values of the experimentally studied parameters of PbTe(Ga) ($|\mu|, T<\Delta$, $\Delta\sim0.1$ eV, $\Lambda v\sim10$ eV, $v\sim10^8$ cm/s)

$$\chi=-\frac{\alpha}{6\pi^2}\frac{v}{c}\ln\frac{\Lambda v}{\Delta}\approx-10^{-5}. \quad (2)$$

That is, in thermodynamic equilibrium as well as under quasichemical equilibrium conditions (provided the nonequilibrium carrier distribution function is of a Fermi character) the susceptibility is diamagnetic and depends on μ and T only through small corrections. It should be noted that an expression of the form (2) was obtained in Ref. 13 by direct summation over the Landau levels.

Using Eq. (1) it is easy to calculate the correction $\delta\chi$ in the case of a Boltzmann gas of carriers with density n_1 and effective mass m^* as

$$\delta\chi=\frac{2}{3}\frac{n}{T}\left(\frac{eh}{2m^*c}\right)^2. \quad (3)$$

Here $m^*=\Delta/v^2$, which is much smaller than the free electron mass m ($m^*/m\sim1/20$). As a result, the effective g factor of the carriers is greater than the g factor of a free electron by the same factor. Physically, the correction (3) describes the total susceptibility of doubly spin-degenerate states singly occupied by electrons.

Direct calculations using Eq. (3) show that under quasithermodynamic equilibrium conditions with $T\sim60$ K and density of singly-occupied levels satisfying $n_1\sim10^{15}$ cm^{-3}, the paramagnetic correction $\delta\chi\sim10^{-7}$ is small compared with the total diamagnetic susceptibility of the system. The estimate presented for the density n is an upper limit for the number of nondegenerate itinerant electrons at this temperature and corresponds to the experimentally observed increase in the resistance by two to three orders of magnitude in the region of the transition from metallic to semiconductor behavior (Fig. 1). This result agrees completely with the absence of any paramagnetic anomalies on the high-temperature shoulder of the resistance peak.

On this basis, no paramagnetic anomalies are to be expected on the low-temperature shoulder of the resistance peak, where the temperature and density of the carriers do not differ much from their values on the high-temperature shoulder. In reality, a sharp paramagnetic peak, comparable in magnitude to the absolute magnitude of the diamagnetic susceptibility of the sample, is observed on the low-temperature shoulder of the resistance.

To obtain from Eq. (3) the scale of the paramagnetic peak actually observed, it must be assumed that the density n_1 of singly-occupied levels is not 10^{15} cm^{-3} but rather of the order of 10^{17} cm^{-3}. There could be several formal explanations for such a density n_1: first, the presence of a giant peak of the density of states near the Fermi level; second, strong Coulomb repulsion between electrons with opposite spins, filling each level; and, finally, the existence of a strongly nonequilibrium carrier distribution function in the system.

The first two hypotheses must be rejected because under these conditions a paramagnetic anomaly would also be observed on the high-temperature shoulder of the resistance peak. Therefore only the third hypothesis need be considered.

In order for a large number of singly-occupied levels to exist in the system, not only must the photoexcited carriers have a long lifetime τ, but the transition time of electrons between these levels must also be long. The latter is possible only for states localized on impurities and a low density of delocalized (itinerant) electrons. Otherwise, impurity–band transitions will rapidly establish quasithermodynamic equilibrium in the system of photoexcited carriers. The temperature region where the paramagnetic peak is observed satisfies these conditions. In this region the lifetime τ of the nonequilibrium carriers is still long enough (comparable to the duration of the experiment) for the number of photoexcited carriers at low temperatures to remain sufficiently large, but it is already inadequate to maintain in the system of these carriers a quasi-Fermi distribution function with the chemical potential located in the conduction band.

The number of singly-occupied levels can be estimated using the following simple kinetic scheme. At low temperatures, as a result of external excitation, photoexcited electrons (two for each trivalent Ga atom) appear in the conduction band because of a change in the configuration of the Ga impurity from s^2p^1 to s^0p^3. This excited state of the carriers is found to be metastable, since the reverse (recombination) process should proceed via an intermediate (s^1p^2) divalent state of Ga (according to the scheme $s^0p^3\rightarrow s^1p^2-s^2p^1$), which has a higher energy than the s^0p^3 and s^2p^1 states, i.e., it is separated from them by an energy barrier U.[2)] In this manner, electrons accumulate in sufficient number to form a quasi-Fermi distribution. At low temperatures this distribution is characterized by the presence of two carriers in each of almost all states occupied by the electrons.

As temperature T increases, the thermal energy of these carriers increases and finally becomes high enough for the carriers to overcome the recombination energy barrier U. In this region the lifetime of a photoexcited carrier decreases according to an activation law ($\tau \propto \exp(U/T)$). The dynamics of the number of pair-filled states in the indicated temperature range can thus be roughly described by the equation

$$\frac{dn_2}{dt} = -\frac{n_2}{\tau}, \tag{4}$$

where τ is the lifetime and n_2 is the number of pair-filled states. Initially, the number $n_2(0)$ of pairs equals essentially half the carrier density because of the Fermi character of the carrier distribution. Hence initially the number n_1 of singly-occupied levels can be set equal to zero and their temporal dynamics can be described by the equation

$$\frac{dn_1}{dt} = -\frac{n_1}{\tau} - \frac{dn_2}{dt}, \tag{5}$$

since a singly-occupied level arises as a result of the recombination of a single carrier from a pair. It should be noted that the kinetic equation (5) is valid only for localized single-electron states. In the opposite case, a term of the type $-\gamma n_1$, describing the reverse transformation of singly-occupied states into pairs, must be added to it. This term is certainly not small for delocalized states (or strongly overlapping states). It is obvious that the existence of such a term would sharply decrease the values of n_1 allowed by Eqs. (4) and (5), while the paramagnetic peak would be unobservable. Solving Eqs. (4) and (5) simultaneously yields the following time dependence of the number of singly-occupied states:

$$n_1(t) = \frac{t}{\tau} n_2(0) \exp\left(-\frac{t}{\tau}\right). \tag{6}$$

This equation implies that the maximum number of singly-occupied levels that can be attained as a result of temporal evolution is only e times smaller than the initial number of pairs.

Thus, the appearance of a sharp paramagnetic peak with high intensity on the low-temperature shoulder of the resistance is described completely naturally in this scheme and attests to the presence in lead telluride of shallow localized states associated with Ga impurity atoms in a trivalent state and lying below the conduction-band bottom.

In closing, we briefly discuss the nature of the Curie law in the susceptibility observed at low temperatures under weak illumination conditions. In principle it can also be attributed to the presence of a large number of shallow singly-occupied states under these conditions. However, it has not been ruled out that here an additional contribution arises from Ga in an unstable divalent state. Direct measurements (for example, by electron spin resonance) of the effective g factor of magnetic centers could answer unequivocally the question of the nature of the appearance of the Curie law at low temperatures.

B. A. Volkov is grateful to the Russian Fund for Fundamental Research for support (Grants Nos. 96-02-16701, 96-02-19022, 96-15-96474) and the international program INTAS–RFBR (Grant No. 95-1136).

*)E-mail: vasil@lt.phys.msu.su

[1)]Here there is no need to consider the possibility of the formation of a hydrogen-like state associated with the long-range part of the Coulomb potential of the impurity. In IV–VI semiconductors the permittivity satisfies $\varepsilon > 100$, the effective is mass $m^* \sim 10^{-2}$ times the free-electron mass, and so the characteristic binding energy of such a state is $\leqslant 10^{-4}$ eV.

[2)]The widely used scheme for producing a barrier U by means of deformational effects is not entirely satisfactory, although, of course, the change in the valence of impurities is accompanied by a deformation.

[1] P. M. Mooney and J. Appl. Phys. **67**, R1 (1990).
[2] J. Emsley, *Elements*, Oxford University Press, New York, 1991, 2nd edition.
[3] G. Martinez, M. Schluter, and M. L. Cohen, Phys. Rev. B **11**, 651 (1975).
[4] O. A. Pankratov and B. A. Volkov, Sov. Sci. Rev., Sect. A **9**, 357 (1987).
[5] V. I. Kaĭdanov and Yu. I. Ravich, Usp. Fiz. Nauk **145**, 51 (1985) [Sov. Phys. Usp. **28**, 31 (1985)].
[6] B. A. Akimov, N. B. Brandt, S. A. Bogoslovskiĭ, L. I. Ryabova, and S. M. Chudinov, JETP Lett. **29**, 9 (1979).
[7] B. M. Vul, I. D. Voronova, G. A. Kalyuzhnaya, T. S. Mamedov, and T. Sh. Ragimova, JETP Lett. **29**, 18 (1979).
[8] Yu. Kagan and K. A. Kikoin, JETP Lett. **31**, 335 (1980).
[9] B. A. Volkov and O. A. Pankratov, Dokl. Akad. Nauk SSSR **255**, 93 (1980) [Sov. Phys. Dokl. **25**, 922 (1980)].
[10] B. A. Volkov and O. M. Ruchaĭskiĭ, JETP Lett. **62**, 217 (1995).
[11] M. Matyas, Czech. J. Phys. **8**, 301 (1958).
[12] B. A. Volkov and O. M. Ruchaĭskiĭ, Fiz. Tverd. Tela (St. Petersburg) **40**, 57 (1998) [Phys. Solid State **40**, 50 (1998)].
[13] L. A. Fal'kovskiĭ, A. V. Brodovoĭ, and G. V. Lashkarev, Zh. Éksp. Teor. Fiz. **80**, 334 (1981) [Sov. Phys. JETP **53**, 170 (1981)].

Translated by M. E. Alferieff

Thermal expansion of thin C_{60} films

A. T. Pugachev, N. P. Churakova, N. I. Gorbenko, and Kh. Saadli

Khar'kov State Polytechnical University, 310002 Khar'kov, Ukraine

E. S. Syrkin*)

Physicotechnical Institute of Low Temperatures, Ukrainian National Academy of Sciences, 310164 Khar'kov, Ukraine

(Submitted 18 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1868–1875 (November 1998)

The thermal expansion coefficient α and structure of C_{60} films with thickness $t \sim 3$–10 nm were investigated in the temperature interval from room to liquid-nitrogen temperature by electron-optical methods. The thermal expansion coefficient was determined from the temperature shift of the diffraction maxima in the electron diffraction patterns. The objects of investigation were epitaxial C_{60} films condensed in vacuum on a (100) NaCl cleavage surface and oriented in the (111) plane. A surface-induced size effect in the thermal expansion coefficient was observed. It was established that as t decreases α_f increases and is described well by the relation $\alpha_f = 17 \cdot 10^{-6}\,\mathrm{K}^{-1} + 8.3 \cdot 10^{-5}\,\mathrm{nm\,K}^{-1}\,t^{-1}$. This relation was used to estimate the linear expansion coefficient α_s of the C_{60} surface in the (111) plane as $\alpha_s = 60 \cdot 10^{-6}\,\mathrm{K}^{-1}$, which is several times larger than the bulk value. The experimental results agree satisfactorily with the theoretical calculations of the mean-square displacements of molecules located in a region near the surface. © *1998 American Institute of Physics.* [S1063-7761(98)02411-1]

1. INTRODUCTION

Fullerites — a new molecular form of condensed carbon — have been the object of intensive investigations in the last few years. Many investigations have been devoted to various aspects of fullerite synthesis, structure, properties, and applications.[1–3] Investigations of fullerite in a film state are of interest in themselves because the structural and geometric parameters can be varied over wide limits by altering the conditions of condensation.[4] For thicknesses $t \leqslant 10$ nm the surface layers should make a considerable contribution to film properties.[5] As a result, the properties determined by the anharmonicity of the forces acting between C_{60} molecules should differ. One such property is the thermal expansion, characterized by the coefficient α. For the range of geometric sizes studied, the effect of the surface should be more pronounced for bodies with a large lattice period. Fullerite C_{60} is such an object. It should be noted that because of the van der Waals nature of the interaction between C_{60} molecules thin fullerite films are a convenient object for checking theoretical models.

The present work is devoted to investigations of the thermal expansion of vacuum-condensed thin C_{60} films in the temperature interval from room to liquid-nitrogen temperature. The thermal expansion was investigated experimentally by transmission high-energy electron diffraction, using the shift of the diffraction peaks in the electron diffraction patterns.[6] The structure of the C_{60} films was also tested by electron microscopy methods. An important advantage of a diffraction study of the properties is that it is possible to follow the structure and state of the experimental object during the course of a thermophysical experiment. This becomes especially important for studying films of thickness 1–10 nm. The study of the thermal expansion of such films is of intrinsic interest and in many cases yields helpful information about surface properties.

2. EXPERIMENT

The fullerite films were obtained by evaporating and condensing, in a vacuum $\sim 10^{-3}$ Pa, crystal particles with a purity of at least 99.9%. C_{60} crystal particles of mass $10^{-4} - 10^{-3}$ g were evaporated from a quartz crucible Joule-heated to ~ 700 K. The substrate consisted of a NaCl single crystal. Condensation was performed on a (001) NaCl surface. The substrate temperature was ~ 360 K. The condensation rate was $\sim 0.1\,\mathrm{nm\,s}^{-1}$. For subsequent electron-optical investigations the fullerite films were separated from the substrate and secured on electron-microscope meshes. To do this, a single crystal with a film was immersed at angle $\sim 20°$ in distilled water. The fullerite film detached from the substrate in several seconds as a result of partial dissolution of the NaCl. Next, the film was recovered from the water onto copper meshes with cell-size ~ 0.05 mm.

The experimental scheme used to determine the thermal expansion coefficient of C_{60} films by transmission high-energy electron diffraction is shown in Fig. 1. A sample was secured in a cryostat and placed in an electron diffraction camera. In the electron diffraction camera the sample was surrounded by a screen cooled to liquid-nitrogen temperature. The sample temperature could be varied from room to liquid-nitrogen temperature. The sample and a control were placed in the same plane, and the electron beam passed through them simultaneously. The control consisted of a thin

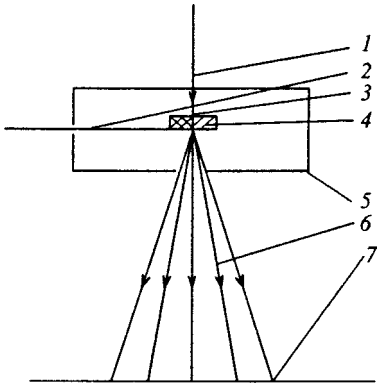


FIG. 1. Experiment arrangement: *1* — Electron beam, *2* — cryostat, *3* — control, *4* — sample, *5* — cooled screen, *6* — diffracted beams from sample and control, *7* — fluorescent screen (photographic plate).

annealed Al film. An electron diffraction pattern consisting of a superposition of the patterns from the sample and the control was recorded on a photographic plate. The (transmission) electron diffraction patterns were obtained with an accelerating voltage of 40 kV and electron beam density less than $10^{-6}\,\mathrm{A\,cm^{-2}}$ so as to reduce the effect of the electron beam on the experimental object to a minimum. The possible instrumental errors in the measurements of the interplanar distances were taken into account by using the control. Published data on the temperature dependence of the lattice period of bulk aluminum were used to determine the temperature dependence of the electron diffraction camera constant $2L\lambda$.[7] The attachment and screen temperatures were measured to within $\pm 3°$ with copper–constantan thermocouples.

The thickness of these thin condensed C_{60} films was preset by the mass of the charge and the evaporation geometry. It was monitored with a quartz resonator during the experiment and then refined on the section being studied using electron diffraction by finding the size of a reciprocal-lattice site in the direction normal to the film.[8] For this, the film was tilted relative to the electron beam, and the angle within which the reflection sphere intersects a given reciprocal-lattice site was determined with a goniometer. The film thickness was found using the (220) and (422) reflections. The error in measuring the thickness is estimated to be 10%. Films of thickness 3–10 nm were investigated.

The thermal expansion coefficient α was determined from data on the variation of the interplanar distances as a function of temperature:

$$\alpha = \Delta d/d\Delta T, \tag{1}$$

where $\Delta d/d$ is the relative change occurring in the interplanar distances as a result of the thermal expansion due to a temperature change ΔT. Since $\Delta d/d = -\Delta 2r/2r$, the ratio $\Delta d/d$ was measured experimentally according to the change in the distances $2r$ between the diffraction reflections in the electron diffraction pattern. The values of $\Delta d/d$ were found using the reflections (422). The ratio $\Delta 2r/2r$ was measured in an optical microscope with an error of $5\cdot 10^{-4}$, which leads to an error of $\pm 3\cdot 10^{-6}\ \mathrm{K^{-1}}$ in the thermal expansion coefficient. Note that the average value of α over the indicated temperature interval was determined. In addition, for normal incidence of the electron beam on the film the values of $\Delta d/d$ and therefore α are obtained in a direction parallel to the film surface.

3. RESULTS AND DISCUSSION

According to electron-diffraction and electron-microscopy data the C_{60} films were continuous and possessed a face-centered (fcc) lattice. The electron diffraction

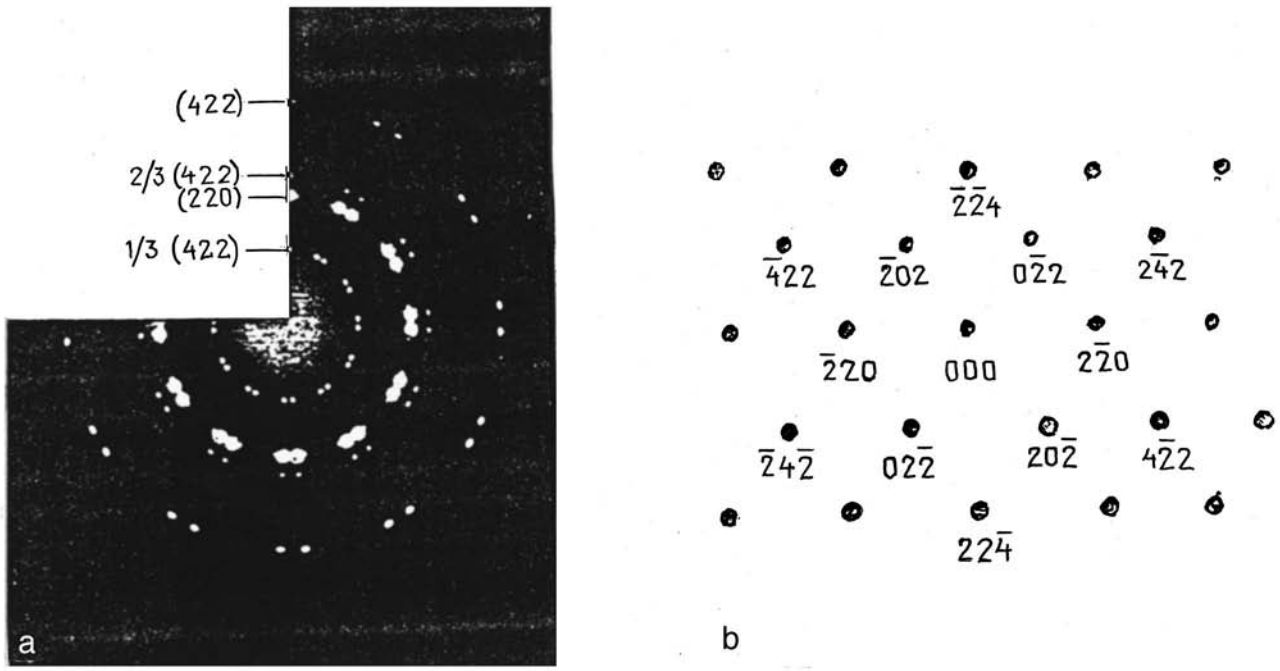


FIG. 2. Electron diffraction pattern of a fullerite film of thickness 4.5 nm (a) and section of the reciprocal lattice of a fcc crystal with (111) orientation (b).

TABLE I. Results of precise measurement of the relative changes $\Delta d/d$ in the interplanar distances of C_{60} films with thickness t in the temperature range T_1-T_2.

t, nm	T_1-T_2, K	$\Delta d/d\cdot 10^3$	$\alpha_f\cdot 10^6$, K^{-1}
3.5	269 – 83	11.6	44 (Ref. 12)
4.5	275 – 88	10.2	36
6	273 – 80	8.2	25
10	274 – 80	8.4	27
∞	273 – 80	7.1	19.5 (Ref. 10)
∞	273 – 80	6.8	19 (Ref. 11)

patterns (Fig. 2) contained the reflections (220) and (422) typical of the (111) orientation, for which the (111) plane of C_{60} is parallel to the (100) NaCl cleavage surface. However, besides this, reflections with interplanar distances 0.86 and 0.43 nm, which can be identified as the reflections 1/3(422) and 2/3(422) due to stacking faults,[9] were present in the electron diffraction pattern. It should be noted that, while the lattice period of NaCl ($a=0.564$ nm) differed substantially from that of C_{60} fullerite ($a=1.42$ nm) this film–substrate system contains favorable orientation ratios for epitaxial growth of C_{60} films. Thus, to within 1% two C_{60} lattice periods equal five NaCl lattice periods and to within 2% two C_{60} [110] diagonals equal seven lattice periods of NaCl. However, the analysis performed showed that neither the parallel nor the 45-degree orientation is realized. The presence of 24 reflections of the type (220) and (422) together with the six reflections expected for this orientation attest to the fact that the structure of the films is the result of four-position epitaxial nucleation and subsequent growth. According to data from dark-field electron-microscope photographs, the average size of the crystal particles is 30–40 nm. According to the perfection of the structure, the experimental epitaxial C_{60} films, oriented in the (111) plane fall between textured and single-crystalline.

The results of a precision measurement of the interplanar distances $\Delta d/d$ of C_{60} fullerite films in the temperature range T_1-T_2 are collected in Table I. The quantity $\Delta d/d$ increases as film thickness decreases. Thus, the ratio $\Delta d/d$ for a film of thickness 3.5 nm is 1.6 times greater than for bulk fullerite.[10,11] The measured dilatometric effect of the lattice period of thin C_{60} films in the experimental temperature range is due to both the change in $\Delta a/a$ on account of thermal expansion and the change due to the phase transition of fullerite from a fcc lattice into a simple cubic (SC) lattice at $T\approx 260$ K. At temperatures $T<260$ K the C_{60} molecules remain in the same positions, but their three-fold axes become oriented along the $\langle 111\rangle$ directions. Information on the temperature dependence of the lattice period of thin C_{60} films is therefore required for quantitative interpretation of the experimental data.

Figure 3 shows the temperature dependence of the lattice period a of C_{60} films of thickness 4.5 nm. According to this figure, a jump due to a phase transition of fullerite from a SC to a fcc lattice is observed at $T\approx 260$ K. The transition temperature as well as the magnitude of the jump in the films agree well with the dilatometric and x-ray diffraction data

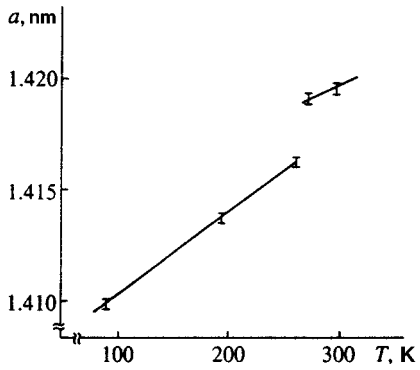


FIG. 3. Lattice period a of C_{60} films of thickness 4.5 nm versus temperature T.

obtained from investigations of this phase transition in bulk C_{60}, where $\Delta a/a\sim 3\cdot 10^{-3}$.[10,11] These data were used to calculate average linear thermal expansion coefficients α_s of the films for the indicated temperature range. We note that the experimental temperature interval lies above the Debye temperature of fullerites $\Theta_D\approx 70$ K.[13] This makes it possible to neglect the temperature dependence of the thermal expansion coefficient. The linear temperature dependence of the lattice period of C_{60} films attests to this (Fig. 3). The experimental temperature interval included the region (80–260) K of the oriented SC phase and part of the region (260–275) K of the fcc phase. It follows from Fig. 3 that the oriented phase makes an overwhelming contribution to the change in the lattice period due to thermal expansion and therefore its thermal expansion coefficient is measured. Thus, the average values found for the thermal expansion coefficient correspond to this phase.

The experimental thickness range of C_{60} films oriented in the (111) plane corresponds to 4–12 interplanar spacings. For such objects the contribution of surface and near-surface layers to the properties becomes considerable, since the coordination numbers of molecules in surface and near-surface layers differ from those of molecules in the interior volume. This has the effect that, for example, the mean-square amplitude $\overline{u_s^2}$ of the vibrations of molecules on the surface and in the first two layers near the surface is larger than the corresponding values $\overline{u_v^2}$ in the interior volume.[5] According to Ref. 14, the mean-square displacements and thermal expansion coefficients are related simply by

$$\overline{u_s^2}/\overline{u_v^2}\approx \alpha_s/\alpha_v\,. \tag{2}$$

Hence the values of the linear expansion coefficients for the surface and near-surface layers should also be overestimated.

The relation (2) is extremely convenient for theoretical interpretation of the results of measurements of the linear expansion coefficients, since it connects anharmonic quantities (linear expansion coefficients) and the mean-square displacements calculated in the harmonic approximation. The method of Jacobian matrices,[15,16] also known in the literature as the recursive method,[17] is an effective method for calculating the temperature dependences of the mean-square displacements of both molecules at the surface and in the interior volume.

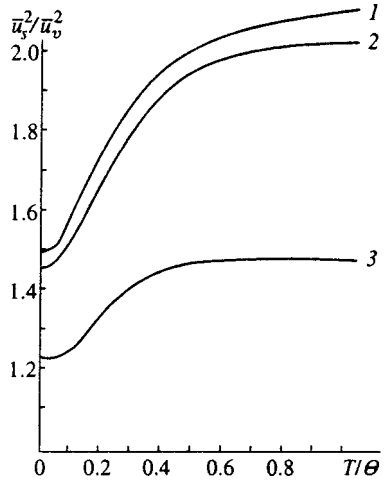


FIG. 4. Temperature dependences of $\overline{u_s^2}/\overline{u_v^2}$ for a (110) surface plane: *1, 2, 3* — for vibrations of molecules along [001], [110], and [110], respectively.

The mean-square displacements of molecules in an arbitrary system can be written as[15,16]

$$\overline{u_h^2}=\frac{\hbar}{2m}\int_0^{\lambda_m}\frac{1}{\sqrt{\lambda}}\coth\left(\frac{\hbar\sqrt{\lambda}}{2kT}\right)\rho_h(\lambda)d\lambda, \tag{3}$$

where λ is the squared vibrational frequency, λ_m is the maximum value of λ, and $\rho_h(\lambda)$ is the so-called spectral density generated by the vector h. The generating vector h corresponds to the displacement of a selected molecule in one direction. The mean-square displacements $\overline{u_s^2}$ or $\overline{u_v^2}$ of surface or interior molecules, respectively, are calculated according to whether the chosen molecule is located on the surface or in the interior volume.

Figures 4 and 5 display the computational results obtained for the temperature dependences of $\overline{u_s^2}/\overline{u_v^2}$ by the method of Jacobian matrices using Eq. (3) for a fcc crystal

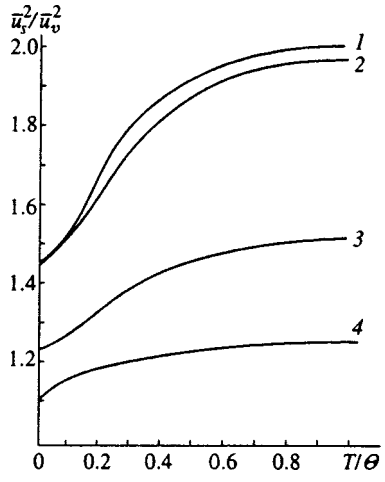


FIG. 5. Temperature dependences of $\overline{u_s^2}/\overline{u_v^2}$ for surface planes of the type (001) and (111): *1, 2* — for vibrations of molecules in the direction perpendicular to the planes (001) and (111), respectively; *3, 4* — for vibrations of molecules in a direction parallel to the planes (001) and (111), respectively.

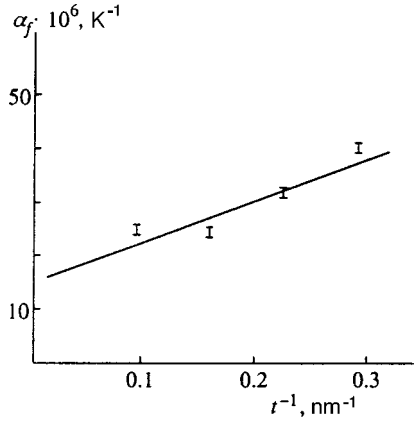


FIG. 6. Thermal expansion coefficient of C_{60} films versus thickness.

with a central interaction between nearest neighbors with various orientations of the surface plane. It is easy to see that the largest differences in the mean-square displacements (~ 2.2) are associated with the (110) plane, since for this orientation a molecule on the surface possesses seven nearest neighbors (instead of 12 in the interior) and for the (100) and (111) orientations it possesses eight and nine nearest neighbors, respectively. Moreover, for the (110) orientation of the surface plane the vibrations in this plane are strongly anisotropic.

For (111) orientation, in the present model the ratio $\overline{u_s^2}/\overline{u_v^2}$ is 1.2 for displacements in the plane (for displacements in the normal direction the ratio ≈ 2). However, it should be noted that the results are presented for a model in which the surface distortion was neglected. Obviously, if this circumstance is taken into account, then the ratio $\overline{u_s^2}/\overline{u_v^2}$ can change considerably.[18,19] In real systems, as a rule, when a surface is formed, the interaction between surface and subsurface layers is much weaker than between layers in the interior. As a result, the mean-square displacements of molecules located in the surface region increase appreciably.

To a first approximation the thermal expansion coefficient of a film with thickness t and two free surfaces can be represented as

$$\alpha_f=\alpha_v+2(\overline{\alpha}_s-\alpha_v)\Delta t/t, \tag{4}$$

where Δt is the outer layer, whose lattice dynamics differs from that of the interior layers, and $\overline{\alpha_s}$ is the average value of the thermal expansion coefficient for this layer.

The experimental data on the thickness dependence of the thermal expansion coefficient of films are shown in Fig. 6. As follows from Fig. 6, the values of α_f in the experimental thickness range are described quite well by expression (4). The value found from Fig. 6 for the thermal expansion coefficient of bulk C_{60} is $17\cdot 10^{-6}$ K^{-1}, which is close to the published data.[10,11] The quantity $2(\overline{\alpha}_s-\alpha_v)\Delta t=8.3\cdot 10^{-5}$ nm K^{-1}. To determine α_s it is necessary to have information about the value of Δt. Theoretical estimates show that Δt corresponds to two interplanar distances.[5] For the indicated film orientation and $d_{(111)}=0.82$ nm it was assumed that the surface layer is 1.64 nm thick. For these val-

ues of Δt the quantity $\overline{\alpha_s}=40\cdot 10^{-6}\ \mathrm{K}^{-1}$. Information about $\overline{\alpha_s}$ makes it possible to estimate the linear expansion coefficient α_s of a surface atomic layer, using the expression $\overline{\alpha_s}=(\alpha_s+\alpha_v)/2$. Hence it follows that for the outer atomic layer with (111) orientation the linear expansion coefficient α_s in a direction parallel to the film surface is $60\cdot 10^{-6}\mathrm{K}^{-1}$. This is 3.5 times greater than the bulk value.

In closing, we note that the value of α_s found in this manner is somewhat too high. Treating the film as a very thin, perfect, plane-parallel plate is a simplification. A real film surface is well known to be several times larger than the geometric surface because of natural roughness. Consequently, not only the thermal expansion component parallel to the surface but also the component perpendicular to the surface contribute to the experimentally measured dilatometric effect $\Delta d/d$. Regions adjoining defects in the film, primarily the boundaries of crystal particles, contribute to the measured effect $\Delta d/d$. However, simple estimates show that this contribution can be neglected for 30–40 nm crystal particles. Hence it follows that the experimental and calculated thermal expansion coefficient of an atomic surface layer of C_{60} fullerite agree well with one another, considering the approximation of the model and the real structure and geometry of the experimental objects.

This work was supported by the Ukrainian Ministry of Education and the Ukrainian Fund for Fundamental Research (project USKO-97).

*)E-mail: syrkin@ilt.kharkov.ua

[1] H. Kroto, Rev. Mod. Phys. **69**, 703 (1997).
[2] A. V. Eletskiĭ and B. M. Smirnov, Usp. Fiz. Nauk **163**, 33 (1993).
[3] V. M. Loktev, Fiz. Nizk. Temp. **18**, 217 (1992) [Sov. J. Low Temp. Phys. **18**, 149 (1992)].
[4] L. S. Palatnik, M. Ya. Fuks, and V. M. Kosevich, *Formation Mechanism and Substucture of Condensed Films*, Nauka, Moscow, 1972.
[5] A. Maradudin, Solid State Phys. **18**, 273 (1966).
[6] A. T. Pugachev and N. P. Churakova, Izv. Akad. Nauk, Ser. Fiz. **57**, 126 (1993).
[7] M. E. Straumanis and C. L. Woodard, Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr. **27**, 549 (1971).
[8] P. B. Hirsch, A. Howie, R. B. Nicholson, D. W. Pashley, and M. J. Whelan, *Electron Microscopy of Thin Crystals*, Butterworth, London, 1967 [Russian translation, Mir, Moscow, 1968].
[9] W. B. Zhao, X.-D. Zhang, K. J. Luo *et al.*, Thin Solid Films **232**, 149 (1993).
[10] F. Gugenberger, R. Heid, C. Meingast *et al.*, Phys. Rev. Lett. **69**, 3774 (1992).
[11] L. S. Fomenko, V. D. Natsik, S. V. Lubenets *et al.*, Fiz. Nizk. Temp. **21**, 465 (1995) [Low Temp. Phys. **21**, 364 (1995)].
[12] A. T. Pugachev, N. P. Churakova, and N. I. Gorbenko, Fiz. Nizk. Temp. **23**, 854 (1997) [Low Temp. Phys. **23**, 642 (1997)].
[13] W. P. Beyermann, M. F. Hundley, and Thompson, Phys. Rev. Lett. **68**, 2046 (1992).
[14] V. E. Kenner and R. E. Allen, Phys. Lett. A **39**, 245 (1972).
[15] V. I. Peresada, in *Condensed-State Physics*, Physicotechnical Institute of Low Temperatures, Ukrainian SSR Academy of Sciences, Khar'kov, 1968, p. 172.
[16] V. I. Peresada and E. S. Syrkin, Surf. Sci. **54**, 293 (1976).
[17] R. Haydock in *Solid-State Physics*, edited by H. Ehrenreich *et al.*, Academic Press, New York, 1980, Vol. 35, p. 129.
[18] I. A. Gospodarev, E. S. Syrkin, and S. B. Feodos'ev, Poverkhnost' **2**, 23 (1996).
[19] E. S. Syrkin, Fiz. Nizk. Temp. **2**, 1211 (1976) [Sov. J. Low Temp. Phys. **2**, 591 (1976)].

Translated by M. E. Alferieff

Magnetic resonance of intrinsic defects in the spin–Peierls magnet $CuGeO_3$

A. I. Smirnov and V. N. Glazkov

P. L. Kapitsa Institute of Physics Problems, Russian Academy of Sciences, 117334 Moscow, Russia

L. I. Leonyuk and A. G. Vetkin

Moscow State University, 119899 Moscow, Russia

R. M. Eremina

E. K. Zavoĭskiĭ Kazan' Physicotechnical Institute, Russian Academy of Sciences, 420029 Kazan', Russia

(Submitted 17 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 1876–1896 (November 1998)

Magnetic resonance in pure single-crystal $CuGeO_3$ at frequencies 9–75 GHz in the temperature range 1.2–25 K is investigated. Splitting of the magnetic-resonance line into several spectral components is observed at temperatures below 5 K, where spin–Peierls dimerization suppresses the magnetic susceptibility and the ESR signal intensity. Analysis of the magnetic resonance spectra over a wide frequency range with different directions of the magnetic field at different temperatures makes it possible to identify among these components the ESR signals due to defects, having effective spin $S=1/2$ and spin $S=1$, in the spin–Peierls phase. The g factor corresponding to these ESR signals is the same and close to the value characteristic for the ion Cu^{2+}. Another magnetic-resonance line is characterized by a strongly anisotropic g factor and an increase (at a threshold in the excitation power) in the susceptibility both at resonance and in the line wings. These signals are tentatively attributed to two possible types of planar defects arising on the walls of domains of the spin–Peierls state with different values of the dimerization phase. © *1998 American Institute of Physics.* [S1063-7761(98)02511-6]

1. INTRODUCTION

The inorganic compound $CuGeO_3$, which possesses magnetic and crystallographic properties typical of spin–Peierls compounds, has been studied intensively in the last few years by various methods. The magnetic structure of this crystal is based on one-dimensional chains of Cu^{2+} ions running along the c axis of an orthorhombic crystal.

A sharp decrease of the magnetic susceptibility[1] with a simultaneous displacement of the atoms accompanied by doubling of the lattice period in the directions a and c[2,3] has been observed below the spin–Peierls transition temperature. The accompanying change in the magnetic properties is attributed to the formation of dimers consisting of magnetic atoms between which the distance became smaller and the exchange integral larger than in the initial state at a temperature above the transition. The ground state of a spin–Peierls crystal is a singlet state, while the excited triplet states are separated from the ground state by an energy gap. The magnetic susceptibility should vanish at absolute zero temperature. The gap width and the spin–Peierls transition temperature are determined by varying the exchange integral from a state with the atoms moving closer to one in which they are moving apart.[4,5] We note for comparision that a nondimerized chain of $S=1/2$ spins with antiferromagnetic exchange possesses a gapless excitation spectrum, and the ground state does not have Néel order.[6]

The transition described above stems from an instability that develops in the presence of an interaction between the one-dimensional spin chains and the three-dimensional elastic subsystem of the crystal. While the spin chains form a quasi-one-dimensional magnet, the restructuring of the crystal is three-dimensional and the dimers form an ordered sublattice. The displacements of the copper atoms occur in the c direction, while rotations of the oxygen octahedra surrounding the copper ions occur in the ab plane.[2] Investigations of the structure show that the displacements of the copper ions in neighboring chains are correlated in antiphase, i.e.,, a translation by the vector $\mathbf{a}+\mathbf{c}$ or $\mathbf{b}/2+\mathbf{c}$ brings the dimers into coincidence. Here $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$ are the primitive translations of the nondimerized phase. The period of the arrangement of the ions in the direction b in the high-temperature phase is half the lattice period, since two copper ions we are associated with one another by a translation by $\mathbf{b}/2$ lie inside a primitive cell. The displacements of the oxygen ions are correlated similarly.

These translation bring the octahedra into coincidence. A rearrangement of the lattice is therefore accompanied by dimerization of copper ions along the c axis and oxygen ions along the a and b axes. The displacements δz_{klm} of the copper ions relative to their positions in the nondimerized lattice can be described by the relation

$$\delta z_{klm}=\xi\cos[(k+l+m)\pi+\psi]. \tag{1}$$

Here ξ is the amplitude of the displacement; k, l, and m are the coordinates of the copper ions relative to a reference

copper ion, expressed in the units a, $b/2$, and c in a coordinate system with axes directed along the crystal axes a, b, and c. The quantity ψ is called the dimerization phase and can assume one of two values — 0 or π.[7] The state of the crystal is doubly degenerate with respect to this parameter.

The main properties of the spin–Peierls phase in $CuGeO_3$ according to Ref. 3: are as follows: transition temperature $T_{SP}=14.2$ K, in-chain exchange integral $J_c=10.6$ meV, zero-temperature energy gap $\Delta=2$ meV, relative variation of the exchange integral in the dimerized chain $\delta=0.042$. The ratios of the interchain to in-chain exchanges are $J_b/J_c=0.11$ and $J_a/J_c=-0.011$. The maximum dimerization-induced displacement of the copper ions is 0.007 Å .

The data from measurements of the magnetic susceptibility, neutron diffraction investigations of the magnetic and crystallographic structures, and the study of the excitation spectra show that the idea of a spin–Peierls transition leads to a correct qualitative description of the magnetic and lattice properties of $CuGeO_3$ (see, for example, Ref. 3).

Freezing of the magnetic susceptibility is incomplete in real samples. The susceptibility in typical samples decreases approximately ten-fold. The susceptibility is a minimum at 5 K and increases somewhat, approximately by the factor 1.5, as temperature decreases further. This residual susceptibility is ordinarily attributed to the presence of impurities or dangling chain ends. Other conjectures, described below, have also been made.

When defects are introduced into the lattice or magnetic subsystem, the spin–Peierls transition temperature decreases, and three-dimensional antiferromagnetic order is observed at sufficiently low temperature.[8,9] For example, the introduction of 0.07% Si or 2% Zn makes the $CuGeO_3$ crystal an antiferromagnet with Néel temperature of about 4 K. A characteristic feature of the long-range antiferromagnetic order induced by introducing impurities is that the Néel state and spin–Peierls dimerization coexist. The average spin at a site is several tenths of the nominal value. Impurity-induced transformation of a nonmagnetic ground state into an antiferromagnetic state occurs because dimerization is supprssed near a lattice defect or a missing spin.[7,10] The absence of dimerization gives rise to antiferromagnetic correlations of the spins near a defect both along a chain and in the transverse directions due to exchange interactions. The average magnitude of the spin projection at a site decreases away from a defect. The correlated regions of neighboring defects overlap, producing long-range magnetic order. Order breaks down at finite temperature when the energy of the thermal fluctuations is sufficient to destroy the correlations between neighboring defects.

In relatively pure samples, where the magnetic susceptibility below the transition temperature decreases by more than a factor of 10, long-range magnetic order has not been observed down to 1.2 K. Nonetheless, pure $CuGeO_3$ crystals exhibit unusual properties at low temperatures. Investigations of electron spin resonance in $CuGeO_3$ crystals[11,12] show that at low temperatures the spectrum of the signal in the region of residual magnetic susceptibility becomes more complicated, splitting into several components whose origin is still unclear. In Ref. 13 it was observed that in this temperature range an electric field influences the magnetic susceptibility, likewise indicating that the magnetic state of the crystal is unusual. It was surmised that the residual susceptibility could be due to magnetic-lattice defects of the spin–Peierls phase, which form at temperature T_{SP} on the boundaries of crystallites with different values of the dimerization phase. Then the residual magnetic susceptibility can correspond to the density of magnetic defects, which is higher than that of the defects in the paramagnetic phase.

In the present investigation the magnetic resonance spectra in the region of the residual magnetic susceptibility in pure $CuGeO_3$ single crystals are studied in detail for the purpose of identifying further the magnetic defects in the spin–Peierls phase and the effect of the boundaries of the spin–Peierls crystallites on the low-temperature magnetic properties is discussed.

2. PROCEDURE AND SAMPLES

The $CuGeO_3$ samples were grown from highly purified reagents by spontaneous crystallization from a melt with slow cooling. The velocity v_{cr} of the crystallization front was 10^{-3} cm/h. The impurity content was monitored by activation analysis and atomic plasma spectroscopy. The data from these control experiments show that the content of the impurities Fe, Ni, Mn, and Co in samples of the main series studied in this work did not exceed 10^{-4} per copper ion for each of the indicated types of impurities. The samples consisted of transparent blue plates oriented along the c axis and possessing well-expressed bc planes. The crystals were 4 mm long, 2 mm wide, and 0.5 mm thick.

The magnetic impurities and defects of a spin–Peierls crystal lead to the presence of a residual magnetic susceptibility. Hence the quality of a spin–Peierls crystal can be characterized by the ratio Q of the magnetic susceptibility at 15 K to the minimum magnetic susceptibility measured at $T=5$ K. The fewer defects in the crystal, the larger the quality factor Q is. The samples of the main series yielded $Q=20$.

To compare samples containing a different number of defects of different nature, other crystals were also studied. Magnetic resonance spectra were obtained for the sample No. 2, produced by the float-zone method, from Ref. 12. This crystal contained 10^{-3} Fe impurity atoms per cell and its quality factor $Q=7$.

To monitor the contribution of structural defects to the residual susceptibility another series of samples was prepared using the same reagents as for the samples in the main batch with $Q=20$ but with faster cooling (the velocity of the crystallization front was 1 cm/h). For these samples $Q=6$. The quality factor of the samples grown with a crystallization velocity of 6 cm/h was $Q=3$. To monitor the influence of nickel impurity crystals with the composition $Cu_{0.995}Ni_{0.005}GeO_3$, grown with crystallization velocity 1 cm/h, were prepared.

The magnetic resonance lines in the frequency range 18–75 GHz were recorded as the magnetic-field-dependence of the power of the microwave signal transmitted through the

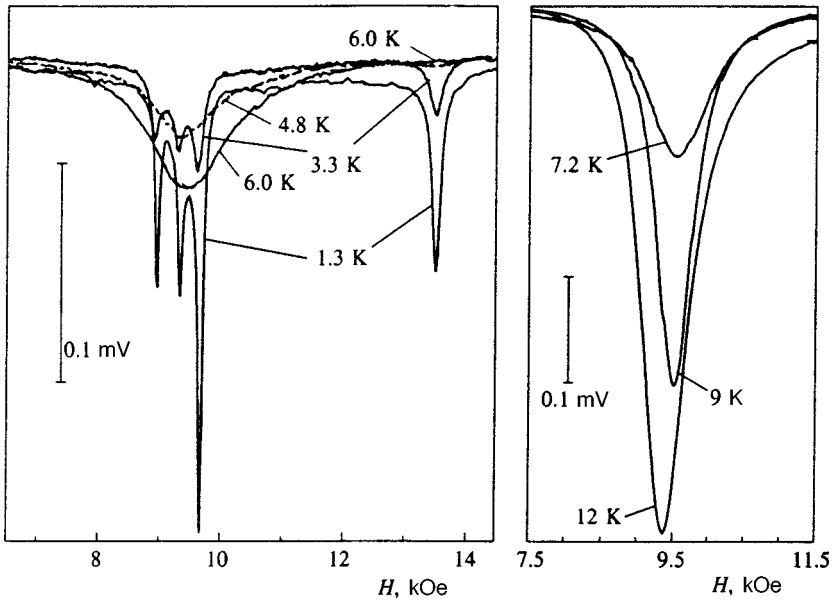


FIG. 1. Magnetic resonance line at 26.7 GHz at various temperatures. Curves of the variation of the power of the microwave signal transmitted through the resonator are shown.

resonator with the sample. At 9 GHz an ESR spectrometer with field modulation and tracing of the magnetic-field derivative of the absorption line was used. The measurements were performed at temperatures 1.2–25 K in magnetic fields up to 60 kOe.

3. MAGNETIC RESONANCE SPECTRUM OF $CuGeO_3$

The ESR spectrum in $CuGeO_3$ consists of one line at temperatures above and near T_{SP}. As described in Refs. 11 and 12, the line broadens as temperature decreases. Near 5 K the spectrum splits and at lower temperatures contains four principal lines and several weak lines having different intensities relative to the principal lines at different frequencies and in different magnetic field orientations.

The low-temperature ESR signals are relatively weak. Using the known value of the molar susceptibility of $CuGeO_3$ at $T=15$ K[1] and the value of the quality factor, it is possible to find the effective density of paramagnetic defects which are responsible for the observed magnetic-resonance signal. Thus, for samples in the main series the total ESR signal intensity at the minimum at 5 K is 10^{-3} times the intensity of resonance absorption in a paramagnet with one $S=1/2$ spin per copper ion.

The properties of magnetic resonance are shown in Figs. 1–6. The change in the lineshape with temperature and the transformation of one line into four lines are illustrated in Fig. 1. It is evident here how a portion of the total intensity of the wide line splits off as temperature decreases and at $T=3.5$ K forms a new line to the right of the main line, while the main line splits into three components.

The magnetic-resonance lines recorded at different frequencies at the lowest temperature $T=1.3$ K with $\mathbf{H}\|\mathbf{c}$ are shown in Fig. 2. The four main lines, labeled 1, 2, 3, and 4, and several weaker lines, labeled by the letters $\alpha, \beta, \gamma, \epsilon$, and ν, are clearly seen. The magnetic-field dependence of the resonance-absorption frequencies $f_i(H)$ for $\mathbf{H}\|\mathbf{c}$ is illustrated in Fig. 3. The index i corresponds to one of the lines labeled in Fig. 2 by letters and numbers. The triplet of close lines 1, 2, 3 exhibits frequency–field dependence in the form of parallel straight lines. The central straight line, corresponding to line 2, passes through the origin. The magnetic-field-dependence of the resonance frequency for line 4 is a straight line with a different slope and passes through the origin. The resonance frequencies $f_{1,2,3,4}$ do not depend on temperature in the range 1.3–4 K. The data in Fig. 3 and the results of measurements with $\mathbf{H}\|\mathbf{a}$ and $\mathbf{H}\|\mathbf{b}$ show that the dependences $f_{1,2,3,4}(H)$ in the frequency range 9–75 GHz for rational directions have the form

$$f_i(H_\kappa)=\frac{\mu_B}{2\pi\hbar}g_{i\kappa}H_\kappa+d_{i\kappa}. \tag{2}$$

The index κ denotes one of the directions of the magnetic field along the a, b, or c axis. The values $g_{i\kappa}$ of the g factor and the constants $d_{i\kappa}$ are given in Table I. Nonzero

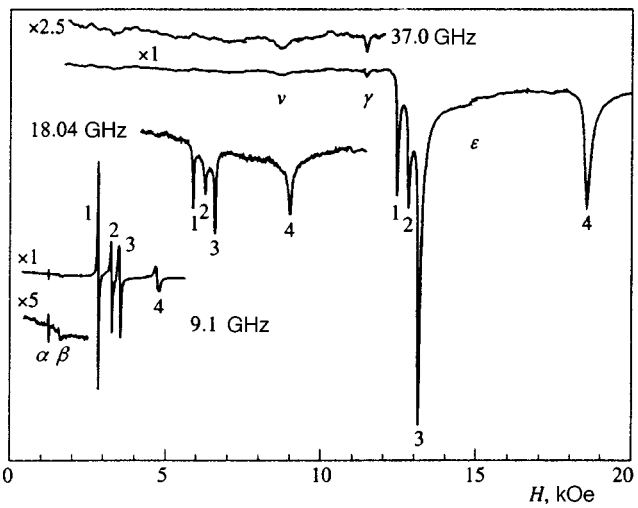


FIG. 2. Magnetic resonance lines of $CuGeO_3$ with $\mathbf{H}\|\mathbf{c}$. The lines were recorded at frequencies 37.0, 18.0, and 9.1 GHz at temperature 1.2 K. The derivative of the absorption line is shown for the frequency 9.1 GHz.

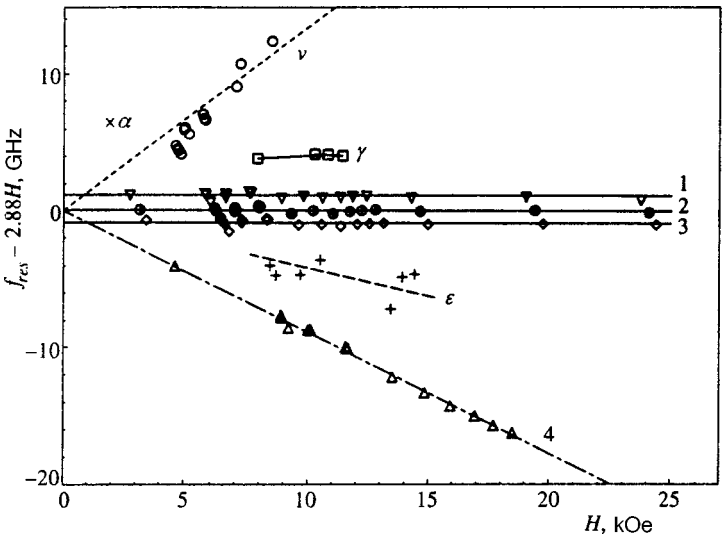


FIG. 3. ESR spectrum with $\mathbf{H} \| \mathbf{c}$, $T = 1.3$ K. The deviations of the resonance frequency from $f = 2.88H$, corresponding to the resonance frequency of line 2, are shown.

values of $d_{1\kappa}$ and $d_{3\kappa}$ correspond to zero-field splitting of the magnetic energy levels. The zero-field splitting is largest for $\mathbf{H} \| \mathbf{b}$ and vanishes for $\mathbf{H} \| \mathbf{a}$.

Representing the spectrum in the region of the main lines 1, 2, 3, and 4 as a superposition of four Lorentzian lines makes it necessary to introduce another, fifth, line with a g factor of approximately 2.0 and linewidth greater than 600 Oe. We shall designate this line by the number 0. The need to introduce such a line can be understood, for example, from Fig. 2, where one can see on the trace of the derivative of the absorption that the points of this curve that correspond to the resonance fields of the spectral components 1, 2, and 3 do not lie on the horizontal axis but rather are shifted upwards or downwards relative to it. This shift corresponds to the presence of another wider line. The intensity of line 0 at temperature 1.3 K is 0.07 times the total intensity of the lines 1, 2, 3, and 4 for the sample with $Q = 20$. For samples with $Q = 6$ line 0 becomes dominant.

Curves of the values of the resonance field versus its orientation are shown in Fig. 4. The lines 1 and 3 change places as the field rotates from the c to the b axis, so that the frequency differences $f_{1,3} - f_2$ change sign. As the field rotates from the c to the a axis, the lines 1, 2, and 3 merge into a single line. Line 4 corresponds to a strongly anisotropic g factor, varying from 1.43 to 1.86 depending on the angle between the magnetic field and the crystal axes.

It was found that the line α can be recorded only at the frequencies 9.1 and 9.4 GHz. The difference of the resonance fields for these frequencies shows that this line possesses zero frequency in zero field and corresponds to a g factor of 5.4 with the field oriented along the c axis. The line β demonstrates a direction-independent resonance value of the field with g factor 4.21, which is typical of the Fe^+ ion.[14]

Figure 5 shows the temperature dependence of the total intensity of the ESR spectrum and the lines 3 and 4 at 9.4 GHz. The temperature dependence of the linewidths is shown in Fig. 6. A peak having the width of all the lines is observed at temperature near 5 K, where the general line splits into four individual lines.

For comparison, the properties of magnetic resonance in

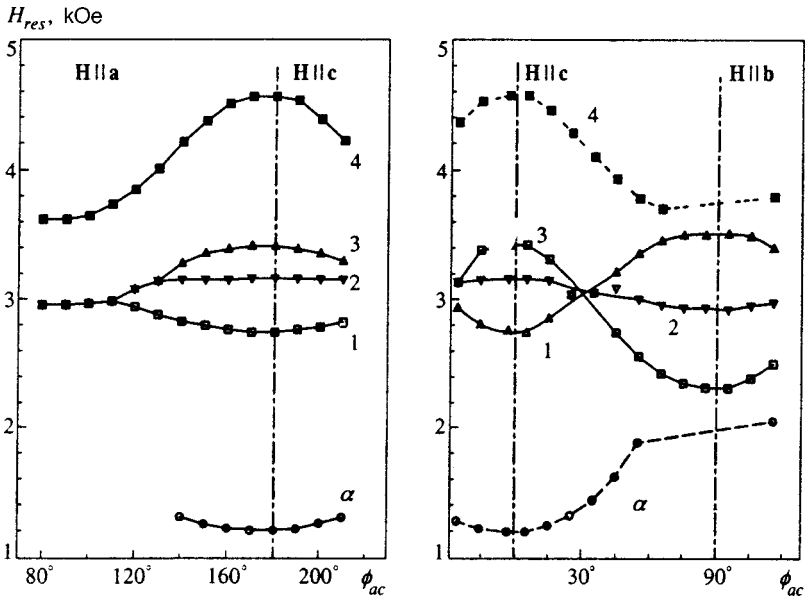


FIG. 4. Angular dependence of the ESR magnetic field at 9.4 GHz and $T = 1.5$ K for the field rotating in the ac and bc planes.

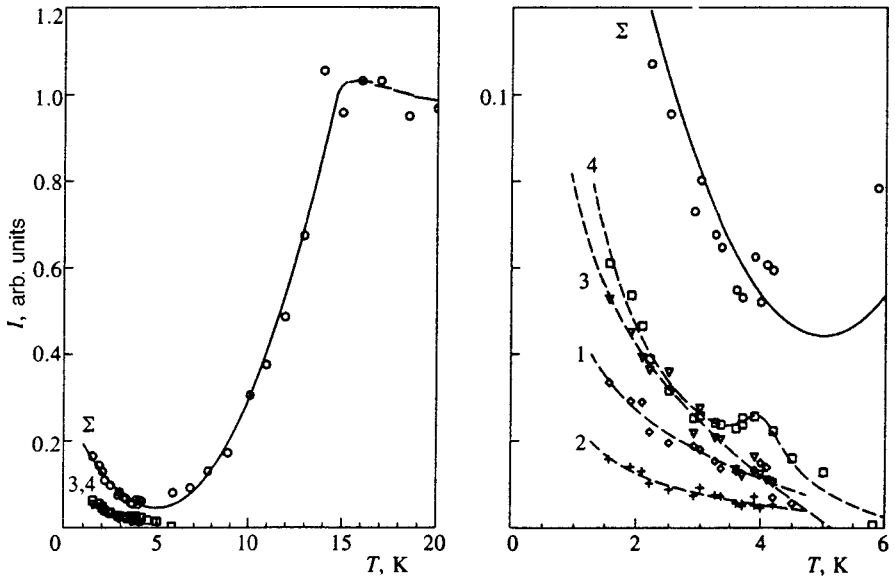


FIG. 5. Temperature dependence of the total intensity of ESR and the intensity of the lines 2, 3, 4 for frequency 9.4 GHz. The numbering of the resonance lines follows the designations in Fig. 2; the total intensity is denoted by the letter Σ.

samples of different quality are shown in Figs. 7 and 8 and in Table II. The intense line 0 is typical of samples with a low quality factor Q. In Table II this is illustrated by the values of the ratio Σ_{1-4}/I_0 of the total intensity of the lines 1, 2, 3, and 4 to the intensity of line 0 at $T=1.3$ K. Besides the data obtained in the present investigation, Table II also gives the values of Q and Σ_{1-4}/I_0 obtained from the Ref. 11. For samples with small Q the intensity of line 0 increases and the lines 1, 2, and 3 broaden and become indistinguishable in a crystal grown with a high cooling rate. Line 4 in samples with defects broadens and becomes more intense. Moreover, comparing the samples of the main series with sample No. 2 from Ref. 12 shows that line 2 is more intense in the low-Q sample. The lines 1 and 3 have the same intensity and linewidth as the pure sample. The intensity of the line α is close to that in purest sample, while the intensity of the line β is higher, in accordance with the iron concentration data. As one can see from Fig. 8, for samples grown from identical initial materials the lower the crystallization velocity, the lower the intensity of line 4 is. This attests to the fact that line 4 is due to structural defects in the magnetic subsystem of the copper ions and not to the impurity ions of other metals.

4. NONLINEAR MAGNETIC RESONANCE IN $CUGEO_3$

The high-frequency resonance susceptibility is power-dependent for line 4. Figure 9 shows that the magnetic resonance absorption lines at 20.2 GHz for different values of the microwave power. When the power increases above a certain threshold level, the additional high-frequency susceptibility increases. The intensity of line 4 becomes much greater than the total intensity of lines 1, 2, and 3. Line 4 becomes strongly asymmetric and the increase in the additional susceptibility on the low-field wing of the line stretches out more.

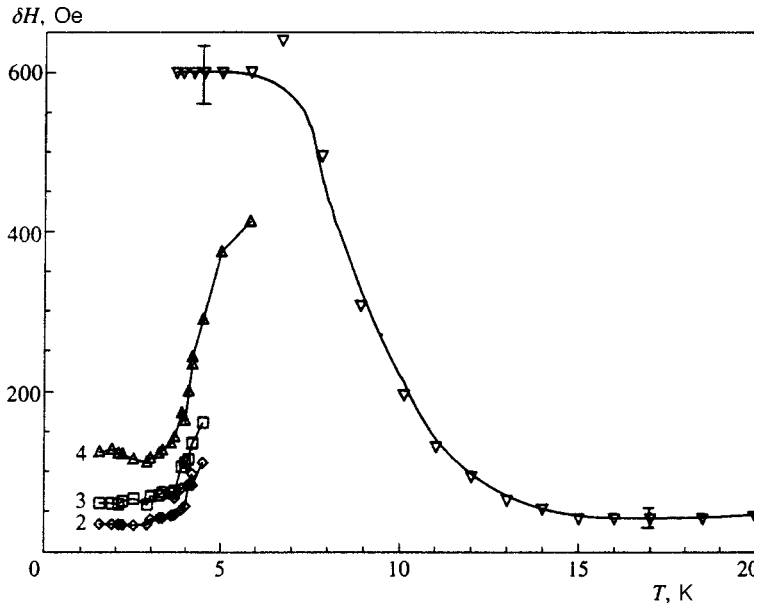


FIG. 6. Width of the magnetic-resonance line versus temperature at 9.4 GHz. The numbering of the lines follows the designations of the lines in Fig. 2.

TABLE I.

i	g_{ia}	g_{ib}	g_{ic}	d_{ia}, GHz	d_{ib}, GHz	d_i, GHz
1	2.17	2.26	2.10	0.0	−1.7	1.15
2	2.17	2.26	2.10	0.0	0.0	0.0
3	2.17	2.26	2.10	0.0	2.05	−0.95
4	1.82	1.86	1.43	0.0	0.0	0.0

TABLE II.

No.	v_{cr}, cm/h	$C_{Ni,Co,Mn}$	C_{Fe}	Q	Σ_{1-4}/I_0	T_{SP}	Reference
1	10^{-3}	$<10^{-4}$	$<10^{-4}$	20	15	14.5	
2	1	$<10^{-4}$	$<10^{-4}$	6	0.1	14.0	
3	6	$<10^{-4}$	$<10^{-4}$	3	0.01	13.0	
4	0.1	$<10^{-4}$	5×10^{-4}	7	1	14.5	12
5	unknown	unknown	unknown	100	20	14.5	11
6	1	5×10^{-3} (Ni)	$<10^{-4}$	2		12.5	

Figure 10 shows the power-dependence of the imaginary part of the susceptibility for a resonance magnetic field of line 4 and on the wing of this line. This dependence demonstrates the threshold character of the additional susceptibility. The threshold power levels marked in the figure by arrows correspond to microwave power of the order of 1 mW flowing into the resonator and to less than 10 μW absorption in the sample. The intensity of the microwave magnetic field at the sample is approximately 0.1 Oe.

The nonlinear increase in the imaginary part of the high-frequency susceptibility is greatest at a frequency of about 20 GHz. This effect is observed at 18 GHz and 23 GHz but the nonlinear growth of the susceptibility at comparable powers is approximately three times smaller. At all other frequencies which we employed (the closest is 26 GHz) the nonlinear increase of the susceptibility did not exceed the noise level.

5. DISCUSSION

5.1. Temperature evolution of the line shape

The temperature evolution of a magnetic resonance line as temperature decreases, i.e., the transformation of the line from a single narrow line first into a single wide line and then into four individual lines, distinct from one wide line, can be explained by taking account of the exchange interaction of paramagnetic centers, which are defects in the spin–Peierls phase, and thermally-activated triplet excitations.[11] The characteristic exchange-interaction frequency is found as the product of the corresponding exchange integral, expressed in frequency units, and the relative concentration of thermally-activated excitations.[15] When the characteristic exchange-interaction frequency is higher than the difference of the transition frequencies corresponding to different spectral lines, one resonance line is observed (exchange narrowing effect). As temperature decreases the concentration of thermally-activated triplets decreases and the exchange narrowing effect disappears, and as the exchange-interaction frequency passes through a value of the order of the difference of the transition frequencies individual lines appear which become narrower with further freezing of excitations.[16]

This scenario of the evolution of the line shape has been observed in TCNQ organic crystals,[16] which have a nonmagnetic ground state and triplet excited states. Broadening of the line and its splitting into two lines corresponding to triplet excitations with effective spin $S=1$ in a crystal field have been observed in pure crystals of these substances with decreasing temperature. A line corresponding to residual defects with effective spin $S=1/2$ has also been seen in experiments with irradiated TCNQ crystals.[17] The excited doublet with $S=1$ decreases in intensity as the temperature decreases. In crystals with defects the lines corresponding to $S=1/2$ and $S=1$ do not disappear as temperature decreases, demonstrating the temperature dependence of the intensity that is typical of paramagnetic impurities.

As shown in Ref. 11, the width of the ESR line in Cu GeO_3 in regions of both ''fast'' (above 5 K) and ''slow'' (below 5 K) exchange follows the theoretical formulas cor-

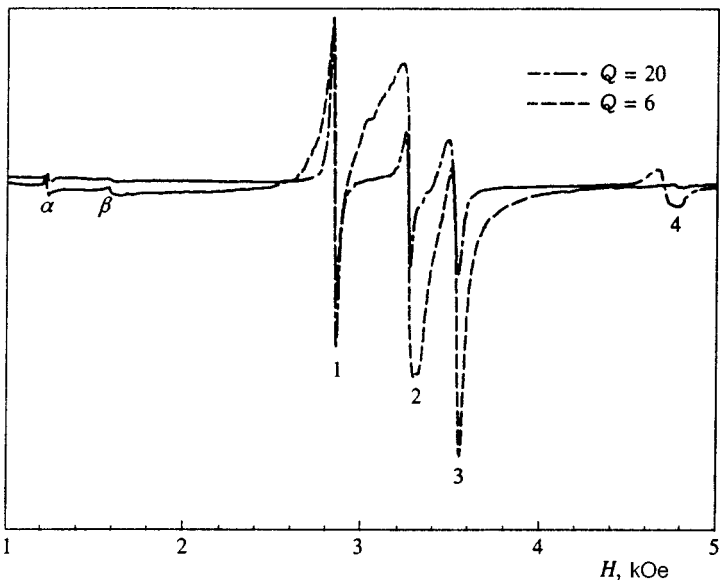


FIG. 7. Plots of the derivatives of the absorption at 9.1 GHz and $T=1.5$ K for two $CuGeO_3$ samples with $Q=20$ and $Q=6$. The samples differ by the Fe impurity content and the growth method. The amplitudes of the signals are normalized to the same intensity at $T=15$ K.

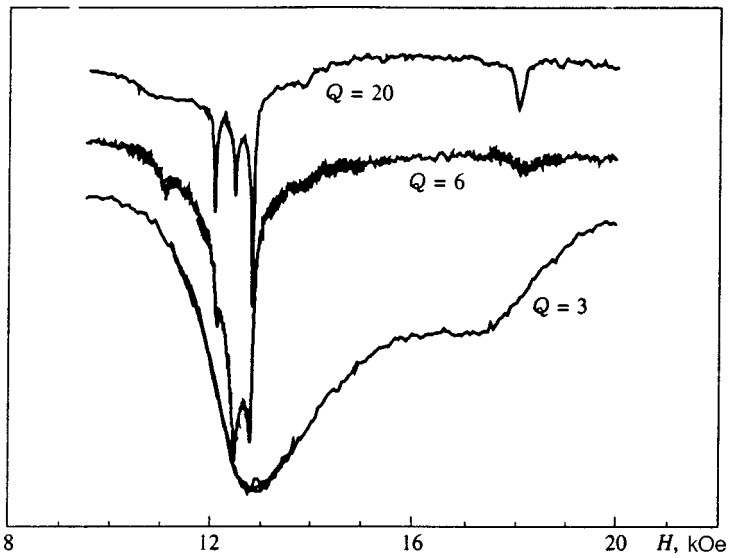


FIG. 8. ESR lines of samples prepared from the same batch with different crystallization velocities. The lines were recorded at $T=1.8$ K. The amplitudes of the signals are normalized to the same intensity at $T=15$ K.

responding to the above mechanism for the change in line shape with temperature.

5.2. Effective spin and origin of defects

At temperatures below 4 K the magnetic susceptibility is suppressed by a transition into the spin–Peierls phase, whose ground state is a singlet and nonmagnetic, while the excited states are separated from the ground state by a gap. In a pure defect-free crystal the magnetic susceptibility and the ESR signal intensity should approach zero exponentially at low temperatures. The nonzero susceptibility and the ESR signal are due to magnetic defects of the spin–Peierls phase. Evidently, the observed ESR signals belong to different types of such defects. In what follows we shall endeavor to identify them.

The g factors of the lines 1, 2, and 3 are close to the g factor of divalent copper in the paramagnetic phase of Cu GeO_3. This suggests that these lines are associated with copper ions. The g factor for line 4 differs strongly from the values typical of Cu^{2+} ions, equal to approximately 2.2. It can be inferred that this line is associated with impurities. However, analysis of the atomic composition gives for the concentration of magnetic impurities limiting values that correspond to a magnetic resonance line intensity less than half of that observed in our experiment. We studied the magnetic resonance spectrum in a $Cu_{0.995}Ni_{0.005}GeO_3$ single crystal. In this crystal the impurity ESR line corresponds to $g_a=1.92$, $g_b=2.00$, and $g_c=1.70$, i.e., its ESR frequency differs substantially from that of line 4. Therefore the presence of impurity nickel in the sample could not explain the appearance of this signal. Magnetic resonance in the presence of doping with nonmagnetic impurities reveals no change in the g factor of more than 3%.[18] Moreover, as indicated above, the difference in the intensity of line 4 for the samples grown with different crystallization velocities indicates that it originates from the copper ions. Therefore the results obtained indicate that the magnetic resonance lines 1, 2, 3, and 4 are associated with copper ions.

The relative intensities of line 2 in relation to lines 1 and 3 in samples of different quality differ (Fig. 7), while the

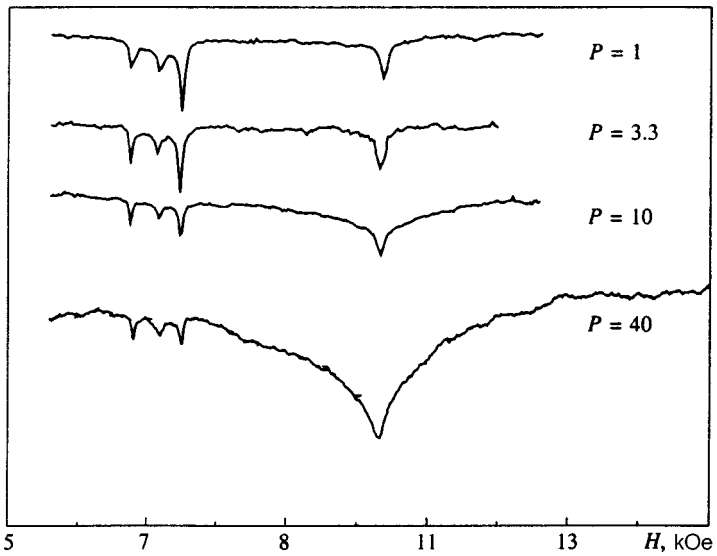


FIG. 9. Magnetic field dependence of the power of the microwave signal transmitted through the resonator for different power levels at temperature 1.2 K–1.6 K, $\mathbf{H}\parallel\mathbf{c}$, $f=20.2$ GHz. The power levels are in arbitrary units. The temperature variation within the indicated limits is due to the microwave heating of the resonator.

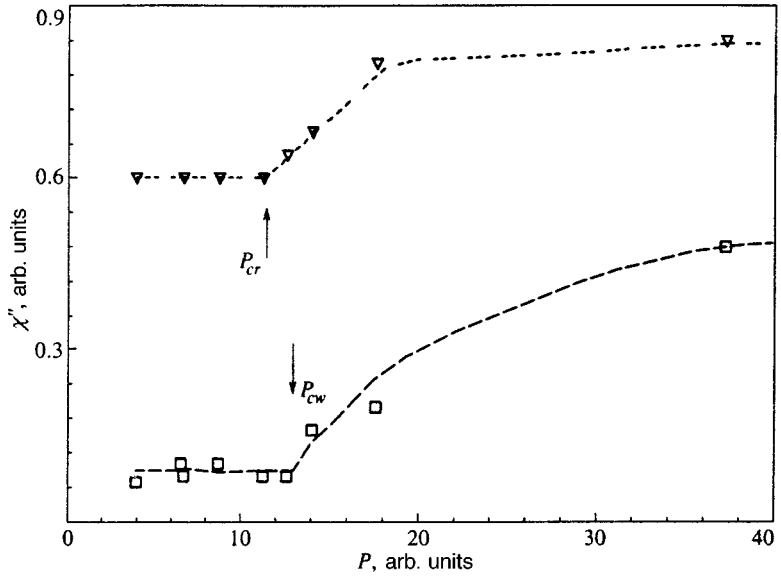


FIG. 10. Imaginary part of the susceptibility versus the microwave pump power for line 4 with a resonant field (triangles) and on the left wing of the line (squares); P_{cr} and P_{cw} are threshold powers at resonance and on the line wing, respectively; $T=1.5$ K, $\mathbf{H}\parallel\mathbf{c}$, and $f=20.2$ GHz.

ratios of the intensities of the lines 1 and 3 are the same. Hence it can be concluded that the lines 1 and 3 are associated with defects of one type while line 2 is associated with a defect of a different type. The absence of crystal-field splitting for line 2 attests to an effective spin $S=1/2$ for defects of this type (Ref. 19, Chap. I, § 5).

The zero-field splitting and the characteristic angular dependences of the lines 1 and 3 together with a change in their relative position with respect to line 2 imply that these lines are associated with defects with effective spin $S=1$. For $S=1$ the magnetic resonance typically splits in a crystal field into two lines with angular dependence similar to that which we observed (Fig. 5) and with a line spacing that is frequency-independent in the limit $g\mu_B H \gg D$, where D is the single-ion anisotropy constant of the spin Hamiltonian (Ref. 19, Chap. I, § 5). It is natural to infer that the lines 1 and 3 which we observed could be associated with exchange pairs of copper ions with effective spin $S=1$. The ESR line of these pairs can be split by a dipole–dipole interaction or anisotropic exchange (Ref. 19, Chap. 9, § 5). For example, such splitting has been observed in the ESR spectrum of copper acetate monohydrate (Ref. 19, Chap. 9, § 5). Here there is no need to introduce the dipole–dipole interaction or anisotropic exchange, since for Cu^{2+} ions with $S=1/2$ the splitting of the spectrum of an exchange pair by single-ion anisotropy is absent.

5.3. Intrinsic defects of the spin–Peierls phase

In our experiments the line 2 of spin-1/2 defects and the lines 1 and 3 of exchange-coupled defects have comparable intensities. For a random distribution of a small number of defects the resonance line of the exchange-coupled pairs should be much weaker than the line due to single defects. To resolve this discrepancy it is necessary to examine the structure of the magnetic defects in a spin–Peierls crystal which arise at a transition into the spin–Peierls phase.

As described in the Introduction, the low-temperature phase is characterized by one of two values of the dimerization phase. During a transition it is possible for crystallites (domains) with different values of this parameter to form so that on the domain wall, where $\psi=0$ switches to $\psi=\pi$, at least one atomic layer is nondimerized. Crystallite boundaries are usually pinned on defects which are present above the lattice rearrangement temperature, and therefore a single point defect of the crystal can give rise to an entire plane of magnetic defects. The boundaries of antiferromagnetic domains can serve as an example of the formation of such planar defects.[20] Figure 11 shows schematically the arrangement of copper ions in a sample containing two domains. The orientation of the planar sections of the wall between the domains is chosen to be close to three rational directions. A type-I boundary, lying in the plane ac, contains nondimerized chains which run along the c axis and are coupled by weak ferromagnetic exchange J_a in the a direction. Crossing this boundary destroys the order of the oxygen atom displacements. In the similar type-II boundaries (not shown in the figure) lying in the bc plane, the nondimerized chains of spins are coupled by a weak antiferromagnetic exchange J_b. In both cases there is a strong exchange interaction, characterized by the exchange integral J_c, within the chains located inside a domain wall. In type–III domain walls, lying in ab planes or in the {101} family of planes, the order of the dimerization of the copper ions breaks down. Walls of this type contain nondimerized spins from different chains. Weak exchange interactions J_a and J_b exist in the plane of such a wall.

A type-I wall contains disordered chains with strong in-chain antiferromagnetic exchange J_c. Weak ferromagnetic exchange acts in a direction perpendicular to the chains. The magnetic susceptibility of these chains is suppressed by strong exchange J_c and equals 1/200 of the susceptibility of the same number of paramagnetic spins at temperature 1.5 K. Structural defects in this wall (vacancies, dangling chain ends, or steps) should with probability of order 1 give rise to exchange pairs of copper ions with effective spin $S=1$ on account of the ferromagnetic exchange J_a with the spins in

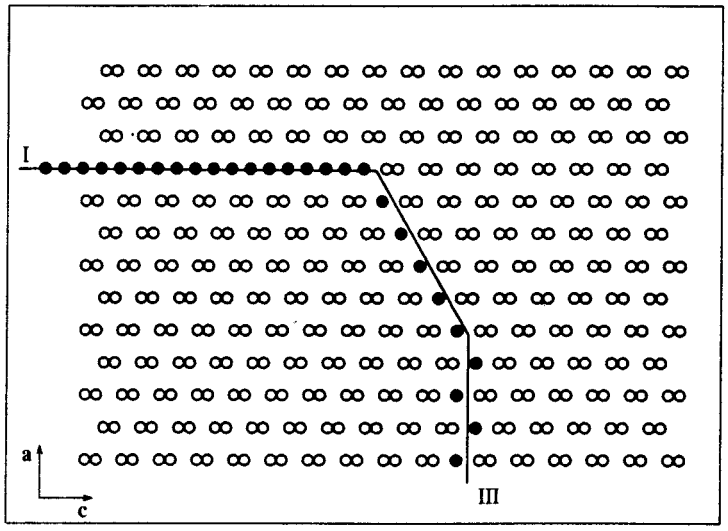


FIG. 11. Schematic diagram of dimerization domains and type-I and -III domain walls. The filled circles represent nondimerized copper ions.

neighboring chains. Breaks in chains in type-I walls are apparently the sources of the exchange-coupled pairs giving the resonance lines 1 and 3.

The angular and frequency dependences of the magnetic-resonance fields corresponding to lines 1 and 3 can be described by means of a spin Hamiltonian with $S=1$ (Ref. 20, Chap. I, § 5):

$$\mathscr{H}=g_c\mu_B H_c+g_b\mu_B H_b+g_a\mu_B H_a+D_c\hat{S}_c^2+D_b\hat{S}_b^2 + D_{cb}(\hat{S}_c\hat{S}_b+\hat{S}_b\hat{S}_c) \quad (3)$$

with $g_c=2.10$, $g_b=2.26$, $g_a=2.17$, $D_c=0.04$ K, $D_b=-0.05$ K, and $D_{cb}=-0.03$ K.

With overwhelming probability such defects in the main spin–Peierls matrix and type-II and -III walls either remain unpaired or form pairs with spin $S=0$ because of the stronger antiferromagnetic exchanges along the c and b axes.

It is natural to attribute the signal responsible for line 2 to isolated breaks in chains far from domain walls or in type-II walls. Free spins $S=1/2$ result from these breaks.

Splitting of the initial ESR line into three components as temperature decreases has also been observed in organic spin–Peierls crystals.[21] In this case an analysis of the angular and frequency dependence that would permit triplet and doublet states to be distinguished was not performed for the components of the signal. The authors interpreted the observed components as signals corresponding to different types of magnetic ions with different values of the g factor.

The typical behavior of the ESR line in a spin–Peierls magnet with one line splitting into several lines, including a triplet, is confirmed by the observation of this scenario of the evolution of the shape of the 36 GHz line in a second, recently discovered, inorganic spin–Peierls crystal, NaV_2O_5.[22]

5.4. Magnetic clusters in a spin–Peierls matrix

We shall now consider type–III domain walls, where nondimerized spins from different chains lie in the plane of the wall. According to the ideas developed in Refs. 7 and 10, each such spin is a center of a region, several lattice periods in size, with antiferromagnetically correlated spins in this chain fragment. Such a region (soliton) can be estimated on the basis of a theoretical calculation to contain approximately seven magnetic ions along the c axis. The total spin of the soliton is 1/2, while away from the nondimerized spin the average spin at a site decreases because of dimerization (Fig. 12).

These objects are unique magnetic clusters in a nonmagnetic spin–Peierls matrix. Magnetic resonance of these objects, which possess internal magnetic structure and corresponding internal degrees of freedom, is an urgent problem. Among the known related magnetic-resonance problems, we single out the problem of a three-spin cluster of identical magnetic ions with spin 1/2.[24] For a cluster in the form of an isosceles triangle, differing very little from an equilateral triangle, it has been shown that the energy levels have the form

$$E_{1,2,3,4}=\pm\frac{1}{2}\sqrt{G^2+\delta^2+g^2\mu_B^2H^2\pm 2g\mu_B H\sqrt{\delta^2+G^2\cos^2\theta}}. \quad (4)$$

Here θ is the angle between the symmetry axis z of the crystal and the magnetic field, $J_{0,1}$ are the exchange integrals, $\delta=|J_1-J_0|$, and G is the Dzyaloshinskiĭ–Moriya antisymmetric exchange constant. The constant G is zero if a center of symmetry exists between the magnetic ions in a pair. The transition energies between the levels (2), which are small compared with δ and G, satisfy

$$\hbar\omega=\frac{g}{\sqrt{1+G^2\cos^2\theta/\delta^2}}\mu_B H. \quad (5)$$

Thus, the magnetic-resonance spectrum of a triangular cluster at low frequencies looks like a paramagnetic resonance spectrum with a strongly anisotropic effective g factor and $g_z<2$. Experiments with crystals of organic complexes containing triads of copper ions show that the observed magnetic resonances[24] and the static[25] properties correspond to this scenario with nonzero G.

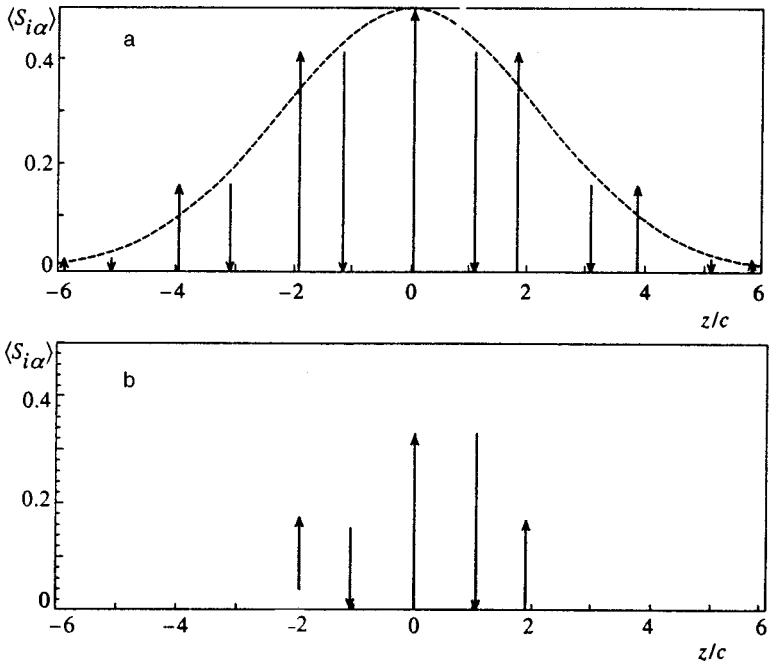


FIG. 12. Inferred distribution of the average values of the spin projection near a dimerization defect (a) and the result of a calculation of the average values of the projection for a five-spin cluster with symmetric and antisymmetric exchange (b).

For solitons localized on a domain wall there is no center of symmetry between the central and neighboring spins, since the environment around the central ion is closer to that existing in the nondimerized phase, while the environment around the neighboring ions is closer to that of ions in the dimerized phase. Therefore the symmetry admits a nonzero value of G.

To describe a linear magnetic cluster in a spin–Peierls nonmagnetic matrix near a nondimerized spin, we shall examine the three- and five-spin approximations. For the three-spin linear cluster we write the Hamiltonian in the form

$$\mathscr{H}=J_{12}\hat{\mathbf{S}}_1\cdot\hat{\mathbf{S}}_2+J_{23}\hat{\mathbf{S}}_2\cdot\hat{\mathbf{S}}_3+J_{13}\hat{\mathbf{S}}_1\cdot\hat{\mathbf{S}}_3+\mathbf{G}_{12}\hat{\mathbf{S}}_1\times\hat{\mathbf{S}}_2+\mathbf{G}_{23}\cdot\hat{\mathbf{S}}_2\times\hat{\mathbf{S}}_3, \tag{6}$$

where $\mathbf{G}_{12}$ and $\mathbf{G}_{23}$ are vector parameters of the Dzyaloshinskiĭ–Moriya interaction. The energies of the two lowest states of the cluster for an arbitrary orientation of the external magnetic field are given by

$$E_{1,2}=\frac{\varepsilon_1+\varepsilon_2}{2}-\frac{1}{2}\sqrt{(\varepsilon_1-\varepsilon_2)^2+\mathbf{G}^2+\mathbf{h}^2\pm2\sqrt{(\varepsilon_1-\varepsilon_2)^2\mathbf{h}^2+(\mathbf{hG})^2}}. \tag{7}$$

Here ε_1 and ε_2 are the energies of two possible states of a cluster with total spin $S=1/2$ in the absence of an antisymmetric exchange interaction and a magnetic field

$$\varepsilon_{1,2}=-\frac{1}{4}(J_{12}+J_{13}+J_{23})\pm\frac{1}{2}\sqrt{\left[J_{23}-\frac{1}{2}(J_{12}+J_{13})\right]^2+\frac{3}{4}(J_{12}-J_{13})^2}, \tag{8}$$

where $\mathbf{G}$ and $\mathbf{h}$ denote the vectors $\mathbf{G}=(\mathbf{G}_{12}+\mathbf{G}_{23})/\sqrt{3}$ and $\mathbf{h}=g\mu_B\mathbf{H}$, where g are the corresponding components of the g tensor of a single copper ion.

For $\mathbf{h}\perp\mathbf{G}$ the effective g factor is given by

$$g_{\text{eff}}\approx g\left[1-\frac{G_x^2+G_y^2}{2(\varepsilon_2-\varepsilon_1)^2}\right]. \tag{9}$$

Setting $J_{23}=J_{12}=10$ meV (Ref. 3) and $J_{13}=3.6$ meV (Ref. 27, we find that to explain the observed value $g=1.43$ it must be assumed that there exists a vector $\mathbf{G}$ perpendicular to the plane of the fragment CuO_2-CuO_2-Cu with $|G_{12}+G_{23}|\approx8$ meV. We note that $\mathbf{G}=0$ in a cluster which is symmetric with respect to the central spin. However, as one can see from Fig. 11, because of the distortions of dimerization on neighboring pairs of copper ions there is no center of symmetry on nondimerized spins of a domain wall lying in the {101} planes.

This estimate gives a value for the antisymmetric exchange constant on the order of the exchange interaction. Switching to a model with five spins (Cu5–Cu4–Cu1–Cu2–Cu3) it is easy to see that the number of excited states of a cluster with total spin $S=1/2$, which are admixed to the main doublet by antisymmetric exchange, increases. In this connection expression (9) becomes

$$g_{\text{eff}}=g\left[1-\frac{(\mathbf{G}_{12}+\mathbf{G}_{23}+\mathbf{G}_{41}+\mathbf{G}_{54})^2}{3E_{31}^2}-\frac{(\mathbf{G}_{12}+\mathbf{G}_{23}-\mathbf{G}_{41}-\mathbf{G}_{54})^2}{12E_{21}^2}\right], \tag{10}$$

where E_{31}^2 and E_{21}^2 are given by the expressions

$$E_{21}^2=\left[J_{23}-\frac{1}{2}(J_{12}+J_{13})+\frac{1}{4}J_{24}\right]^2+\frac{3}{2}(J_{12}-J_{13})^2, \tag{11}$$

$$E_{31}^2=\left\{J_{23}-\frac{1}{2}(J_{12}+J_{13})-\frac{3}{4}J_{24}+\left[\left(J_{23}-\frac{1}{2}(J_{12}+J_{13})+\frac{1}{4}J_{24}\right)^2+\frac{3}{2}(J_{12}-J_{13})^2\right]^{1/2}\right\}^2+\frac{3}{2}(J_{12}-J_{13})^2. \quad (12)$$

We note that only the antisymmetric combination $G_{12}+G_{23}$ of the parameters enters in Eq. (9), while in Eq. (10) a contribution from the symmetric combination $G_{12}+G_{23}-G_{41}-G_{54}$ also appeared. This fundamental difference arises because for a three-spin cluster the excited state with total spin $S=1/2$ is unique and antisymmetric. Among the excited states of a five-spin cluster with total spin $S=1/2$ there are both antisymmetric and symmetric (relative to the center of the cluster) states. Consequently, both combinations contribute to the change in the g factor: antisymmetric $G_{12}+G_{23}+G_{41}+G_{54}$ and symmetric $G_{12}+G_{23}-G_{41}-G_{54}$. For a five-spin cluster there are only four excited states with $S=1/2$. Only the two lowest of these states are taken into account in Eq. (10), since otherwise the formula becomes very complicated. A calculation assuming $G_{12}=G_{23}=G_{54}=G_{41}$ gave $G_{12}=2.9$ meV, for which $g_c=1.5$. The following numerical values were used here: $J_{12}=J_{14}=10$ meV, $J_{23}=J_{45}=10.4$ meV (dimerized pairs of spins), and $J_{13}=J_{15}=3.6$ meV.

Numerical calculations performed by exact diagonalization of the energy matrices in a magnetic field, taking account of all excited states in a cluster of five spins, give $G_{12}=3.0$ meV. The average values of the projection of the spins of a cluster in a particular direction, obtained by means of the present calculation using the numerical values indicated above, are shown in Fig. 12. The assumption that antisymmetric exchange is present in $CuGeO_3$ with the Dzyaloshinskiĭ vector perpendicular to the c axis, was first advanced in Ref. 27 on the basis of an analysis of the reasons for the broadening of the ESR lines. As stressed in Ref. 27, this assumption contradicts the data of Ref. 2 on the crystal structure of CuO_2 chains. In this connection, it should be stated that the symmetry is destroyed near nondimerized spins in a type–III domain wall and the existence of a nonzero Dzyaloshinskiĭ–Moriya interaction constant becomes allowed, in any case below the temperature T_{SP}.

Evidently, the line 4 must be attributed to defects of the latter type if we stay within the magnetic subsystem formed by the copper ions. These considerations make it possible to explain the marked departure of the g factor from the value 2.0 and the large anisotropy of the g factor for the line 4 by analyzing a cluster of five spins, taking account of the antisymmetric exchange interaction in which the parameter is of order 30% of the exchange integral.

Of course, these considerations concerning the magnetic resonance of a cluster whose internal structure is determined by the Dzyaloshinskiĭ–Moriya interaction give only a qualitative explanation of the pronounced departure of the g factor from its typical value exhibited by copper ions. The following discrepancies exist between this model and the experimental results. In the first place, the value of the g factor in other orientations does not approach 2.0, as follows from the proposed model. In the second place, a strong decrease of the g factor should be observed for all directions of the field which lie in a plane perpendicular to **G** and not only in the direction c. These discrepancies could be due to the presence of antisymmetric exchange for the interaction of the next-to-nearest neighbors and misalignment of the Dzyaloshinskiĭ vectors of different pairs of ions.

The ESR at 9.31 GHz in a $CuGeO_3$ crystal of even higher quality was studied in Ref. 11. From the data presented by the authors, Q is estimated to approach 100. The lines α, 1, 2, 3, and 4 were also observed, and the values of their resonance field agree well with the data obtained in the present work (Fig. 2). This confirms that the observed spectrum, consisting of a triplet line and a line due to clusters, is typical of pure spin–Peierls crystals.

5.5. Two-dimensional magnet on domain walls of a spin–Peierls phase

A frozen soliton can be viewed as a magnetic quasiatom, since its structure with antiferromagnetic correlation of neighboring spins is fixed by the strong exchange J_c. Then a wall of the latter type is a two-dimensional magnet consisting of the described quasiatoms, coupled by ferromagnetic exchange NJ_a in the direction a and antiferromagnetic exchange NJ_b along the b axis. Here N is the effective number of spins in the quasiatom. Using the estimate $N\approx 5$, we obtain 50 K and -5 K for the values of the exchange of quasiatoms along the b axis and a axis, respectively. Because of the presence of anisotropy with characteristic energy $E_a=0.5$ K per quasiatom, it can be expected that such a two-dimensional magnet becomes ordered at a temperature $T_c\approx N\sqrt{J_bJ_a}/\ln(NJ_b/E_a)\approx 3$ K.

Thus, at the temperature of our experiment, where $T\ll NJ_b$ holds, this flat object can be a disordered but strongly correlated magnet. Its excitation spectrum in the long-wavelength region is similar to that of an antiferromagnet without anisotropy.[28] One branch of this spectrum, $\omega_{1\mathbf{k}}$, has no gap in a magnetic field. The second branch has a gap $\omega_{20}=g_\alpha\mu_B H_\alpha$. In this case, excitation of a uniform mode of oscillations with frequency ω_{20} is possible in a homogeneous high-frequency magnetic field. Since the precession of the magnetic moment is elliptical because of the anisotropy, parametric excitation of pairs of waves of the gapless mode, which satisfy the condition of paramagnetic resonance

$$\omega_{mw}=\omega_{1,\mathbf{k}}+\omega_{1,-\mathbf{k}}, \quad (13)$$

becomes possible. Here ω_{mw} is the frequency of the high-frequency pump field. The absorption of energy in the process of parametric excitation has a threshold with respect to the pump power and the energy flowing from the pump to the wave modes is resonant with respect to the pump frequency (see, for example, Refs. 29 and 30). The resonant character is due to energy transfer via oscillations of the magnetization of one of the magnetic-resonance modes.

Parametric excitation of spin waves belonging to the branch that possesses a gap in the presence of a magnetic

field, according to the parametric resonance condition, should give rise to the absorption of energy of the high-frequency pump in fields below the magnetic resonance field at half the frequency. Absorption bands such as, for example, in Figs. 2 and 3 (line ν), were observed in our experiments.

The observed nonlinear threshold increase in the susceptibility for line 4 allows one to conclude that two-dimensional flat magnetic defects are present in a spin–Peierls matrix. Point defects could not produce a nonlinear effect with a threshold switching-on of the susceptibility. It is well-known that magnetic resonance saturates for them, the imaginary part of the susceptibility decreasing as the power increases. The existence of three-dimensionally correlated regions is unlikely, since in this case the ordering temperature should be higher, of order $\sqrt{J_c J_b} \geqslant 10$ K, which will give rise to appreciable gaps for both branches of the spectrum. This is incompatible with the linear and gapless dependence $f_4(H)$.

6. CONCLUSIONS

Spin-1/2 and -1 defects of a spin–Peierls phase, which are formed by breaks in chains of copper ions in domain walls of the spin–Peierls phase and by exchange-coupled pairs of these defects, were identified by analyzing the magnetic resonance spectra of pure crystal of a spin–Peierls magnet at temperatures substantially below the transition temperature.

An additional magnetic-resonance signal confirming the presence of two-dimensional magnetic domains with long-range correlation of the spins was observed. Such regions in a spin–Peierls phase are presumed to be the walls of domains with a different value of the dimerization phase.

We thank L. A. Prozorova, V. A. Atsarkin, V. V. Demidov, I. A. Zaliznyak, V. I. Marchenko, S. S. Sosin, L. E. Svistov, and H. Benner for many discussions, H.-A. Krugg von Nidda for assisting in the measurements at 9 GHz, and Yu. M. Tsipenyuk, V. I. Firsov, and M. Fedoroff for analyzing the atomic composition of the samples.

This work was supported by the Russian Fund for Fundamental Research (Project No. 98-02-16572) and CIDF (Grant No. RP 1-207). The measurements at 9 GHz were performed at Technische Hochschule Darmstadt (Germany) with the support of Sonderforschungbereich (No. 185).

[1] M. Hase, I. Terasaki, and K. Uchinokura, Phys. Rev. Lett. **70**, 3651 (1993).
[2] K. Hirota, D. E. Cox, J. E. Lorenzo, G. Shirane, J. M. Tranquada, M. Hase, K. Uchinokura, H. Kojima, Y. Shibuya, I. Tanaka, Phys. Rev. Lett. **73**, 736 (1994).
[3] L.-P. Regnault, M. Aïn, B. Hennion, G. Dhalenne, and A. Revcholevschi, Phys. Rev. B **53**, 5579 (1996).
[4] J. W. Bray, H. Hart, L. Interrante, I. Jacobs, J. Kasper, G. Watkins, S. Wee, and J. Bonner, Phys. Rev. Lett. **35**, 744 (1975).
[5] L. N. Bulaevskiĭ, Fiz. Tverd. Tela (Leningrad) **11**, 1132 (1969) [Sov. Phys. Solid State **11**, 921 (1969)].
[6] H. Bethe, Z. Phys. **71**, 205 (1931).
[7] D. Khomskii, W. Geertsma, and M. Mostovoy, Chech. J. Phys. **46**, Suppl. S **6**, 3239 (1996).
[8] J. G. Lussier, S. M. Coad, D. F. McMorrow, and D. Paul, J. Phys.: Condens. Matter **7**, L325 (1995).
[9] L.-P. Regnault, J.-P. Renard, G. Dhallenne, and A. Revcholevschi, Europhys. Lett. **32**, 579 (1995).
[10] H. Fukuyama, T. Tanimoto, and M. Saito, J. Phys. Soc. Jpn. **65**, 1182 (1996).
[11] M. Honda, T. Shibata, K. Kindo, Sh. Sugai, T. Takeuchi, and H. Hori, J. Phys. Soc. Jpn. **65**, 691 (1996).
[12] A. I. Smirnov, V. N. Glazkov, A. N. Vasil'ev, S. Coad, D. McK Paul, G. Dhalenne, and A. Revcolevschi, JETP Lett. **64**, 305 (1996).
[13] A. I. Smirnov, A. N. Vasil'ev, and L. I. Leonyuk, JETP Lett. **64**, 695 (1996).
[14] S. A. Al'tshuler and B. M. Kozyrev, *Electron Paramagnetic Resonance in Compounds of Transition Elements*, Halsted, New York, 1975 [Russian original, Nauka, Moscow, 1972 p. 429].
[15] M. T. Jones and D. B. Chesnut, J. Chem. Phys. **38**, 1311 (1963).
[16] D. B. Chesnut and W. D. Philips, J. Chem. Phys. **35**, 1002 (1961).
[17] H. M. McConnell, H. O. Griffith, and D. Pooley, J. Chem. Phys. **36**, 2518 (1962).
[18] P. Fronzes, M. Poirier, A. Revcolevschi, and G. Dhalenne, Phys. Rev. B **56**, 7827 (1997).
[19] A. Abragam and B. Bleaney, *Electron Paramagnetic Resonance of Transition Ions*, Clarendon Press, Oxford, 1970 [Russian translation, Mir, Moscow, 1972, Vol. 1].
[20] D. N. Astrov, Zh. Éksp. Teor. Fiz. **38**, 984 (1960) [Sov. Phys. JETP **11**, 708 (1960)].
[21] I. S. Jacobs, J. W. Bray, H. R. Hart, Jr., L. V. Interrante, J. S. Kasper, G. D. Watkins, D. E. Prober, and J. C. Bonner, Phys. Rev. B **14**, 3036 (1976).
[22] A. N. Vasil'ev, A. I. Smirnov, M. Isobe, and Y. Ueda, Phys. Rev. B **56**, 5065 (1997).
[23] M. I. Belinskiĭ, B. S. Tsukerblat, and A. V. Ablov, Fiz. Tverd. Tela (Leningrad) **16**, 989 (1974) [Sov. Phys. Solid State **16**, 639 (1974)].
[24] Yu. V. Yablokov, V. K. Voronkova, and L. V. Mosina, *Electron Paramagnetic Resonance of Exchange Clusters*, Nauka, Moscow, 1988; Yu. K. Yablokov, B. Ya. Kuyavskaya, A. V. Ablov, L. V. Mosina, and M. D. Mazus, Dokl. Akad. Nauk SSSR **256**, 1182 (1981).
[25] B. S. Tsukerblat, V. M. Novotvortsev, B. Ya. Kuyavskaya, M. I. Belinskiĭ, A. V. Ablov, A. N. Bazhan, and V. T. Kalinnikov, JETP Lett. **19**, 277 (1974).
[26] J. Riera and A. Dobry, Phys. Rev. B **51**, 16098 (1995).
[27] I. Yamada, M. Nishi, and J. Akimitsu, J. Phys.: Condens. Matter **8**, 2625 (1996).
[28] G. Mueller, H. Thomas, H. Beck, and J. C. Bonner, Phys. Rev. B **24**, 1429 (1981).
[29] V. I. Ozhogin, Zh. Éksp. Teor. Fiz. **58**, 2079 (1970) [Sov. Phys. JETP **31**, 1121 (1970)].
[30] V. V. Kveder, B. Ya. Kotyuzhanskiĭ, and L. A. Prozorova, Zh. Éksp. Teor. Fiz. **63**, 2205 (1972) [Sov. Phys. JETP **36**, 1165 (1972)].

Translated by M. E. Alferieff

Diffusive Toda chains in models of nonlinear waves in active media

V. M. Zhuravlev*)

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia
(Submitted 14 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1897–1914 (November 1998)

A class of models of autowaves in the form of nonlinear diffusion equations, which are closely related to the Liouville equation and two-dimensionalized Toda chains, is investigated. Exact solutions of these equations are constructed and analyzed. A simple method for constructing diffusive Toda chain models from known basic models is proposed. © *1998 American Institute of Physics.* [S1063-7761(98)02611-0]

Soliton models based on the inverse scattering problem presently comprise one of the most important classes of models of nonlinear wave processes. However, despite the broad applicability of this method in applied problems, there are some areas of physics where the nature of the solutions obtained with its aid (solitons, kinks, breathers, etc.) cannot be considered consistent with the nature of the processes studied. This applies, for example, to nonlinear diffusion processes, which have recently attracted special attention in connection with the development of such areas of physics as the theory of self-organizing systems, the theory of wave processes in active media, etc.[1–5] New approaches to constructing basic models having fairly rich classes of exact solutions are needed to describe these processes.

As a rule, autowave models are investigated by approximate analytical methods[4–7] or numerical simulation.[5] However, the significance of the conclusions which can be drawn using these methods depends largely on the possibility of testing the reliability of the calculated results against the exact solutions of the equations. Just such a test reveals whether the simplifications and approximations made remain within the original model. Therefore, the lack of well developed methods for constructing exact solutions for models with diffusion should be regarded as an obstacle to a thorough understanding of the nonlinear wave processes which take place in such systems and are associated with the appearance of regular or coherent structures in them.

An entire class of models of nonlinear wave processes in active media with diffusion that can be solved exactly was proposed in Ref. 8. Such models represent a generalization of two-dimensionalized Toda chains[9–11] to the case of two-dimensional diffusion processes. The class of exact solutions obtained in Ref. 8 is a new useful formulation of the solutions of two-dimensionalized Toda chains, which were first considered in Ref. 11 and 12, and the Liouville equation[13] in a Hermitian form of order 2. We also note that new integrable Toda chains belonging to a class of diffusion equations in one-dimensional space were found in Ref. 14 via a symmetry approach.

The method proposed in Ref. 8 for constructing and analyzing models which can be termed diffusive Toda chains is developed in this paper. The properties of these models, which lead to the appearance of regular structures in diffusion systems, and the principles which can be used to obtain diffusive Toda chain models from known basic models are also considered. As in Ref. 8, the method for constructing solutions is based on a novel reformulation of the method for constructing exact solutions of two-dimensionalized Toda chains using quadratic forms.[8,11,12] Exact solutions of some basic reaction–diffusion models are constructed on the basis of this approach.

1. GENERAL CONSTRUCTION OF SOLUTIONS OF TWO-DIMENSIONALIZED TODA CHAINS IN A CLASS OF QUADRATIC FORMS

Two-dimensionalized Toda chains comprise a physical system, whose dynamics are described by a set of differential equations in partial derivatives of the following form:

$$\Delta\Phi_n=\exp\{\Phi_{n-1}-2\Phi_n+\Phi_{n+1}\},\quad n=1,\ldots,N-1, \tag{1}$$

where $\Delta=\partial^2/\partial z\partial\bar{z}$ is the Laplacian operator in the case of the complex coordinates $z=x+iy$ and $\bar{z}=x-iy$ [or the D'Alembertian operator in the case of the real coordinates $z=x+y$ and $\bar{z}=x-y$ (conical coordinates)]. The system for the variables $G_n=\Phi_n-\Phi_{n-1}$ has a more familiar form:

$$\Delta G_n=\exp\{G_{n+1}-G_n\}-\exp\{G_n-G_{n-1}\},$$
$$n=1,\ldots,N-1. \tag{2}$$

There is a definition for generalized diffusive Toda chains (see, for example, Refs. 10 and 15), but they will not be considered in this paper.

Let us briefly describe the basic principles for constructing the exact solutions of two-dimensionalized Toda chains using quadratic forms. Let us consider a function $H(z,\bar{z})$ of the following form

$$H_1(z,\bar{z})=\sum_{\alpha,\beta=1}^{N}h^{\alpha\beta}\psi_\alpha(z)\psi^*_\beta(\bar{z}),$$

where $h^{\alpha\beta}$ is an $N\times N$ Hermitian matrix. The function $H_1(z,\bar{z})$ is an Hermitian form of order N, which is defined in the complex space $\mathbf{F}^N$. The set of the functions $\psi_\alpha(z)$, which depend on z, can be represented in the form of the vector $\mathbf{\Psi}=\{\psi_1,\psi_2,\ldots,\psi_N\}$ in the space $\mathbf{F}^N$, in which the length

of the vector is assigned by $H_1(z,\bar{z})$. For the sake of brevity, it is useful to replace the complex-conjugate functions $\psi^*_\beta(\bar{z})$ by functions with superscripts according to the rule

$$\psi^\alpha(\bar{z})=\sum_{\beta=1}^{N} h^{\alpha\beta}\psi^*_\beta(\bar{z}).$$

Then for any nondegenerate form $h^{\alpha\beta}$ the function $H_1(z,\bar{z})$ can be represented as

$$H_1(z,\bar{z})=\sum_{\alpha=1}^{N}\psi_\alpha(z)\psi^\alpha(\bar{z})=\boldsymbol{\Psi}\bar{\boldsymbol{\Psi}},$$

where $\bar{\boldsymbol{\Psi}}=\{\psi^1,\psi^2,\ldots,\psi^N\}$.

For convenience, we introduce the notation

$$[f(z),g(z)]\equiv f(z)\frac{dg(z)}{dz}-g(z)\frac{df(z)}{dz}=fg'-gf',$$

$$f'=\frac{df(z)}{dz},\quad g'=\frac{dg(z)}{dz} \tag{3}$$

for the two arbitrary functions $f(z)$ and $g(z)$. Then, it can be shown by direct calculations that for any fixed $N>1$

$$H_1^2(z,\bar{z})\Delta\ln H_1(z,\bar{z})=H_2(z,\bar{z}), \tag{4}$$

where

$$H_2(z,\bar{z})=\sum_{\alpha<\beta=1}^{N} w_{\alpha\beta}(z)w^{\alpha\beta}(\bar{z}),$$

$$w_{\alpha\beta}=[\psi_\alpha(z),\psi_\beta(z)],\quad \alpha\neq\beta.$$

In fact,

$$\begin{aligned}H_1^2(z,\bar{z})\Delta\ln H_1(z,\bar{z})&=H_1(z,\bar{z})\Delta H_1(z,\bar{z})\\&\quad-H_{1,z}(z,\bar{z})H_{1,\bar{z}}(z,\bar{z})\\&=(\boldsymbol{\Psi}\bar{\boldsymbol{\Psi}})\Delta(\boldsymbol{\Psi}\bar{\boldsymbol{\Psi}})-(\boldsymbol{\Psi}\bar{\boldsymbol{\Psi}})_z(\boldsymbol{\Psi}\bar{\boldsymbol{\Psi}})_{\bar{z}}\\&=\sum_{\alpha,\beta=1}^{N}[\psi_\alpha(z)\psi^\alpha(\bar{z})\psi'_\beta(z)\psi'^\beta(\bar{z})\\&\quad-\psi'_\alpha(z)\psi^\alpha(\bar{z})\psi_\beta(z)\psi'^\beta(\bar{z})]\\&=\sum_{\alpha<\beta=1}^{N}\{[\psi_\alpha(z)\psi'_\beta(z)-\psi'_\alpha(z)\psi_\beta(z)]\\&\quad\times[\psi^\alpha(\bar{z})\psi'^\beta(\bar{z})-\psi^\alpha(\bar{z})\psi'^\beta(\bar{z})]\}\\&=\sum_{\alpha<\beta=1}^{N}w_{\alpha\beta}(z)w^{\alpha\beta}(\bar{z}),\end{aligned} \tag{5}$$

which was to be proved. The function $H_2(z,\bar{z})$ is also an Hermitian form, but of dimension $N(N-1)/2$. Thus, the same transformation (4) can be applied to it. As a result of the successive application of this transformation to the sequence of Hermitian forms appearing in each step, we obtain the chain of functions $H_n(z,\bar{z})$ $(n=1,2,\ldots,N)$ which is cut off in the Nth step. These functions satisfy Eqs. (1). No constraints, apart from sufficient smoothness and differentiability, are imposed on the functions $\psi_i(z)$. This result was formulated and proved in a somewhat different form in Refs. 11 and 12. The lack of constraints on the functions $\psi_i(z)$ allows us to consider various reductions and to impose additional conditions on them in order to satisfy more general equations. For example, such equations can be the equations of reaction–diffusion models with nonlinear sources.

2. CLASSIFICATION OF DIFFUSIVE TODA CHAIN MODELS

The general form of the models investigated in the present work is

$$\frac{d\mathbf{U}}{dt}=\mathbf{F}(\mathbf{U})+\mathbf{D}\Delta\mathbf{U}, \tag{6}$$

where $\mathbf{U}=\{u_1,u_2,\ldots,u_M\}$ is the state vector of an element in the medium, for example, the concentrations of the chemical substances participating in a reaction, $\mathbf{F}(\mathbf{U})$ is a certain nonlinear vector function, and $\mathbf{D}$ is the matrix of diffusion coefficients of the components of the medium. These equations are usually called reaction–diffusion equations. These models are usually classified on the basis of the form of the nonlinear source on the right-hand side, the so-called null isocline.[3,5] The form of the null isocline and the classification corresponding to it reflect the character of the equilibrium state in the medium and the ways to achieve it for an assigned form of the nonlinear source. However, such a classification, which is useful from the standpoint of distinguishing models with respect to some general features that characterize the physical processes as a whole, is relatively insensitive to the form of the permissible structures in such models and their dynamics, which are determined by more refined characteristics of the models.

The theory of two-dimensionalized Toda chains briefly described in the preceding sections permits the general description of several different classes of diffusive Toda chain models. The models will obviously differ with respect to the order N of the quadratic forms and the number of components M in the unknown functions H_i describing physical components of the model. For two-dimensionalized Toda chains these numbers are rigidly related if no additional constraints are imposed on the functions $\psi_i(z)$. In this case we have $N=M$. In the reduction of two-dimensionalized Toda chains, the number of independent components can decrease. The relevant reducing conditions, which are important for the discussion below, include, for example, the periodicity conditions of two-dimensionalized Toda chains. Thus, in the general case we have $M\leqslant N$.

The order of the quadratic forms, which is equal to the number of the linearly independent functions $\psi_i(z)$ from which they are constructed, determines (somewhat conditionally) the number of independent spatial structures (modes) evolving simultaneously in the system. Therefore, this order is logically called the number of spatial modes or simply the number of modes in the system. The number of independent quadratic forms determines the number of different physical components evolving simultaneously in the system. In the case of the interpretation of such models from the standpoint of reaction–diffusion models, the number M is

essentially equal to the number of different components participating in chemical reactions. This number will henceforth be called simply the number of components in the system.

In this paper we shall consider only systems of diffusive Toda chains with two and three modes. This is because the structure of the nonlinear sources in diffusive Toda chains depends significantly on the number of modes. Therefore, the closure conditions for the systems of equations for diffusive Toda chains with different numbers of modes are different, and a transition from quadratic forms of one order to another requires separate analyses of the corresponding systems of equations. It is natural to begin these analyses from quadratic forms of low orders, i.e., numbers of modes in the system.

3. MULTICOMPONENT TWO-MODE MODELS

Two-component two-mode models were previously considered in Ref. 8. These models do not have nonsingular periodic solutions in the form of two-component quadratic forms.[8] However, nonsingular periodic solutions exist for multicomponent two-mode systems. Such systems are described by equations of the form[8]

$$\frac{\partial u_i}{\partial t}-D_i\Delta u_i=e^{-2u_i}\left(l_i+e^{u_i}\sum_{k=1}^{n}m_{ik}e^{u_k}\right), \tag{7}$$

where $u_i=\ln\Psi_i(z,\overline{z},t)$, and the solutions for the functions $\Psi_i(z,\overline{z},t)$ are represented by quadratic forms in ψ_1 and ψ_2:

$$\Psi_i(z,\overline{z},t)=|g(z)|^2\{a_i(t)|\psi_1|^2+b_i(t)|\psi_2|^2+c_i(t)\psi_1\psi_2^*+c_i^*(t)\psi_2\psi_1^*\},\quad i=1,\ldots,n, \tag{8}$$

whose coefficients satisfy the equations

$$\dot a_i=\sum_{j=1}^{n}m_{ij}a_j,\quad \dot b_i=\sum_{j=1}^{n}m_{ij}b_j,\quad \dot c_i=\sum_{j=1}^{n}m_{ij}c_j,\quad i=1,\ldots,n, \tag{9}$$

$$(a_ib_i-|c_i|^2)f_i^2D_i=l_i=\text{const},\quad |W_{12}|^2|g|^4=1. \tag{10}$$

Here $W_{12}(z)=[\psi_1(z),\psi_2(z)]$. The requirement that the $\Psi_i(z,\overline{z},t)$ be real implies that $a_i(t)$ and $b_i(t)$ must be real, but $c_i(t)$ can be complex.

We describe the general construction of these solutions first in the general case. Let $\mathbf{n}^{(a)}$, where $a=1,\ldots,n$, be the set of eigenvectors of the real matrix $\mathbf{M}=(m_{ij})$ that correspond to the eigenvalues λ_a:

$$m_{ij}=\sum_{a=1}^{n}\lambda_a\overline{n_j^{(a)}}n_i^{(a)}. \tag{11}$$

Here the $\overline{n_i^{(a)}}$ are the components of the conjugate eigenvectors. The eigenvectors and their conjugates satisfy the following relations:

$$\sum_{i=1}^{n}\overline{n_i^{(a)}}n_i^{(b)}=\delta^{ab}.$$

Because $\mathbf{M}$ is real, each of its complex eigenvalues λ_a has a corresponding complex conjugate $\lambda_b=\lambda_a^*$. The corresponding eigenvectors are also complex-conjugate. Then the solutions for the coefficients can be represented in the form

$$a_i(t)=\sum_{a=1}^{n}A_an_i^{(a)}\exp(\lambda_at),\quad b_i(t)=\sum_{a=1}^{n}B_an_i^{(a)}\exp(\lambda_at),\quad c_i(t)=\sum_{a=1}^{n}C_an_i^{(a)}\exp(\lambda_at). \tag{12}$$

The constants A_a, B_a, and C_a must be chosen such that $a_i(t)$ and $b_i(t)$ would be real and n relations of the form (10) would hold. Generally speaking, the substitution of these relations into (12) leads to a set of n^2 algebraic equations for A_a, B_a, and C_a. In the relations

$$\sum_{a,b=1}^{n}D_in_i^{(a)}n_i^{(b)}\exp[(\lambda_a+\lambda_b)t](A_aB_b-C_aC_b^*-C_a^*C_b)=l_i=\text{const},\quad i=1,\ldots,n,$$

the coefficients accompanying exponential functions in which the exponents $\lambda_a+\lambda_b$ are nonzero should vanish. However, in some cases the number of equations is less than n^2, and then there are solutions with constant values for the model parameters l_i.

Let us consider the case of $n=3$ as an example. We require that the eigenvalues of $\mathbf{M}$ have the values $(0,-i\Omega,+i\Omega)$. Let $\mathbf{n}^{(1)},\mathbf{n}^{(2)},\mathbf{n}^{(3)}=\mathbf{n}^{(2)*}$ be the eigenvectors corresponding to them. The vector $\mathbf{n}^{(1)}$ has real components. The vectors $\overline{\mathbf{n}}^{(2,3)}$, which are conjugate to eigenvectors of $\mathbf{M}$, are simply equal to the complex-conjugate vectors: $\overline{\mathbf{n}}^{(2,3)}=\mathbf{n}^{2,3*}$. In this case

$$\sum_{i=1}^{3}n_i^{a*}n_i^a=0,\quad a=2,3.$$

Therefore,

$$m_{ij}=i\Omega(n_j^{(2)*}n_i^{(2)}-n_i^{(2)*}n_j^{(2)}),\quad i,j=1,2,3,$$

and the functions $a_i(t)$, $b_i(t)$, and $c_i(t)$ can be represented in the form

$$a_i(t)=a_0n_i^{(1)}+An_i^{(2)}e^{i\Omega t}+A^*(n_i^{(2)})^*e^{-i\Omega t},$$

$$b_i(t)=b_0n_i^{(1)}+Bn_i^{(2)}e^{i\Omega t}+B^*(n_i^{(2)})^*e^{-i\Omega t},$$

$$c_i(t)=c_0n_i^{(1)}+C_1n_i^{(2)}e^{i\Omega t}+C_2(n_i^{(2)})^*e^{-i\Omega t},$$

$$c_i^*(t)=c_0^*n_i^{(1)}+C_2^*n_i^{(2)}e^{i\Omega t}+C_1^*(n_i^{(2)})^*e^{-i\Omega t}.$$

The numbers a_0 and b_0 are real, and A, B, c_0, C_1, and C_2 are complex. Substituting these solutions into the relations (10), we obtain the following system of algebraic equations

$$a_0B+b_0A=c_0C_2^*+c_0^*C_1,\quad AB=C_1C_2^*,$$

$$D_i[(a_0b_0-|c_0|^*)(n_i^{(1)})^2+(A^*B+AB^*-|C_1|^2-|C_2|^2)n_i^{(2)}(n_i^{(2)})^*]=l_i,\quad i=1,2,3.$$

The complex constants A and B, for example, can be found from the first two equations of this system:

$$B=\frac{1}{2a_0}(c_0C_2^*+c_0^*C_1)$$
$$\pm\sqrt{\frac{1}{4a_0^2}(c_0C_2^*+c_0^*C_1)^2-C_1C_2^*\frac{b_0}{a_0}},$$
$$A=\frac{C_1C_2^*}{B}. \tag{13}$$

After this there are still eight real constants, three of which can be found for assigned values of D_i and l_i from the last three real algebraic conditions. In order for the corresponding solutions of the system (7) not to be singular (a necessary condition), the functions a_i, b_i, and c_i must not vanish simultaneously. The requirements

$$|a_0|>|A|, \quad |b_0|>|B|, \quad |c_0|^2>\frac{1}{2}(|C_1|^2+|C_2|^2)$$

and the requirements placed on $\psi_1\psi_2$ that a) these functions not have poles and b) they not vanish simultaneously are sufficient conditions.

Thus, it has been shown that three-component diffusive Toda chain models with two-mode excitation have exact periodic nonsingular solutions.

4. THREE-MODE MODELS: GENERAL STATEMENT OF THE PROBLEM

Let us consider models described by equations of two types as examples of multicomponent models with three-mode excitation. The equations of one type have the form

$$\frac{\partial}{\partial t}\Psi_j^{-1}-D_j\Delta\ln\Psi_j=F_j(\Psi_1,\Psi_2,\dots), \tag{14}$$

and the equations of the other type have the more familiar form

$$\frac{\partial}{\partial t}\ln\Psi_j-D_j\Delta\ln\Psi_j=G_j(\Psi_1,\Psi_2,\dots). \tag{15}$$

For convenience, we shall call models of the form (14) models with nonlinear diffusion, and we shall call models of the second form models with linear diffusion. After a replacement of variables according to the formulas $u_j=\ln\Psi_j$, these equations acquire the form of the respective diffusive Toda chains:

$$\frac{\partial}{\partial t}u_j+D_je^{u_j}\Delta u_j=e^{u_j}F_j(e^{u_1},e^{u_2},\dots),$$
$$\frac{\partial}{\partial t}u_j-D_j\Delta u_j=G_j(e^{u_1},e^{u_2},\dots).$$

We note that equations of the form (14) can be written in a somewhat different form using the replacement $\Psi_j\to 1/\Psi_j$:

$$\frac{\partial}{\partial t}\Psi_j+D_j\Delta\ln\Psi_j=F_j\left(\frac{1}{\Psi_1},\frac{1}{\Psi_2},\dots\right). \tag{16}$$

The goal of this section is to calculate of the type of nonlinearity and to construct classes of exact solutions for several models.

According to the general classification proposed in Sec. 2, the solutions for three-mode models have a representation in the form of third-order quadratic forms,

$$\Psi_j(z,\bar{z},t)=|w(z)|^2[a_j(t)|\psi_1|^2+b_j(t)|\psi_2|^2+c_j(t)|\psi_3|^2], \tag{17}$$

in the linearly independent differentiable functions $\psi_1(z)$, $\psi_2(z)$, and $\psi_3(z)$ with the multiplier $|w(z)|^2$, which was introduced so that the solutions would be general. According to the general relations considered in Sec. 1, these functions obey the following identities:

$$\Delta\ln\Psi_j=\frac{a_jb_j|W_{12}|^2+b_jc_j|W_{23}|^2+a_jc_j|W_{13}|^2}{\Psi_j^2}, \tag{18}$$

where

$$W_{\alpha\beta}(z)=[\psi_\alpha,\psi_\beta], \quad \alpha\neq\beta, \quad \alpha,\beta=1,2,3.$$

We additionally introduce functions of the form

$$\Phi_j(z,\bar{z},t)=a_jb_j|W_{12}|^2+b_jc_j|W_{23}|^2+a_jc_j|W_{13}|^2$$
$$=A_j|W_{23}|^2+B_j|W_{31}|^2+C_j|W_{12}|^2, \tag{19}$$

where

$$A_j(t)=c_jb_j=\frac{Q_j}{a_j}, \quad B_j(t)=a_jc_j=\frac{Q_j}{b_j}, \quad C_j(t)=a_jb_j=\frac{Q_j}{c_j},$$
$$Q_j(t)=a_j(t)b_j(t)c_j(t). \tag{20}$$

According to (5), for arbitrary $\psi_1(z)$, $\psi_2(z)$, and $\psi_3(z)$ the functions Φ_j satisfy the identity

$$\Delta\ln\Phi_j=a_jb_jc_j|W_{123}|^2\frac{a_j|\psi_1|^2+b_j|\psi_2|^2+c_j|\psi_3|^2}{\Phi^2}, \tag{21}$$

where

$$W_{123}(z)=\det\begin{pmatrix}\psi_1 & \psi_1' & \psi_1''\\ \psi_2 & \psi_2' & \psi_2''\\ \psi_3 & \psi_3' & \psi_3''\end{pmatrix}$$

is the Wronskian of the functions ψ_i. General relations for the functions $a_j(t)$, $b_j(t)$, and $c_j(t)$, under which the functions Ψ_j and Φ_j satisfy equations of the form (14) or (15), can be obtained on the basis of these identities.

If no constraints are imposed on the form of $\psi_i(z)$, it is difficult to find conditions under which there will not be any functions which depend explicitly on time on the right-hand side of Eq. (14) or (15). In fact, in the general case the functions $|\psi_i|^2$ and $|W_{ij}|^2$ are linearly independent. Therefore, the derivatives of the functions Ψ_j with respect to time can depend only on Ψ_j, and the derivatives of the functions Φ_j can depend only on Φ_j. This situation and the condition of constancy of the coefficients in the original equations lead to the equations

$$\dot{a}_j=\sum_{k=1}^{L}m_{jk}a_k, \quad \dot{b}_j=\sum_{k=1}m_{jk}b_k \quad \dot{c}_j=\sum_{k=1}^{L}m_{jk}c_k, \tag{22}$$

where $m_{ik}=\text{const}$ and L is the number of components in the original system. Now, the relations for A_j, B_j, and C_j follow automatically from (22). The form (linearity) and dimension of the systems of equations for A_j, B_j, and C_j should be the same as for (22). Therefore, the number of different eigenvalues of the system of characteristic equations for A_j, B_j, and C_j should not exceed L. Since A_j, B_j, and C_j are expressed in terms of pairwise products of a_j, b_j, and c_j, the number of different eigenvalues for them doubles in the general case. This imposes significant constraints on the matrix of the coefficients (m_{ik}) in (22), which are difficult to satisfy.

Another way to construct a solvable system of equations of mode dynamics is to adopt some constraints on the form of the functions $\psi_i(z)$, which have the property of the periodicity conditions for two-dimensionalized Toda chains.

We require fulfillment of the relations

$$W_{12}=k_3 g(z)\psi_3, \quad W_{23}=k_1 g(z)\psi_1, \quad W_{31}=k_2 g(z)\psi_2, \tag{23}$$

where k_1, k_2, and k_3 are complex constants, and $g(z)$ is an arbitrary function. As a result, Φ will have the same form (17) as Ψ.

The system of ordinary differential equations (23) in ψ_1, ψ_2, and ψ_3 has a general solution of the form

$$\psi_1(z)=q_1 f(z)\cos\xi(z), \quad \psi_2(z)=q_2 f(z)\sin\xi(z),$$

$$\psi_3(z)=f(z), \quad \frac{d}{dz}\xi(z)=\frac{q_3}{g(z)f(z)},$$

where

$$q_1=i\sqrt{\frac{k_3}{k_1}}, \quad q_2=i\sqrt{\frac{k_3}{k_2}}, \quad q_3=\sqrt{k_1 k_2},$$

and $f(z)$ is an arbitrary function.

5. THREE-MODE MODELS WITH NONLINEAR DIFFUSION

We first consider a very simple one-component model of the form (14). The substitution of (23) into (21) leads to the following identity:

$$\Delta \ln\Psi=|w(z)|^4|g(z)|^2 \times\frac{ab|k_3|^2|\psi_3|^2+bc|k_1|^2|\psi_1|^2+ac|k_2|^2|\psi_2|^2}{\Psi^2}. \tag{24}$$

In a one-component model the requirement that the nonlinearity have the form of a function which depends only on the unknown function leads to the requirement $g(z)=1/w(z)$ and the following equations for $a(t)$, $b(t)$, and $c(t)$:

$$\dot a-D_1 bc|k_1|^2=\lambda a, \quad \dot b-D_1 ac|k_2|^2=\lambda b,$$

$$\dot c-D_1 ab|k_3|^2=\lambda c. \tag{25}$$

This system is similar to the equations encountered in three-wave interaction problems and can be integrated in quadratures. Its general solution is constructed from the following relations, which are obtained directly from (25):

$$a^2=\left(\alpha_0+2D_1|k_1|^2\int \exp(-2\lambda t')Q\,dt'\right)\exp(2\lambda t),$$

$$b^2=\left(\beta_0+2D_1|k_2|^2\int \exp(-2\lambda t')Q\,dt'\right)\exp(2\lambda t),$$

$$c^2=\left(\gamma_0+2D_1|k_3|^2\int \exp(-2\lambda t')Q\,dt'\right)\exp(2\lambda t),$$

where $Q(t)=a(t)b(t)c(t)$, and α_0, β_0, and γ_0 are constants. Multiplying through these equations, we find that $Q(t)$ satisfies the equation

$$Q^2=\left(\alpha_0+2D_1|k_1|^2\int \exp(-2\lambda t')Q\,dt'\right)\times\left(\beta_0+2D_1|k_2|^2\int \exp(-2\lambda t')Q\,dt'\right)\times\left(\gamma_0+2D_1|k_3|^2\int \exp(-2\lambda t')Q\,dt'\right)\exp(6\lambda t).$$

If we introduce the function $P=\int\exp(-2\lambda t')Q\,dt'$, we obtain the following equation for it:

$$\left(\frac{dP}{dt}\right)^2=(\alpha_0+2D_1|k_1|^2P)(\beta_0+2D_1|k_2|^2P)\times(\gamma_0+2D_1|k_3|^2P)e^{2\lambda t},$$

which reduces to the elliptic integral

$$\int\frac{dP}{\sqrt{(\alpha_0+2D_1|k_1|^2P)(\beta_0+2D_1|k_2|^2P)(\gamma_0+2D_1|k_3|^2P)}}=\frac{1}{\lambda}e^{\lambda t}+C_0.$$

Here C_0 is an integration constant.

Two particular solutions of these equations can be written in elementary functions. One can be written in trigonometric functions:

$$a(t)=\frac{a_0 e^{\lambda t}}{\cos\theta}, \quad b(t)=\frac{b_0 e^{\lambda t}}{\cos\theta}, \quad c(t)=c_0 e^{\lambda t}\tan\theta, \tag{26}$$

where

$$\theta(t)=\theta_0\pm D_1 c_0|k_1||k_2|\frac{e^{\lambda t}}{\lambda}, \quad a_0=\pm c_0\frac{|k_1|}{|k_3|},$$

$$b_0=\pm c_0\frac{|k_2|}{|k_3|}, \tag{27}$$

and c_0 is an arbitrary real constant. The other can be written in hyperbolic functions:

$$a(t)=\frac{a_0 e^{\lambda t}}{\sinh\theta}, \quad b(t)=\frac{b_0 e^{\lambda t}}{\sinh\theta}, \quad c(t)=c_0 e^{\lambda t}\coth\theta, \tag{28}$$

where all the constants and the function $\theta(t)$ satisfy the same relations (27). We note that both solutions can be rendered nonsingular in any time interval $t>t_0$ by adjusting the con-

stants, since $\theta(t)$, which plays the role of the phase of the oscillations in the medium, can be made bounded function in that time interval by adjusting the constants.

Thus, exact solutions of a one-component diffusion model of the form

$$-\frac{\partial}{\partial t}\Psi=\Psi^2 D_1\Delta\ln\Psi+\lambda\Psi, \tag{29}$$

or, in the form of diffusive Toda chains, of the form

$$-\frac{\partial}{\partial t}u=D_1 e^u\Delta u+\lambda, \tag{30}$$

where $u=\ln\Psi$, have been constructed. The same model can be represented in the following form:

$$\frac{\partial}{\partial t}v=-D_1\nabla\left(\frac{1}{v}\nabla v\right)+\lambda v, \tag{31}$$

where $v=1/\Psi$. The form of these equations accounts for the name of the corresponding models, viz., models with nonlinear diffusion.

For a two-component model of the form (14) with the components

$$\Psi(z,\overline{z},t)=|w(z)|^2(a(t)|\psi_1|^2+b(t)|\psi_2|^2+c(t)|\psi_3|^2),$$

$$\Phi(z,\overline{z},t)=|w(z)|^2(A(t)|\psi_1|^2+B(t)|\psi_2|^2+C(t)|\psi_3|^2)$$

the closure conditions for the model reduce to the condition $g(z)=1/w(z)$ and to a system of ordinary differential equations for the coefficients of the quadratic forms of the following form:

$$\begin{aligned}
&\dot a+D_1 bc|k_1|^2=\mu_1 A+\lambda_1 a,\\
&\dot b+D_1 ac|k_2|^2=\mu_1 B+\lambda_1 b,\\
&\dot c+D_1 ab|k_3|^2=\mu_1 C+\lambda_1 c,
\end{aligned} \tag{32}$$

$$\begin{aligned}
&\dot A+D_2 BC|k_1|^2=\mu_2 A+\lambda_2 a,\\
&\dot B+D_2 AC|k_2|^2=\mu_2 B+\lambda_2 b,\\
&\dot C+D_2 AB|k_3|^2=\mu_2 C+\lambda_2 c.
\end{aligned} \tag{33}$$

This system is apparently not completely integrable and allows complicated behavior for the solutions. The equations of this model then have the following form:

$$\frac{\partial}{\partial t}\Psi_1^{-1}-D_1\Delta\ln\Psi_1=\frac{\mu_1\Psi_2+\lambda_1\Psi_1}{\Psi_1^2}, \tag{34}$$

$$\frac{\partial}{\partial t}\Psi_2^{-1}-D_2\Delta\ln\Psi_2=\frac{\mu_2\Psi_2+\lambda_2\Psi_1}{\Psi_2^2}, \tag{35}$$

or

$$\frac{\partial}{\partial t}u+D_1\nabla\left(\frac{1}{u}\nabla u\right)=\frac{u}{v}(\mu_1 u+\lambda_1 v), \tag{36}$$

$$\frac{\partial}{\partial t}v+D_2\nabla\left(\frac{1}{v}\nabla v\right)=\frac{v}{u}(\mu_2 u+\lambda_2 v), \tag{37}$$

where $u=\Psi_1^{-1}$ and $v=\Psi_2^{-1}$.

In the case of the conditions (23) multicomponent systems of the form (14) have closure equations which reduce to a set of equations of a form similar to (32):

$$\begin{aligned}
&\dot a_j+D_j b_j c_j|k_1|^2=\sum_{k=1}^{L}\mu_{jk}a_k,\\
&\dot b_j+D_j a_j c_j|k_2|^2=\sum_{k=1}^{L}\mu_{jk}b_k,\\
&\dot c_j+D_j a_j b_j|k_3|^2=\sum_{k=1}^{L}\mu_{jk}c_k,
\end{aligned}$$

for $j=1,2,\ldots,L$, where L is the number of components in the system. The corresponding equations of the model have the following form:

$$\frac{\partial}{\partial t}\Psi_j^{-1}-D_j\Delta\ln\Psi_j=\frac{\sum_{k=1}^{L}\mu_{jk}\Psi_k}{\Psi_j^2},\quad j=1,2,\ldots,L.$$

6. THREE-MODE MODELS WITH LINEAR DIFFUSION

Let us now consider models of the form (15). A one-component model of this form has trivial temporal dynamics. Let Ψ have the form (17). Then the closure conditions for the model reduce to the equations

$$a(t)=a_0 e^{\lambda t},\quad b(t)=b_0 e^{\lambda t},\quad c(t)=c_0 e^{\lambda t},$$

where a_0, b_0, and c_0 are constants, and the model is described by the equation

$$\frac{\partial}{\partial t}\ln\Psi-D_1\Delta\ln\Psi=\lambda+D_1 q_0\frac{e^{\lambda t}}{\Psi}$$

with several additional conditions on the constants k_1, k_2, k_3, and q_0.

Two-component models have more complicated dynamics, which are similar to the dynamics of models of the form (14) considered above. Let one component correspond to the quadratic form (17), and let the other to (19). Then the closure conditions for the system of equations of this model with allowance for (24) and (23) reduce to the redefined system of ordinary differential equations

$$\begin{aligned}
&\dot a=\lambda a+\mu A(t),\quad D_1 bc|k_1|^2=\nu a+\kappa A(t),\\
&\dot b=\lambda b+\mu B(t),\quad D_1 ac|k_2|^2=\nu b+\kappa B(t),\\
&\dot c=\lambda c+\mu C(t),\quad D_1 ab|k_3|^2=\nu c+\kappa C(t).
\end{aligned} \tag{38}$$

This system should be supplemented by equations for the coefficients A, B, and C, which follow from the closure conditions for the model and should have the following form:

$$\begin{aligned}
&\dot A=\lambda_1 a+\mu_1 A,\quad D_2 BC|k_1|^2=\lambda_2 a+\mu_2 A,\\
&\dot B=\lambda_1 b+\mu_1 B,\quad D_2 AC|k_2|^2=\lambda_2 b+\mu_2 B,\\
&\dot C=\lambda_1 c+\mu_1 C,\quad D_2 AB|k_3|^2=\lambda_2 c+\mu_2 C.
\end{aligned} \tag{39}$$

However, the set of equations (38) already completely specifies the form of the functions a, b, and c and the form of the

functions A, B, and C. Therefore, the explicit forms of the coefficients $\lambda_{1,2}$ and $\mu_{1,2}$, which are now already functions of time, can also be calculated from this system. In fact, it follows from (38) that a, b, and c satisfy the same system of equations (25):

$$\dot{a}=\Lambda a+R_1 bc|k_1|^2, \quad \dot{b}=\Lambda b+R_1 ac|k_2|^2,$$

$$\dot{c}=\Lambda c+R_1 ab|k_3|^2, \tag{40}$$

where $\Lambda=\lambda-\nu\mu/\kappa$ and $R_1=\mu D_1/\kappa$. This system yields the relations

$$\frac{d}{dt}(bc)=2\Lambda bc+R_1 a(|k_2|^2c^2+|k_3|^2b^2),$$

$$\frac{d}{dt}(ac)=2\Lambda ac+R_1 b(|k_3|^2a^2+|k_1|^2c^2),$$

$$\frac{d}{dt}(ab)=2\Lambda ab+R_1 c(|k_2|^2a^2+|k_1|^2b^2).$$

From the same equations we obtain

$$\frac{d}{dt}(|k_2|^2c^2+|k_3|^2b^2)-2\Lambda(|k_2|^2c^2+|k_3|^2b^2)$$
$$=2R_1|k_2k_3|^2Q,$$

$$\frac{d}{dt}(|k_3|^2a^2+|k_1|^2c^2)-2\Lambda(|k_3|^2a^2+|k_1|^2c^2)$$
$$=2R_1|k_1k_3|^2Q,$$

$$\frac{d}{dt}(|k_2|^2a^2+|k_1|^2b^2)-2\Lambda(|k_2|^2a^2+|k_1|^2b^2)$$
$$=2R_1|k_1k_2|^2Q,$$

where $Q(t)=a(t)b(t)c(t)$. The solutions of these equations can be written

$$|k_2|^2c^2+|k_3|^2b^2=[\alpha_0+2R_1|k_2k_3|^2P(t)]e^{2\Lambda t},$$

$$|k_3|^2a^2+|k_1|^2c^2=[\beta_0+2R_1|k_1k_3|^2P(t)]e^{2\Lambda t},$$

$$|k_2|^2a^2+|k_1|^2b^2=[\gamma_0+2R_1|k_2k_1|^2P(t)]e^{2\Lambda t}.$$

Here $P(t)=\int e^{-2\Lambda t}Q(t)\,dt$. To close the system, the integration constants α_0, β_0, and γ_0 should be set equal to

$$\alpha_0=\frac{\sigma_0}{|k_1|^2}, \quad \beta_0=\frac{\sigma_0}{|k_2|^2}, \quad \gamma_0=\frac{\sigma_0}{|k_3|^2}.$$

Combining the relations obtained, we ultimately find

$$\lambda_1(t)=\left(\lambda-2\frac{\mu\nu}{\kappa}\right)\frac{\nu}{\kappa}+\frac{D_1^2\mu}{\kappa^2}$$
$$\times[\sigma_0+2R_1|k_1k_2k_3|^2P(t)]e^{2\Lambda t},$$

$$\mu_1(t)=2\lambda-3\frac{\mu\nu}{\kappa}=\text{const},$$

$$\lambda_2(t)=\frac{\nu^3D_2}{\kappa^2D_1}-\frac{D_1D_2\nu}{\kappa^2}[\sigma_0+2R_1|k_1k_2k_3|^2P(t)]e^{2\Lambda t}$$
$$+\frac{D_1^2D_2}{\kappa^2}|k_1k_2k_3|^2Q(t),$$

$$\mu_2(t)=\frac{\nu^2D_2}{\kappa D_1}=\text{const}.$$

As a result the equations of the model take the form

$$\frac{\partial}{\partial t}\ln\Psi-D_1\Delta\ln\Psi=\frac{\Psi(\lambda\Psi+\mu\Phi)+\nu\Psi+\kappa\Phi}{\Psi^2},$$

$$\frac{\partial}{\partial t}\ln\Phi-D_2\Delta\ln\Phi=\frac{\Phi(\lambda_1(t)\Psi+\mu_1\Phi)+\lambda_2(t)\Psi+\mu_2\Phi}{\Phi^2}, \tag{41}$$

or

$$\frac{\partial}{\partial t}u-D_1\Delta u=e^{-2u}(\lambda e^{2u}+\mu e^{u+v}+\nu e^u+\kappa e^v)$$

$$\frac{\partial}{\partial t}v-D_2\Delta v=e^{-2v}$$
$$\times[\lambda_1(t)e^{u+v}+\mu_1(t)e^{2v}+\lambda_2(t)e^u+\mu_2e^v], \tag{42}$$

where $u=\ln\Psi$ and $v=\ln\Phi$.

As can be seen from the examples presented, the closure conditions for one-component and two-component models of the form (15) lead to nonlinearities, which depend explicitly on time in a special way. In all likelihood, this also applies to multicomponent models of that type. Therefore, such models can be less effective as basic models for wave propagation processes in active media and the formation of regular structures in them than models of the form (14) or models with two-mode excitation. However, the coefficients in Eqs. (41) and (42), which are explicitly time-dependent, are of order D_1D_2 and $D_1^2D_2$. Therefore, in situations where the dimensionless diffusion coefficients are small quantities, the time dependence of the coefficients can be neglected. Thus, these models are no less effective than the models (14) in many systems actually encountered.

7. GENERAL PRINCIPLES FOR CONSTRUCTING DIFFUSIVE TODA CHAIN MODELS

As was demonstrated above and in Ref. 8, diffusive Toda chain models, which allow exact solutions in quadratic forms, comprise a fairly rich class of models. This class of models can be expanded by considering multicomponent systems and systems with multimode excitation using the more general theory of exact periodic two-dimensionalized Toda chains, which can also be formulated on the basis of quadratic forms (see Refs. 11 and 12). The solutions allowed by the models considered form classes of exact solutions, which depend on arbitrary functional parameters. Although the stability of the solutions was not analyzed in the present work, it can be expected that the classes of solutions constructed are fairly stable by virtue of the arbitrary nature of

the functional parameters on which the solutions depend. All this points to the uniqueness of the models considered among other basic models. Therefore, it would be of interest to find the general physical properties of diffusive Toda chains which distinguish them from other basic models and would account for the existence of special conditions in them that give rise to regular structures with simple temporal evolution. This raises the question of the physical principles for constructing them and employing them as basic models.

The usual means for constructing autowave models are kinetic equations or balance equations for physical parameters of the systems supplemented by terms that are responsible for the diffusion of individual components in these systems and by nonlinear sources, which take into account the mutual influence of the individual components on one another. The type of nonlinearity is usually dictated by some simple physical arguments, for example, by the probabilistic nature of models of the Lotka–Volterra type, or by the form of the law of mass action and the requirements that the nonlinearity vanish at equilibrium points of the medium for reaction–diffusion models.

One of the characteristic features of diffusive Toda chain models is that they have a nonlinearity of the type observed in Toda chains, i.e., a sum of exponential functions with exponents that are multiples of the fields in the model. This type of nonlinearity usually does not appear explicitly in an autowave theory, since, for example, the kinetic equations describing the variation of the concentrations of substances contain power-law nonlinearities due to the form of the law of mass action. If these arguments are followed in interpreting and analyzing diffusive Toda chain models, it is natural to try to impart to them a form in which the nonlinearity has the form of rational functions of components of the model. It is not difficult to see that the equations of diffusive Toda chains written in terms of the unknown functions Ψ_i are similar to kinetic equations with diffusion relative to the concentrations of substances with nonlinearities in the form of rational functions like, for example, (34) and (41). However, although the diffusion operator transforms into a new diffusion operator in this formulation, additional terms appear. These terms describe physical processes which are not present in the other formulation corresponding to the general form of diffusive Toda chains. For example,

$$\Psi_i \Delta \ln \Psi_i = \Delta \Psi_i - \frac{(\nabla \Psi_i)^2}{\Psi_i}.$$

The last term on the right-hand side of this expression can be interpreted as the transport of a component with the concentration Ψ_i by a flow having the velocity field $\mathbf{v}=\nabla \Psi_i/\Psi_i$. Although there is, in fact, no advection in the model, the additional source is such that it is completely equivalent to advective transport. Thus, diffusive Toda chain models contain some additional physical mechanisms for the transport of matter not included among the mechanisms usually considered in basic models.[3]

In order to demonstrate the role of these additional mechanisms in the dynamics of diffusive Toda chain models, let us examine the formal method for constructing models of this type from standard kinetic models. We can consider the Lotka–Volterra (predator-prey) model without diffusion as such a model. Its equations have the form

$$\dot{N}_1 = \kappa_1 N_1 - \alpha N_1 N_2, \quad \dot{N}_2 = \beta N_1 N_2 - \kappa_2 N_2.$$

Here κ_1, κ_2, α, and β are constants. Following the general thinking behind diffusive Toda chains, the concentrations N_1 and N_2 should be replaced by the apparent entropies of each component of the system according to the formula

$$S_1 = -\ln N_1, \quad S_2 = -\ln N_2. \tag{43}$$

Then the equations of the model describe the production of entropy in each component of the system:

$$\dot{S}_1 = -\kappa_1 + \alpha \exp(-S_2), \quad \dot{S}_2 = -\beta \exp(-S_1) + \kappa_2. \tag{44}$$

The constants κ_1 and κ_2, which describe the natural birth rate of the prey and the natural death rate of the predators, correspond in such a formulation to the natural constant increase in entropy in component 1 (the prey) and the constant decrease in entropy in component 2 (the predators). In this case the entropy should be regarded as a measure of the number of available states of the system, rather than as a measure of its disorder. Under such an interpretation the remaining elements of the equations also describe the production of entropy components due to the interaction of the subsystems with one another. Thus, the system of equations (44) already has the form corresponding to Toda chain models and has the meaning of equations of entropy component production. Obviously, in kinetic models without diffusion that have rational nonlinearities replacements of the form (43) will always lead to Toda chain equations.

Let us now consider the transition from kinetic models without diffusion to models with diffusion. Under such a transition Eqs. (43) are supplemented by a diffusion flux of ''matter.'' As a result, we obtain equations of the following form:

$$\dot{N}_1 - D_1 \Delta N_1 = \kappa_1 N_1 - \alpha N_1 N_2,$$
$$\dot{N}_2 - D_2 \Delta N_2 = \beta N_1 N_2 - \kappa_2 N_2. \tag{45}$$

Here D_1 and D_2 are the diffusion coefficients of each of the components in the system. The model does not contain any other mechanisms for the transport of ''matter.''

If we now utilize the replacement (43), the equations for the apparent entropies of the components of the system now contain terms, which, according to their form, can be interpreted as the advective transport of matter with the velocities of the medium $\mathbf{v}_1 = -D_1 \nabla S_1$ for component 1 and $\mathbf{v}_2 = -D_2 \nabla S_2$ for component 2:

$$\dot{S}_1 - D_1 \Delta S_1 - D_1 (\nabla S_1)^2 = -\kappa_1 + \alpha \exp(-S_2),$$
$$\dot{S}_2 - D_2 \Delta S_2 - D_2 (\nabla S_2)^2 = -\beta \exp(-S_1) + \kappa_2. \tag{46}$$

The original model does not contain these terms; therefore, to avoid their appearance in constructing a diffusive Toda chain model, Eqs. (44) should be supplemented by a diffusive flux specifically of entropy. Then Eqs. (44) take on the form

$$\dot{S}_1 - D_1 \Delta S_1 = -\kappa_1 + \alpha \exp(-S_2), \tag{47}$$

$$\dot{S}_2 - D_2 \Delta S_2 = -\beta \exp(-S_1) + \kappa_2 .$$

This system of equations already has the form of two-component diffusive Toda chains and is similar to the models considered here and in Ref. 8.

What distinguishes Eqs. (46) from (47) is the presence of additional entropy sources, whose values are always positive in all space:

$$J_1 = D_1 (\nabla S_1)^2 > 0, \quad J_2 = D_1 (\nabla S_2)^2 > 0.$$

The influence of these sources can be understood on the basis of their analogy to the advective entropy transport caused by flow of the medium with the rates

$$\mathbf{u}_1 = -D_1 \nabla S_1, \quad \mathbf{u}_2 = -D_2 \nabla S_2 .$$

As can be seen, the rates of advective transport correspond to the transport of entropy in the direction opposite to its gradient, i.e., these sources accelerate obliteration of the entropy differences. The absence of these sources in (47) means that the both the entropy and concentration differences are obliterated more slowly in that model. This points to the reason why regular structures appear ''more easily'' in diffusive Toda chain models with entropy diffusion: the specific slowing of the diffusive obliteration of these differences favors the formation of regular structures.

This fact, in turn, points out a general way to construct basic models of the diffusive Toda chain type that describe the appearance of regular structures in active media with diffusion. In fact, if there is a certain model of the form

$$\frac{\partial}{\partial t} N_i - D_i \Delta N_i = R(N_1, \ldots, N_n), \quad i = 1, \ldots, n,$$

where $R(N_1, \ldots, N_n)$ is a rational function of both arguments, then we can use the replacement $N_i \to S_i = -\ln N_i$ to go over from it to the diffusive Toda chain equations

$$\frac{\partial}{\partial t} S_i - D_i \Delta S_i = R(\exp(-S_1), \ldots, \exp(-S_n)),$$

$$i = 1, \ldots, n,$$

in which the transition from the diffusion of concentrations to the diffusion of entropies has also been made.

8. CONCLUSION

As has been shown above, there is a rich class of models that describe waves in active media with diffusion (reaction–diffusion models) and, according to the form of their equations, are diffusive Toda chain models. The main distinguishing feature of these models is the fact that they allow fairly representative classes of exact solutions. This permits the use of these models as basic models for investigating diverse phenomena and types of waves in active media. One of the principal results of the present work is the further development of the method for constructing exact solutions of the equations of these models, which was begun in Ref. 8 for several principal types of diffusive Toda chain models. Another important result of the present work is an analysis of the principal methods for constructing diffusive Toda chain models from standard basic models with nonlinear sources in the form of unknown power-law functions (concentrations). As has been shown, this transition can be accomplished formally in many cases that are of interest from the standpoint of practical applications by replacing the concentrations in the equations by quantities having the meaning of entropies of the spatial distribution of the components in the system and replacing the diffusive fluxes of matter by diffusive fluxes of entropy.

In this paper we have not presented any comparisons of the models investigated with the already fairly standard classification based on the form of the null isoclines of nonlinear sources.[3,5] This was due, on the one hand, to the great variety of types of sources in diffusive Toda chains and, on the other hand, to the fact that, as was shown in Ref. 8, sources with different forms of isoclines according to the classification in Ref. 3 and formal constancy of the form of the solutions in the class of quadratic forms can be obtained by combining the variables and equations of multicomponent systems.

*)E-mail: zhuravl@themp.univ.simbirsk.su

[1] H. Haken, *Synergetics: An Introduction: Nonequilibrium Phase Transitions and Self-Organization in Physics, Chemistry, and Biology*, 2nd enl. ed., Springer-Verlag, Berlin–New York (1978) [Russ. transl., Mir, Moscow (1980)].
[2] I. Prigogine, *From Being to Becoming, Time and Complexity in the Physical Sciences*, W. H. Freeman, San Francisco (1980) [Russ. transl., Nauka, Moscow (1985)].
[3] V. A. Vasil'ev, Yu. M. Romanovskiĭ, and V. G. Yakhno, in *Autowave Processes* [in Russian], D. S. Chernavskiĭ (ed.), Nauka, Moscow (1987).
[4] V. A. Davydov and V. G. Morozov, Usp. Fiz. Nauk **166**(3), 327 (1996) [Phys. Usp. **39**, 305 (1996)].
[5] B. S. Kerner and V. V. Osipov, *Autosolitons: a New Approach to Problems of Self-Organization and Turbulence*, Kluwer Academic, Dordrecht–Boston (1994).
[6] V. A. Davydov, V. S. Zykov, and A. S. Mikhaĭlov, Usp. Fiz. Nauk **161**, 45 (1991) [Sov. Phys. Usp. **34**, 665 (1991)].
[7] P. K. Brazhnik, V. A. Davydov, and A. S. Mikhaĭlov, Teor. Mat. Fiz. **74**, 440 (1988).
[8] V. M. Zhuravlev, JETP Lett. **65**, 300 (1997).
[9] M. Toda, *Theory of Nonlinear Lattices*, Springer-Verlag, Berlin–New York (1981) [Russ. transl., Mir, Moscow (1984)].
[10] V. G. Drinfel'd and V. V. Sokolov, in *Results of Science and Technology. Current Problems in Mathematics, Vol. 24* [in Russian], VINITI, Moscow (1984), p. 81.
[11] A. N. Leznov and M. V. Saveliev, *Group-Theoretical Methods for Integration of Nonlinear Dynamical Systems*, Birkhauser Verlag, Basel–Boston (1992).
[12] A. Lesnov and M. Savel'ev, Physica D **3**, 6272 (1981).
[13] V. M. Zhuravlev, Prikl. Mat. Mekh. **58**(6), 61 (1994).
[14] V. É. Adler and A. B. Shabat, Teor. Mat. Fiz. **111**, 323 (1997).
[15] V. V. Kozlov, *Symmetries, Topology, and Resonances in Hamiltonian Mechanics*, Springer, Berlin–New York (1996).

Translated by P. Shelnitz

Double atomic photoeffect in the relativistic domain: angular and energy distributions of photoelectrons

A. I. Mikhaĭlov and I. A. Mikhaĭlov*)

St. Petersburg Nuclear Physics Institute, 188350 Gatchina, St. Petersburg, Russia
(Submitted 23 January 1998)
Zh. Éksp. Teor. Fiz. **114**, 1537–1554 (November 1998)

We study the double ionization of the atomic *K*-shell by a single photon in the relativistic energy domain. The differential and total cross sections of the process are calculated. It is shown that the ratio of the cross sections of double and single ionization increases with the photon energy, tending to the limit $0.34/Z^2$, where Z is the atomic number or the nuclear charge. The formulas are found to be valid for $Z \gg 1$ and $\alpha Z \ll 1$, where $\alpha = 1/137$ is the fine-structure constant.
 [S1063-7761(98)00111-5]

1. INTRODUCTION

Double photoionization of atoms has been studied for more than 30 years. This reason for this extended interest in the problem is that the ejection of two electrons by a single photon is determined solely by the electron–electron interaction. In view of this, electron correlations manifest themselves here most vividly. The main research, both theoretical and experimental, has been focused on the nonrelativistic domain of photon energies, $\omega \ll m$ (m is the electron mass[1)]) and on the helium atom as the simplest multielectron system.[1–13]

The characteristic features of the nonrelativistic double photoeffect are the constancy of the ratio R of the cross sections of double and single ionization in the high-frequency region $I \ll \omega \ll m$ (I is the binding energy of a K-electron) and the very nonuniform distribution of energy among the photoelectrons. The region of the electron energy spectrum that contributes the most to the cross section is the edge region, where the energy of one electron, E_1, is much higher than I and that of the other, E_2, is of order I. However, as shown by Amusia *et al.*[5] and Drukarev,[13] if we allow for the contribution of the central (middle) region of the spectrum, where the photon energy is more evenly distributed among the two electrons ($E_1 \sim E_2$), we can get a correction R' to R that increases with ω but remains a small quantity of order $(\omega/m)R$ in the nonrelativistic domain. This suggests that in the relativistic domain $\omega \sim m$ the central part of the electron spectrum is as significant as the edge part.

The relativistic double photoeffect has been studied only by Drukarev and Karpeshin,[14] who derived a formula for the differential (in the electron energy) cross section and found the dependence of the ratio R on the photon energy ω. We believe, however, that their expression for the cross section in the central region of the spectrum contains errors. Bearing in mind the ever growing interest in the problem of double photoionization at higher and higher photon energies, we have rederived the formulas for the relativistic double photoeffect. In this paper we establish the energy and angular distributions of photoelectrons and derive a formula for the ratio of the cross sections of double and single ionization of the *K*-shell of the atom. At high photon energies the formula for this ratio becomes very simple and the ratio tends to a constant limit $0.34/Z^2$. All the formulas are valid for atoms with $Z \gg 1$, since we use perturbation theory in the electron–electron interaction. At the same time, wherever it is possible we do expansions in powers of the Coulomb parameter αZ and $\xi = \alpha ZE/p$, where E and p are the electron energy and momentum, so that αZ must be much smaller than unity.

2. AMPLITUDE AND CROSS SECTION OF THE RELATIVISTIC DOUBLE PHOTOEFFECT

We examine double photoionization at photon energies $\omega \gg \eta = m\alpha Z$. Using the Furry representation,[15] we can represent the double-photoeffect amplitude to first order in the electron–electron interaction by eight Feynman diagrams, four of which are depicted in Fig. 1. The other four diagrams differ from these in the sign and interchange of the final states ψ_{p_1} and ψ_{p_2} (or the initial states ψ_α and ψ_β). We will always assume that $p_1 > p_2$.

Analysis of the diagrams shows that diagrams a and a′ contribute the most to the cross section in the edge region of the electron energy spectrum ($p_1 \gg p_2 \sim \eta$, where η is the average K-electron momentum).[2)] Here the denominators of the photon and electron propagators of these diagrams are small, whereas the denominator of the electron propagator of the diagrams b and b′ is not small ($\sim\omega$), and neither are the denominators of the photon and electron propagators of the diagrams c, d, c′, and d′. However, the momentum transfer to the nucleus in the edge region is large, $q = |\mathbf{k} - \mathbf{p}_1 - \mathbf{p}_2| \approx |\mathbf{k} - \mathbf{p}_1| \gg \eta$, just as in the single photoeffect. The electron can transfer such momentum to the nucleus only over short distances, which reduces the cross section substantially. On the other hand, in the central part of the spectrum the photoelectron momenta may add up and balance the photon momentum. As a result the momentum transferred to the nucleus is low ($q \sim \eta$) and the process takes place at great (atomic) distances from the nucleus, where the probability of detecting both electrons is the highest. Below we will also show that the size of the central region ($\sim\omega$) is much larger than that of the edge region ($\sim I$). These two factors enhance

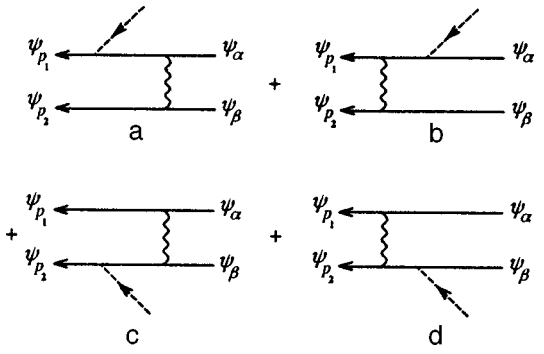


FIG. 1. Feynman diagrams for the double photoeffect. Solid lines denote electrons in the Coulomb of the nucleus, dashed lines denote photons, and wavy lines denote the electron–electron interaction.

the contribution of the central region to the cross section, making it comparable to the contribution of the edge region. The other parts of the spectrum have essentially no effect on the total cross section and are not considered in this paper.

2.1. Energy and angular distributions of electrons in the edge region of the energy spectrum

The amplitude of the double photoeffect in the edge region of the electron spectrum is represented by diagrams a and a′:

$$M_{\text{edge}}^{++}=M_{\text{a}}-M_{\text{a}'}. \tag{1}$$

In the coordinate representation the amplitude M_{a} can be written

$$M_{\text{a}}=\int \overline{\Phi}_{p_1}(\mathbf{r}')\,\gamma^{\mu}\psi_{\alpha}(\mathbf{r}')\,d\mathbf{r}' \times\int \frac{\exp(iR\Delta)}{4\pi R}\overline{\psi}_{p_2}(\mathbf{r})\,\gamma_{\mu}\psi_{\beta}(\mathbf{r})\,d\mathbf{r}; \tag{2}$$

here

$$\overline{\Phi}_{p_1}(\mathbf{r}')=\int \overline{\psi}_{p_1}(\mathbf{r}'')\hat{e}\exp(i\mathbf{k}\cdot\mathbf{r}'')G_C^E(\mathbf{r}'',\mathbf{r}')\,d\mathbf{r}'', \tag{3}$$

$$R=|\mathbf{r}-\mathbf{r}'|,\quad \Delta=E_2-E_{1s}=\varepsilon_2+I,\quad E=E_{1s}-\Delta=m-\varepsilon_2-2I,\quad \varepsilon_2=E_2-m, \tag{4}$$

where G_C^E is the relativistic Coulomb Green's function for an electron with energy E, ψ_{α} and ψ_{β} are the wave functions of K-electrons with different orientation of spin,[3)] $\overline{\psi}_{p_1}$ and $\overline{\psi}_{p_2}$ are the Dirac-conjugate wave functions of the electrons belonging to the continuous spectrum, E_1 and $\mathbf{p}_1$ (E_2 and $\mathbf{p}_2$) are the photoelectron energy and momentum, and γ^{μ} are the Dirac matrices (summation over μ is implied). The notation $\hat{A}$ is used for the scalar product $\gamma A=\gamma_0 A_0-\boldsymbol{\gamma}\cdot\mathbf{A}$. For linearly polarized photons with momentum $\mathbf{k}$ and polarization vector $\mathbf{e}$ we have

$$\hat{e}=-\boldsymbol{\gamma}\cdot\mathbf{e},\quad \mathbf{e}\cdot\mathbf{k}=0. \tag{5}$$

The integrals in (2) converge at $r\sim r'\sim\eta^{-1}$, and the integral in (3) converges at $r''\sim|\mathbf{k}-\mathbf{p}_1|^{-1}\sim m^{-1}$. At such distance the field differs little from the Coulomb field. Hence the use of Coulomb wave functions and the Coulomb Green's function is justified.

Since in the edge region one of the photoelectrons has low energy ($p_2\sim\eta$), the vector part of the electron current is small compared to the scalar part:

$$\overline{\psi}_{p_2}\boldsymbol{\gamma}\psi_{\beta}\sim\frac{\mathbf{p}_2}{m}\varphi_{p_2}^{*}\varphi_{\beta},\quad \overline{\psi}_{p_2}\gamma_0\psi_{\beta}\sim\varphi_{p_2}^{*}\varphi_{\beta} \tag{6}$$

(φ_{p_2} and φ_{β} are the nonrelativistic analogs of the functions ψ_{p_2} and ψ_{β}), and in the sum over μ in (2) it is enough to leave the term with $\mu=0$. If we go over to the momentum representation, we obtain

$$M_{\text{a}}=\int\frac{d\mathbf{f}}{(2\pi)^3}F_1(\mathbf{p}_1,\mathbf{f})D(\mathbf{f})F_2(\mathbf{p}_2,\mathbf{f}), \tag{7}$$

where

$$F_1(\mathbf{p}_1,\mathbf{f})=\int\frac{d\mathbf{f}'\,d\mathbf{f}_1}{(2\pi)^6}\times\langle\psi_{p_1}|\mathbf{f}'+\mathbf{k}\rangle\hat{e}\langle\mathbf{f}'|G_C^E|\mathbf{f}_1\rangle\gamma_0\langle\mathbf{f}_1+\mathbf{f}|\psi_{\alpha}\rangle, \tag{8}$$

$$F_2(\mathbf{p}_2,\mathbf{f})=\int\frac{d\mathbf{f}_2}{(2\pi)^3}\langle\psi_{p_2}|\mathbf{f}_2\rangle\gamma_0\langle\mathbf{f}_2-\mathbf{f}|\psi_{\beta}\rangle, \tag{9}$$

$$D(\mathbf{f})=\frac{1}{f^2-\Delta^2-i0},\quad \Delta=\varepsilon_2+I. \tag{10}$$

The region $f\sim f_1\sim f_2\sim\eta$ provides the main contribution to the integrals in (7)–(9). Since $\varepsilon_2=p_2^2/2m\sim I\sim\alpha Z\eta$, in the lowest order in αZ we can put

$$D(\mathbf{f})=\frac{1}{f^2}, \tag{11}$$

and use nonrelativistic Coulomb wave functions for the wave functions of bound electrons and the low-energy photoelectron. Then

$$F_2(\mathbf{p}_2,\mathbf{f})=w_{\lambda_2}^{\dagger}w_{\beta}N_1\left(-\frac{\partial}{\partial\eta}\right)\langle\varphi_{p_2}|V_{i\eta}|\mathbf{f}\rangle, \tag{12}$$

where $N_1=\eta^3/\pi$, $\eta=m\alpha Z$, and w_{λ} is the Pauli spinor with the zth component of the spin equal to λ. We denote the possible values of λ by α and β, with

$$w_{\alpha}=\begin{pmatrix}1\\0\end{pmatrix},\quad w_{\beta}=\begin{pmatrix}0\\1\end{pmatrix},\quad w_{\lambda}^{\dagger}w_{\lambda}=1. \tag{13}$$

The matrix element of the operator $V_{i\eta}$ in the momentum representation is

$$\langle\mathbf{f}|V_{i\eta}|\mathbf{f}'\rangle=\frac{4\pi}{(\mathbf{f}-\mathbf{f}')^2+\eta^2}. \tag{14}$$

Calculating $F_1(\mathbf{p}_1,\mathbf{f})$ is the most difficult thing here, since F_1 contains the relativistic Coulomb Green's function with a low energy $\varepsilon=E-m\sim I$. For such a function the Coulomb parameter $\xi=\alpha ZE/p$ is order unity and we cannot

use it as an expansion parameter. However, (8) can be simplified if we allow for the fact that the wave function of the high-energy electron can be expanded in powers of the Coulomb parameters αZ and $\xi_1=\alpha Z E_1/p_1\sim\alpha Z$ (we assume that $\alpha Z\ll 1$):

$$\langle\psi_{p_1}|=\overline{u}_{p_1}\{\langle\mathbf{p}_1|-\alpha Z\langle\mathbf{p}_1|\hat{V}_0 G^{E_1}+\cdots\}, \tag{15}$$

where G^{E_1} is the relativistic propagator of the electron with energy E_1 in the absence of an external field, and $\overline{u}_{p_1}=u_{p_1}^{\dagger}\gamma_0$, with u_{p_1} the Dirac bispinor with momentum p_1. Below we will show that both terms are needed in (15) to ensure that the expression for F_1 in the lowest order in αZ is correct.

Let us calculate the contribution F_{10} of the plane wave to the integral in (8), the first term of the expansion (15):

$$F_{10}=N_1\overline{u}_{p_1}\hat{e}U(\mathbf{f})u_0, \quad U(\mathbf{f})=-\frac{\partial}{\partial\eta}\langle\boldsymbol{\kappa}|G_C^E V_{i\eta}|-\mathbf{f}\rangle, \tag{16}$$

where $\boldsymbol{\kappa}=\mathbf{p}_1-\mathbf{k}$, and u_0 is the bispinor of an electron at rest, $\gamma_0 u_0=u_0$. Although the relativistic Coulomb Green's function in (16) corresponds to an electron with nonrelativistic energy, this Green's function cannot be replaced by the nonrelativistic Green's function, since one of the momenta on which it depends is relativistic ($\kappa\sim m$). Hence we transform $U(\mathbf{f})$ so that the Green's operator G_C^E is sandwiched between the nonrelativistic momenta $f\sim f'\sim\eta$. This can be done by using the Lippmann–Schwinger equation for the relativistic Coulomb Green's function:[16]

$$G_C^E=G^E-\alpha Z G^E\hat{V}_0 G_C^E, \tag{17}$$

where $-\alpha Z V_0$ is the operator of the interaction of an electron and the Coulomb field of the nucleus, $\hat{V}_0=\gamma_0 V_0$. The matrix element of V_0 is defined in (14) with $\eta=0$, and the matrix element of G^E has the form

$$\langle\mathbf{f}|G^E|\mathbf{f}'\rangle=G^E(\mathbf{f})(2\pi)^3\delta(\mathbf{f}-\mathbf{f}'),$$
$$G^E(\mathbf{f})=\frac{E\gamma_0-\boldsymbol{\gamma}\cdot\mathbf{f}+m}{p^2-f^2+i0}, \quad p^2=E^2-m^2. \tag{18}$$

Substituting (17) in (16), we obtain

$$U(\mathbf{f})=G^E(\boldsymbol{\kappa})\left(-\frac{\partial}{\partial\eta}\right)\left\{\langle\boldsymbol{\kappa}|V_{i\eta}|-\mathbf{f}\rangle-\alpha Z\right.$$
$$\left.\times\int\frac{d\mathbf{f}'}{(2\pi)^3}\langle\boldsymbol{\kappa}|\hat{V}_0|\mathbf{f}'\rangle\langle\mathbf{f}'|G_C^E V_{i\eta}|-\mathbf{f}\rangle\right\}. \tag{19}$$

The dominant contribution to the integral in (19) is provided by the region where $f'\sim\eta$. At such values of f' the factor $\langle\boldsymbol{\kappa}|V_0|\mathbf{f}'\rangle\approx 4\pi/\kappa^2$ can be taken outside the integral sign, and the relativistic function G_C^E reduces to the nonrelativistic function G_C^{nonrel} (Ref. 17). After we have taken the partial derivative with respect to η the second term in (19) dominates, with the result that[4)]

$$U(\mathbf{f})u_0=\frac{4\pi\alpha Z}{\kappa^2}G^E(\boldsymbol{\kappa})\frac{\partial}{\partial\eta}J(\eta,\mathbf{f}), \tag{20}$$

where

$$J(\eta,\mathbf{f})=\int\frac{d\mathbf{f}'}{(2\pi)^3}\langle\mathbf{f}'|G_C^{\text{nonrel}}V_{i\eta}|-\mathbf{f}\rangle$$
$$=\frac{2ipm}{4\pi}I_y\langle\mathbf{f}|V_{py+i\eta}|0\rangle, \tag{21}$$

$$I_y=\int_1^{\infty}dy\left(\frac{y+1}{y-1}\right)^{i\xi}, \quad \xi=\frac{\alpha Z m}{p}, \tag{22}$$

$$F_{10}=\overline{u}_{p_1}\hat{e}\left(1+\frac{\tilde{\kappa}}{2m}\right)u_0 N_1\frac{8\pi\eta}{\kappa^4}\left(-\frac{\partial}{\partial\eta}\right)J(\eta,\mathbf{f}); \tag{23}$$

here $\tilde{\kappa}=\boldsymbol{\alpha}\cdot\boldsymbol{\kappa}$, and $\boldsymbol{\alpha}=\gamma_0\boldsymbol{\gamma}$ is the Dirac matrix.

Inserting the second term on the right-hand side of Eq. (15) into the integral (8) yields

$$F_{11}=\alpha Z N_1\overline{u}_{p_1}\langle\boldsymbol{\kappa}|\hat{V}_0|0\rangle G^{E_1}(\mathbf{k})\hat{e}\frac{\partial}{\partial\eta}J(\eta,\mathbf{f})u_0$$
$$=\overline{u}_{p_1}(\tilde{k}-\omega)\hat{e}u_0 N_1\frac{4\pi\alpha Z}{\kappa^2(p_1^2-\omega^2)}\left(-\frac{\partial}{\partial\eta}\right)J(\eta,\mathbf{f}). \tag{24}$$

Comparing (23) and (24), we see that both terms in the expansion (15) of the wave functions yield contributions of the same order in αZ to the amplitude of the process. The amplitude (8) becomes

$$F_1(\mathbf{p}_1,\mathbf{f})=F_{10}+F_{11}=T_{\lambda_1\alpha}\left(-\frac{\partial}{\partial\eta}\right)J(\eta,\mathbf{f}), \tag{25}$$

$$T_{\lambda_1\alpha}=N_1\frac{8\pi\eta}{\kappa^4}\overline{u}_{p_1\lambda_1}$$
$$\times\hat{e}\left(1+\frac{\tilde{\kappa}}{2m}+\frac{\tilde{k}-\omega}{2m}\frac{\kappa^2}{p_1^2-\omega_2}\right)u_{0\alpha}, \tag{26}$$

where $u_{p\lambda}$ is the Dirac bispinor for an electron with momentum p and polarization λ. Here $T_{\lambda_1\alpha}$ is the amplitude of the single photoeffect, as a result of which a K-electron with polarization α absorbs a photon and is ejected from the atom having polarization λ_1. Inserting (11), (12), and (25) into (7), using (21), and performing certain transformations (details can be found in Ref. 17), we obtain

$$M_{\text{a}}=-K(\nu)T_{\lambda_1\alpha}w_{\lambda_2}^{\dagger}w_{\beta}, \quad \nu=\frac{\varepsilon_2}{I}=\left(\frac{p_2}{\eta}\right)^2, \tag{27}$$

$$K(\nu)=N_1 N_{p_2}\frac{m}{\eta^4}J(\nu), \quad N_{p_2}^2=\frac{2\pi/\sqrt{\nu}}{1-\exp(-2\pi/\sqrt{\nu})}, \tag{28}$$

$$J(\nu)=\frac{8\zeta^2}{(1+\zeta)^3}\left\{\frac{I_1}{\nu+1}-\frac{I_2}{\nu+2}\right\}, \quad \zeta=(\nu+2)^{-1/2}, \tag{29}$$

$$I_1=\exp\left(-\frac{2}{\sqrt{\nu}}\arctan\sqrt{\nu}\right)\int_0^1\frac{t^{-\zeta}(1-t)}{(1+st)^3}dt, \tag{30}$$

$$I_2=\int_0^1 \frac{t^{-\zeta}(1-t)^3}{(1+st)^3}\Phi_1(t)\Phi_2(t)\,dt, \quad s=\frac{1-\zeta}{1+\zeta}, \tag{31}$$

where

$$\Phi_1(t)=\exp\left(-\frac{2}{\sqrt{\nu}}\arctan\frac{\sqrt{\nu}(1-t)}{a+bt}\right),$$

$$\Phi_2(t)=\frac{(3\zeta^2+1)(1-t)^2+6\zeta(1-t^2)+2(1+t)^2}{[(2\zeta^2+1)(1-t)^2+4\zeta(1-t^2)+(1+t)^2]^2},$$

with $a=\sqrt{\nu+2}+2$, and $b=\sqrt{\nu+2}-2$. The amplitude and cross section of the double photoeffect in the edge region of the spectrum are[5)]

$$M^{++}_{\rm edge}=-K(\nu)[T_{\lambda_1\alpha}w^\dagger_{\lambda_2}w_\beta-T_{\lambda_1\beta}w^\dagger_{\lambda_2}w_\alpha], \tag{32}$$

$$d\sigma^{++}_{\rm edge}=\frac{(4\pi\alpha)^3}{2\omega}\overline{|M^{++}_{\rm edge}|^2}\frac{d\mathbf{p}_2\,d\mathbf{p}_1}{(2\pi)^5}\,\delta(E_1+E_2-2m-\omega). \tag{33}$$

The bar over the square of the amplitude indicates summation over the polarizations λ_1 and λ_2 of the final electrons and averaging over the photon polarizations λ_k:

$$\begin{aligned}\overline{|M^{++}_{\rm edge}|^2}&=\frac{1}{2}\sum_{\lambda_2\lambda_1\lambda_k}|M^{++}_{\rm edge}|^2\\&=K^2(\nu)\frac{1}{2}\sum_{\lambda_1\lambda_k}\{|T_{\lambda_1\alpha}|^2+|T_{\lambda_1\beta}|^2\}\\&=2K^2(\nu)\frac{1}{4}\sum_{\lambda_1\lambda_0\lambda_k}|T_{\lambda_1\lambda_0}|^2\\&=2K^2(\nu)\overline{|M^+|^2}.\end{aligned} \tag{34}$$

Here $\overline{|M^+|^2}$ is the square of the simple (single) photoeffect amplitude summed over the photoelectron polarizations λ_1 and averaged over the polarizations of the photon (λ_k) and the bound electron (λ_0). The differential cross section $d\sigma^+$ of the photoeffect on the K-shell (on two electrons) can be expressed in terms of this quantity:

$$d\sigma^+=\frac{4\pi\alpha}{\omega}\,\overline{|M^+|^2}\,\frac{d\mathbf{p}_1}{(2\pi)^2}\,\delta(E_1-m-\omega). \tag{35}$$

Since in the edge region we have $E_2-m=\varepsilon_2\sim I$, Eqs. (33) and (35) yield

$$d\sigma^{++}_{\rm edge}=\frac{2}{\pi}\alpha^2K^2(\nu)\,d\mathbf{p}_2\,d\sigma^+. \tag{36}$$

We direct the z axis along the photon momentum $\mathbf{k}$ and denote the solid angles into which the high- and low-energy electrons are ejected by $\Omega_1(\theta_1,\varphi_1)$ and $\Omega_2(\theta_2,\varphi_2)$, respectively. Substituting $d\mathbf{p}_2=(\eta^3/2)\sqrt{\nu}\,d\nu\,d\Omega_2$, $K^2(\nu)$ from Eq. (28), and $d\sigma^+$ from Ref. 19 in (36), we obtain

$$\frac{d\sigma^{++}_{\rm edge}}{d\nu\,d\Omega_2\,d\Omega_1}=\frac{Q(\nu)}{4\pi Z^2}\frac{d\sigma^+}{d\Omega_1}, \tag{37}$$

where

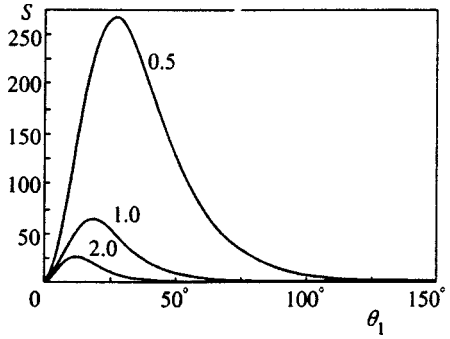


FIG. 2. Angular distributions of high-energy electrons from the edge region, $S(\theta_1)=C^{-1}d\sigma^+/d\Omega_1=Z^2B^{-1}C^{-1}d\sigma^{++}_{\rm edge}/d\Omega_1$, with $C=r_0^2\alpha^4Z^5$, and the value of B defined in Eq. (45). The numbers on the curves indicate the values of ω/m, the photon energies in units of electron mass.

$$Q(\nu)=\frac{8J^2(\nu)}{1-\exp(-2\pi/\sqrt{\nu})}, \tag{38}$$

$$\frac{d\sigma^+}{d\Omega_1}=r_0^2\alpha^4Z^5S(\theta_1), \quad r_0=\frac{\alpha}{m}, \tag{39}$$

$$S(\theta_1)=\left(\frac{2mp_1}{\kappa^2}\right)^3\left\{1-\frac{m}{\omega}+\frac{4m^3}{\omega\kappa^2}\right\}\sin^2\theta_1, \tag{40}$$

with $\kappa^2=(\mathbf{p}_1-\mathbf{k})^2$. The right-hand side of Eq. (37) is independent of the angle of the outgoing low-energy electron. Thus, in the edge region of the spectrum the low-energy electrons are distributed isotropically in solid angle. The angular distribution of the high-energy electrons is determined by the function $S(\theta_1)$, which also gives the angular distribution of the electrons in the single photoeffect. Figure 2 depicts this function for different photon energies. We see that as the photon energy rises, the angular distributions narrow and shift toward smaller angles, but no high-energy electrons are ejected in the forward direction ($\theta_1=0$). Formula (40) implies that there are no high-energy electrons emitted in the backward direction ($\theta_1=\pi$) either. Nonzero fluxes of such electrons for $\theta_1=0$ and $\theta_1=\pi$ can be obtained only if we allow for higher-order corrections in αZ to the amplitude and cross section.[19] Integrating (37) over the ejection angles, we arrive at the following expression for the energy distribution of low-energy electrons ($\varepsilon_2\ll m$):

$$\frac{d\sigma^{++}_{\rm edge}}{d\nu}=\frac{Q(\nu)}{Z^2}\sigma^+(\omega), \quad \nu\ll(\alpha Z)^{-2}, \tag{41}$$

$$\sigma^+(\omega)=\sigma_0Z^5\varphi(\omega), \quad \sigma_0=\pi r_0^2\alpha^4, \tag{42}$$

$$\varphi(\omega)=\frac{4m^2p_1^3}{\omega^5}\left\{\frac{4}{3}+\frac{E_1-2m}{E_1+m}\left(\frac{E_1}{m}-\frac{m}{p_1}\ln\frac{E_1+p_1}{m}\right)\right\}, \tag{43}$$

with $E_1=\omega+m$. If we use the definition (38) and Eqs. (29)–(31), we find that

$$Q(0)=0.168, \quad Q(\nu\gg1)\approx\frac{4}{\pi}\nu^{-7/2}. \tag{44}$$

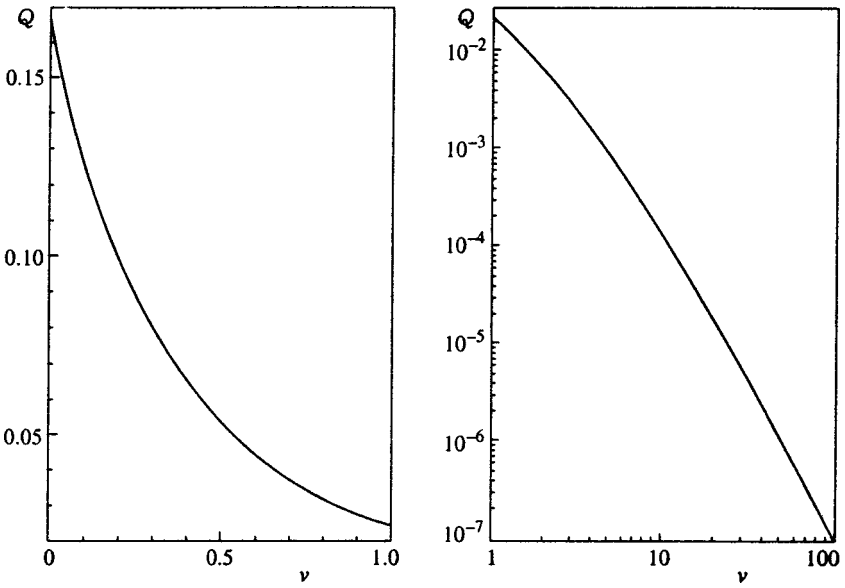


FIG. 3. Energy spectrum of low-energy electrons from the edge region, $Q(\nu)=Z^2\, d\sigma^{++}_{\rm edge}/\sigma^+\, d\nu$, with $\nu=\varepsilon_2/I$.

The values of Q for other values of ν can be found from the diagram in Fig. 3. Bearing in mind that Q rapidly decreases with increasing ν, in calculating the total contribution of the edge region of the spectrum we can replace the upper limit of integration with respect to ν by infinity:

$$\sigma^{++}_{\rm edge}=\frac{B}{Z^2}\sigma^+(\omega), \quad B=\int_0^\infty Q(\nu)\, d\nu=0.090. \tag{45}$$

A very close numerical result ($B=0.093$) was obtained by Amusia *et al.*[5] and Drukarev and Karpeshin,[14] but a consistent QED derivation of Eq. (41) is done only in the present paper. Equations (41) and (45) are valid for both relativistic and nonrelativistic double photoeffects only if the high-energy electron momentum p_1 is much higher than η. This requirement is needed if we want the expansion (15) to be meaningful.

2.2. Distribution of photoelectrons in the central region of the spectrum

The double ionization process may place with low momentum transfer to the nucleus, $q\sim\eta$. In this case the photoelectron energies occupy an interval in the central part of the energy spectrum ($E_1\sim E_2$). The limits of this interval, called the central region,[5] are established below. In the central region we must take into account all four diagrams of Fig. 1 (diagrams a–d), as well as the diagrams with the initial states interchanged (diagrams a′–d′). However, it is enough to calculate only the diagrams a and b, since the other diagrams can be obtained by interchanging the initial or final states or both states (e.g., diagrams c and d are obtained from a and b by the interchanges $\psi_\alpha\leftrightarrow\psi_\beta$ and $\psi_{p_1}\leftrightarrow\psi_{p_2}$). Since both final electrons are relativistic, for their wave functions we can take plane waves. The energy of the intermediate electron is also high. For this reason, we can expand the relativistic Coulomb Green's function in powers of the Coulomb parameters ξ and αZ and keep only the first term, i.e., the free-particle relativistic Green's function. As a result the amplitudes for diagrams a and b are

$$M_{\rm a}=\overline{u}_{p_1}\hat{e}G_{\rm a}(\boldsymbol{\kappa})\gamma^\mu\int\frac{d\mathbf{f}}{(2\pi)^3}\times\langle\mathbf{q}-\mathbf{f}|\psi_\alpha\rangle D(\mathbf{p}_2-\mathbf{f})\overline{u}_{p_2}\gamma_\mu\langle\mathbf{f}|\psi_\beta\rangle,$$

$$\boldsymbol{\kappa}=\mathbf{p}_1-\mathbf{k}, \quad \mathbf{q}=\mathbf{p}_2+\mathbf{p}_1-\mathbf{k}, \tag{46}$$

$$M_{\rm b}=\overline{u}_{p_1}\gamma^\mu\int\frac{d\mathbf{f}}{(2\pi)^3}G_{\rm b}(\mathbf{q}-\mathbf{f}+\mathbf{k})\hat{e}\langle\mathbf{q}-\mathbf{f}|\psi_\alpha\times D(\mathbf{p}_2-\mathbf{f})\overline{u}_{p_2}\gamma_\mu\langle\mathbf{f}|\psi_\beta\rangle, \tag{47}$$

where $G_{\rm a}$ and $G_{\rm b}$ are relativistic electron propagators (18) with energies $E_{\rm a}=m-\varepsilon_2$ and $E_{\rm b}=m+\omega$, respectively. Allowing for the fact that at $q\sim\eta$ the leading contribution to the integrals in (46) and (47) is provided by $f\sim\eta$, we have, to lowest order in αZ,

$$D(\mathbf{p}_2-\mathbf{f})\simeq\frac{1}{2m\varepsilon_2}, \quad G_{\rm b}(\mathbf{q}-\mathbf{f}+\mathbf{k})\simeq G(\mathbf{k}). \tag{48}$$

To the same order in αZ, the wave function of the bound state is the product of the spatial nonrelativistic function $|1s\rangle$ and the Dirac bispinor u_0:

$$|\psi_i\rangle=|1s\rangle u_{0i}, \quad u_{0i}=\begin{pmatrix} w_i \\ 0\end{pmatrix}, \quad i=\alpha,\beta,$$

$$\langle\mathbf{f}|1s\rangle=N_1\left(-\frac{\partial}{\partial\eta}\right)\langle\mathbf{f}|V_{i\eta}|0\rangle. \tag{49}$$

The normalization constant N_1, the Pauli spinors w_i, and the matrix elements $\langle\mathbf{f}|V_{i\eta}|0\rangle$ have been defined in (12)–(14).

Inserting (48) and (49) into (46) and (47), we obtain

$$M_{\rm a}=\frac{\Phi(\mathbf{q})}{2m\varepsilon_2}\overline{u}_{p_1}\hat{e}G_{\rm a}(\boldsymbol{\kappa})\gamma^\mu u_{0\alpha}\overline{u}_{p_2}\gamma_\mu u_{0\beta}, \tag{50}$$

$$M_{\rm b}=\frac{\Phi(\mathbf{q})}{2m\varepsilon_2}\overline{u}_{p_1}\gamma^\mu G_{\rm b}(\mathbf{k})\hat{e}u_{0\alpha}\overline{u}_{p_2}\gamma_\mu u_{0\beta}, \tag{51}$$

$$\Phi(\mathbf{q})=\int\frac{d\mathbf{f}}{(2\pi)^3}\langle\mathbf{q}-\mathbf{f}|1s\rangle\langle\mathbf{f}|1s\rangle=\left(\frac{4\eta^2}{q^2+4\eta^2}\right)^2. \quad (52)$$

The amplitudes for the other diagrams can easily be derived from (50) and (51). The total amplitude of the process in the central region is

$$M^{++}_{\text{central}}=M_{\text{a}}+M_{\text{b}}+M_{\text{c}}+M_{\text{d}} - M_{\text{a}'}-M_{\text{b}'}-M_{\text{c}'}-M_{\text{d}'}. \quad (53)$$

Calculating the square of its absolute value, summing over the polarizations of the final electrons, and averaging over the polarizations of the initial electrons and photon, we arrive at the formula

$$\overline{|M^{++}_{\text{central}}|^2}=\frac{\Phi^2(q)W(E_1)}{(2m)^4E_1E_2}, \quad (54)$$

where

$$W(E_1)=\left(\frac{\varepsilon_1-\varepsilon_2}{\varepsilon_1\varepsilon_2}\right)^2\left\{m^2+E_1E_2-p_{1n}p_{2n}+\left(\frac{m\omega}{\varepsilon_1\varepsilon_2}\right)^2 \times(2m\omega+m^2-E_1E_2+p_{1n}p_{2n})\right\}, \quad (55)$$

with $\varepsilon_1=E_1-m$, $\varepsilon_2=E_2-m$, $p_{1n}=\mathbf{p}_1\cdot\mathbf{n}$, $p_{2n}=\mathbf{p}_2\cdot\mathbf{n}$, and $\mathbf{n}=\mathbf{k}/|\mathbf{k}|$. Equation (55) clearly shows that W depends on E_1, E_2, and the angles of the ejected electrons. However, using the laws of energy conservation ($\varepsilon_1+\varepsilon_2=\omega$) and momentum conservation ($\mathbf{q}=0$), we can express E_2, p_{1n}, and p_{2n} in terms of E_1:

$$E_2=2m+\omega-E_1, \qquad p_{1n}=\frac{E_1^2-E_2^2+\omega^2}{2\omega}, \quad p_{2n}=p_{1n}(1\leftrightarrow 2). \quad (56)$$

Substituting (54) for $\overline{|M^{++}_{\text{edge}}|^2}$ in Eq. (33), we arrive at an expression for the differential cross section of the double photoeffect in the central region:

$$d\sigma^{++}_{\text{central}}=\frac{\alpha r_0^2}{16\pi^2}\Phi^2(q)W(E_1)\frac{d\mathbf{p}_2\,d\mathbf{p}_1}{m^2\omega E_2E_1}\delta(\varepsilon_2+\varepsilon_1-\omega). \quad (57)$$

This cross section (57) is proportional to $\Phi^2(q)$, which is of order unity in the central part of the spectrum and rapidly decreases outside the central part, reaching values of order $(\alpha Z)^8$ at $q\sim m$. In (57) we go from the variable $\mathbf{p}_2$ to the variable $\mathbf{q}$.

As in Sec. 2.1, we direct the z axis along $\mathbf{k}$ and denote the solid angles into which the electrons are ejected by Ω_1 and Ω_2. For $\mathbf{q}$ and Ω_1 fixed, E_2 depends on E_1, so that the removal of the delta function by integration with respect to E_1 gives rise to the factor

$$\int dE_1\,\delta(E_1+E_2-2m-\omega)=\left|1+\frac{\partial E_2}{\partial E_1}\right|^{-1}, \quad (58)$$

where the derivative $\partial E_2/\partial E_1$ is taken at a value of E_1 that satisfies the equations

$$E_2+E_1=\omega+2m, \quad \mathbf{p}_2+\mathbf{p}_1=\mathbf{k}+\mathbf{q}. \quad (59)$$

In the second equation in (59) we can put $q=0$, since the we are dealing with small values of q. As a result we find

$$E_2=\sqrt{E_1^2+\omega^2-2\omega p_1t_1}, \qquad \frac{\partial E_2}{\partial E_1}=\frac{E_1}{E_2}\left(1-\frac{\omega t_1}{p_1}\right), \quad t_1=\cos\theta_1, \quad (60)$$

and the function $E_1(t_1)$, which can easily be found by the formulas written below.

According to (58), the phase volume in (57) can be written

$$d\mathbf{p}_2\,d\mathbf{p}_1\,\delta(\varepsilon_1+\varepsilon_2-\omega)=d\mathbf{q}\,d\Omega_1\frac{p_1E_1}{\chi(t_1)}, \quad (61)$$

where

$$\chi(t_1)=\left|1+\frac{\partial E_2}{\partial E_1}\right|=\frac{E_0p_1-\omega E_1t_1}{E_2p_1}, \quad (62)$$

with $E_0=2m+\omega$. Since the dependence on $\mathbf{q}$ in (57) is contained only in the factor $\Phi^2(q)$, integration over $\mathbf{q}$ is easy:

$$\int\Phi^2(q)\,d\mathbf{q}=\pi^2\eta^3. \quad (63)$$

As a result,

$$d\sigma^{++}_{\text{central}}=\frac{A}{16\pi}\frac{mp_1}{\omega E_2}\frac{W(E_1)}{\chi(t_1)}d\Omega_1, \quad A=\pi r_0^2\alpha^4Z^3=\sigma_0Z^3. \quad (64)$$

The formula become more compact if we put $m=1$ and introduce a new variable x in terms of which the kinetic energies of the electrons can be expressed:

$$\varepsilon_1=\frac{\omega}{2}(1+x), \quad \varepsilon_2=\frac{\omega}{2}(1-x), \quad 0\leqslant x\leqslant 1. \quad (65)$$

Then

$$W(E_1)=\frac{16}{\omega}F(x), \quad (66)$$

where

$$F(x)=\left(\frac{x}{1-x^2}\right)^2\left\{1+\left(\frac{x}{x_0}\right)^2+\frac{2}{\omega}+\left(\frac{4}{\omega}\right)^2\frac{1-x^2/x_0^2}{(1-x^2)^2}\right\}, \quad (67)$$

with $x_0=\sqrt{\omega/(\omega+1)}$, and the angular distribution of the high-energy electrons (when the kinetic exceeds $\omega/2$) from the central region assumes the form

$$\frac{d\sigma^{++}_{\text{central}}}{d\Omega_1}=\frac{A}{\pi\omega^2}\frac{p_1}{E_2}\frac{F(x)}{\chi(t_1)}, \quad (68)$$

where

$$x=x(t_1)=\frac{-E_0\omega(1-t_1^2)+2t_1\sqrt{4\omega+3\omega^2+\omega^2t_1^2}}{E_0^2-\omega^2t_1^2}, \quad (69)$$

for $t_1 \geqslant \sqrt{\omega/(\omega+4)}$. The restriction on t_1 follows from the requirement that x be positive. At $t_1=\sqrt{\omega/(\omega+4)}$ we have $x=0$. As t_1 increases, x reaches its maximum value $x=x_0$ at $t_1=1$. Thus, the range of values of x consistent with the equality $\mathbf{q}=0$ extends from $x=0$ to $x=x_0$.

The angular distribution of the low-energy electrons (when the kinetic energy is less than $\omega/2$) can be obtained from (57) if we replace $d\mathbf{p}_1$ with $d\mathbf{q}$. Then, reasoning along the same lines as we did in deriving (68), we obtain

$$\frac{d\sigma^{++}_{\text{central}}}{d\Omega_2}=\frac{A}{\pi\omega^2}\frac{p_2}{E_1}\frac{F(x)}{\chi(t_2)},$$
$$\chi(t_2)=\frac{E_0p_2-\omega E_2t_2}{E_1p_2}, \quad t_2=\cos\theta_2, \tag{70}$$

where

$$x=x(t_2)=\frac{E_0\omega(1-t_2^2)-2t_2\sqrt{4\omega+3\omega^2+\omega^2t_2^2}}{E_0^2-\omega^2t_2^2}, \tag{71}$$

for $t_2 \leqslant \sqrt{\omega/(\omega+4)}$.

In the central region, the angle of an ejected electron and the electron energy are linked by a rigorous relationship:

$$t_1=\frac{E_1^2-E_2^2+\omega^2}{2\omega p_1}=\frac{E_0x+\omega}{2p_1}. \tag{72}$$

Using the formula that follows from (72),

$$d\Omega_1=\frac{\pi E_2}{p_1}\chi(t_1)\,dx, \tag{73}$$

we can easily go from the angular distribution (68) to the energy distribution:

$$\frac{d\sigma^{++}_{\text{central}}}{dx}=\frac{A}{\omega^2}F(x), \quad 0\leqslant x\leqslant x_0. \tag{74}$$

The contribution of the central region to the total cross section of the double photoeffect is given by the formula

$$\sigma^{++}_{\text{central}}=Af(\omega), \tag{75}$$

where

$$\omega^2 f(\omega)=\int_0^{x_0}F(x)\,dx$$
$$=I_1+\frac{1}{\omega}I_2+\left(\frac{4}{\omega}\right)^2\left(I_3-\frac{1}{\omega}I_4\right), \tag{76}$$

with

$$I_1=x_0(\omega+2)-L, \quad L=\ln\frac{1+x_0}{1-x_0},$$
$$I_2=\frac{x_0}{2}(3\omega+5)-\frac{5}{4}L,$$
$$I_3=\frac{x_0}{4}\left[(\omega+1)^2-\frac{1}{2}(\omega+1)-\frac{1}{4x_0}L\right],$$
$$I_4=\frac{x_0}{6}\left[(\omega+1)^3-\frac{7}{4}(\omega+1)^2+\frac{3}{8}(\omega+1)+\frac{3}{16x_0}L\right].$$

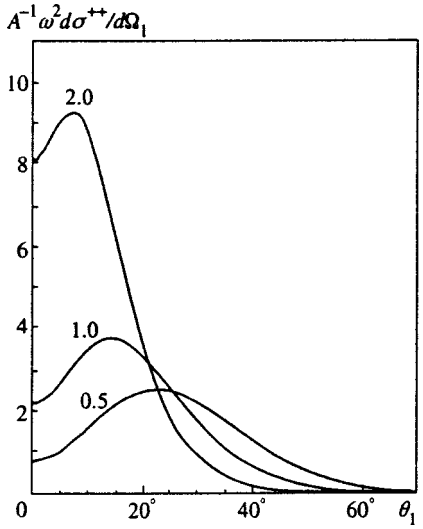


FIG. 4. Angular distributions of high-energy electrons from the central region. The numbers on the curves indicate the values of ω/m, and the value of A is defined in Eq. (64).

3. RESULTS AND DISCUSSION

The angular distributions of the high- and low-energy electrons (Eqs. (68) and (70), respectively) belonging to the central region are depicted in Figs. 4 and 5 for three photon energies: $\omega=0.5$, 1.0, and 2.0 (in units of electron mass). A characteristic feature of the double photoeffect is the nonzero forward ($\theta_1=0$) and backward ($\theta_2=\pi$) electron emission. The differential cross section at $\theta=0$ for the single photoeffect is finite only if we allow for higher-order Coulomb corrections and is of order $r_0^2\alpha^6Z^7$ (see Ref. 19). At moderate values of Z the contribution of the double photoeffect to the forward electron emission ($\sim r_0^2\alpha^4Z^3$) may become much larger than the contribution of the single photoeffect. To separate the contribution of the single photoeffect at small angles, the electron energy must be fixed at $E=\omega+1$.

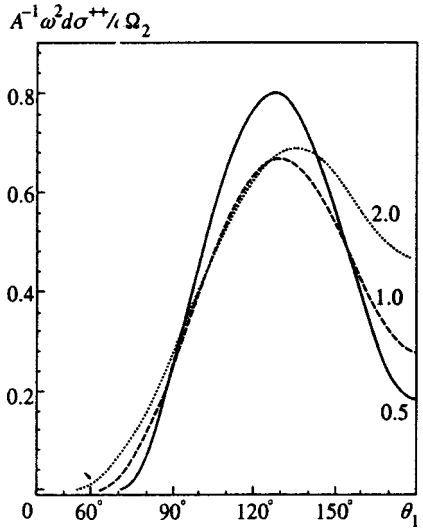


FIG. 5. Angular distributions of low-energy electrons from the central region. The notation is the same as in Fig. 4.

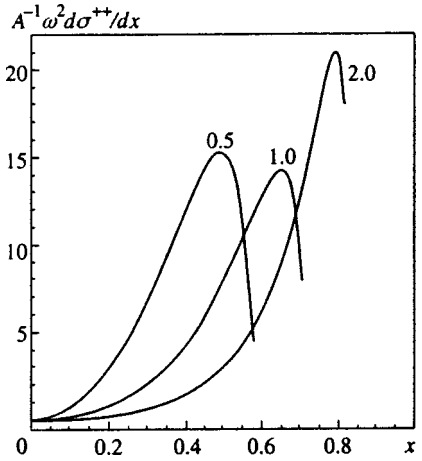


FIG. 6. Energy distribution of high-energy photons from the central region, $x=2\varepsilon_1/\omega-1$. The rest of the notation is the same as in Fig. 4.

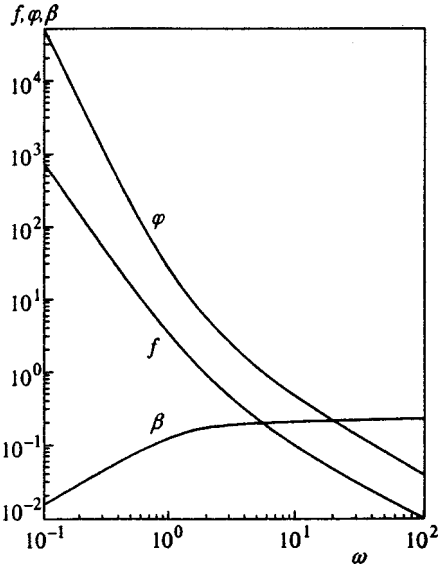


FIG. 7. Dependence on the photon energy ω of the total contribution to the double photoeffect of the electrons from the central region $f(\omega)=\sigma^{++}_{\text{central}}/\sigma_0 Z^3$, of the total cross section of the single photoeffect $\varphi(\omega)=\sigma^{+}/\sigma_0 Z^5$, and of their ratio $\beta(\omega)=Z^2\sigma^{++}_{\text{central}}/\sigma^{+}$. The values of ω are in units of m, and $\sigma_0=\pi r_0^2\alpha^4$.

The energy distribution (74) for three values of ω is depicted in Fig. 6. The vanishing of the cross sections at $x=0$ means that two electrons with the same energy cannot be ejected simultaneously. This is not a trivial fact, since ejection is not forbidden by the energy and momentum conservation laws (Eq. (59)) at $\mathbf{q}=0$. A special study of this problem done in the Appendix shows that the contributions to the amplitude of all the Feynman diagrams describing the double effect in the present approximation (Fig. 1 depicts only half the number of such diagrams) cancel out at $x=0$ and $\mathbf{q}=0$. Note that in Ref. 14 the cross section of the process reaches its minimum at $E_1=E_2$ but does not vanish (and in fact is not small), which points to an error. Drukarev and Karpeshin[14] may not have accounted for all the Feynman diagrams describing the process in the present approximation, since the formula for the cross section simplifies only if all the diagrams are taken into account (in Ref. 14 the formula is much more complicated), and the amplitude of the process vanishes when the electrons have equal energies.

Figure 6 shows that in the central region, as in the edge region, the photon energy is not evenly distributed among the electrons: the maximum of the cross section appears at values of x close to $x_0(\omega)$. All the curves end at $x=x_0(\omega)$. We do not consider the behavior of the cross section for $x>x_0(\omega)$, since here the cross section drops by two to three orders of magnitude when we move away from $x_0(\omega)$ only by $\Delta x\sim\alpha Z$ (see Ref. 17). However, in the narrow edge region $1-x\sim\alpha^2Z^2$ studied in Sec. 2 the cross section rapidly increases with x, reaching its maximum value at $x=1$ (see Eq. (41) and Fig. 2).

The total cross section of the relativistic double photoeffect is determined by the sum of the contributions of the edge (Eq. (45)) and central (Eq. (75)) regions:

$$\sigma^{++}=\sigma^{++}_{\text{edge}}+\sigma^{++}_{\text{central}}=A\{B\varphi(\omega)+f(\omega)\}=\frac{\sigma^{+}(\omega)}{Z^2}\{B+\beta(\omega)\},\tag{77}$$

where

$$\beta(\omega)=\frac{f(\omega)}{\varphi(\omega)}.\tag{78}$$

For the ratio of the cross sections of double and single ionization we obtain a simple formula:

$$R=\frac{\sigma^{++}}{\sigma^{+}}=\frac{B+\beta(\omega)}{Z^2}.\tag{79}$$

The functions $\varphi(\omega)$, $f(\omega)$, and $\beta(\omega)$ are depicted in Fig. 7. While $\varphi(\omega)$ and $f(\omega)$ decrease with increasing ω, their ratio $\beta(\omega)$ increases, with the rapid increase at $\omega<1$ changing to the slow increase at $\omega>1$. Already at $\omega>0.7$ the ratio $\beta(\omega)$ exceeds B, i.e., the contribution of the central region to the cross section becomes larger than the contribution of the edge section.

At low and high photon energies we can derive simple formulas for the functions φ, f, and β. For instance, in the nonrelativistic domain $\alpha^2Z^2\ll\omega\ll1$ we have

$$\varphi(\omega)=\frac{32\sqrt{2}}{3}\omega^{-7/2},\quad f(\omega)\frac{32}{15}\omega^{-5/2},\quad \beta(\omega)=\frac{\omega}{5\sqrt{2}},\tag{80}$$

and the ratio R (Eq. (79)) differs little from the constant B/Z^2. Drukarev[13] used Coulomb functions for his calculations and a value of $\beta(\omega)$ that was larger by a factor of two than the one in (80). The explanation lies in the fact the in removing the delta function Drukarev[13] did not account for the factor χ^{-1} (Eq. (58)), which arises when one replaces the phase volume $d\mathbf{p}_2$ with $d\mathbf{q}$. In the nonrelativistic domain $\chi^{-1}=1/2$, which follows from (62).

At high photon energies ($\omega\gg1$) we have

$$\varphi(\omega)=\frac{4}{\omega}\left(1+\frac{7}{3\omega}\right),$$

$$f(\omega)=\frac{1}{\omega}\left(1+\frac{3-\ln 4\omega}{\omega}\right), \tag{81}$$

$$\beta(\omega)=\frac{1}{4}\left(1+\frac{2/3-\ln 4\omega}{\omega}\right).$$

The σ^{++}-to-σ^{+} ratio (79) tends to a constant limit $R(\infty)$ as $\omega\to\infty$:[6)]

$$R(\infty)=\frac{B+0.25}{Z^2}=\frac{0.34}{Z^2}. \tag{82}$$

This value[7)] is almost four times larger than the corresponding nonrelativistic limit B/Z^2.

ACKNOWLEDGMENTS

We would like to express our gratitude to M. Ya. Amus'ya (Amusia), V. G. Gorshkov, E. G. Drukarev, and L. N. Labzovskiĭ for fruitful discussions.

APPENDIX

We wish to prove that in the approximation adopted in this paper (the first order in Z^{-1} and the lowest order in αZ) the double-photoeffect amplitude vanishes when the ejected electrons have the same energy.

As shown in Sec. 2.2 (Eqs. (50) and (51)), each Feynman diagram in Fig. 1 in the central region of the electron spectrum can be written

$$M_i=\Phi(q)L_i(q=0), \quad i=\mathrm{a,b,\ldots,d'}. \tag{A1}$$

Introducing the four-vectors

$$e=(0,\mathbf{e}), \quad k=(\omega,\mathbf{k}), \quad p_1=(E_1,\mathbf{p}_1), \quad p_2=(E_2,\mathbf{p}_2), \tag{A2}$$

$$j_{1\alpha}=(\bar{u}_{p_1}\gamma_0 u_{0\alpha},\ \bar{u}_{p_1}\boldsymbol{\gamma}u_{0\alpha}),$$

$$j_{2\beta}=(\bar{u}_{p_2}\gamma_0 u_{0\beta},\ \bar{u}_{p_2}\boldsymbol{\gamma}u_{0\beta}),$$

setting $E_1=E_2$ and $\mathbf{q}=0$ in (59), and performing simple transformations, from (50), (51), and (A1) we obtain

$$L_\mathrm{a}+L_\mathrm{b}=(m\omega)^{-2}\{(ej_{1\alpha})(kj_{2\beta}) -(ej_{2\beta})(kj_{1\alpha})-(ep_1)(j_{1\alpha}j_{2\beta})\}. \tag{A3}$$

Doing the interchanges $1\leftrightarrow 2$ and $\alpha\leftrightarrow\beta$, we find the contribution of the diagrams c and d of Fig. 1:

$$L_\mathrm{c}+L_\mathrm{d}=(m\omega)^{-2}\{(ej_{2\beta})(kj_{1\alpha}) -(ej_{1\alpha})(kj_{2\beta})-(ep_2)(j_{2\beta}j_{1\alpha})\}. \tag{A4}$$

The expression ab stands for the scalar product of two four-vectors: $ab=a_0b_0-\mathbf{a}\cdot\mathbf{b}$.

Adding (A3) to (A4) and allowing for (5), we arrive at

$$\sum_{i=\mathrm{a}}^{\mathrm{d}} L_i=(m\omega)^{-2}\{-e(p_1+p_2)(j_{1\alpha}j_{2\beta})\} =(m\omega)^{-2}(\mathbf{e}\cdot\mathbf{k})(j_{1\alpha}j_{2\beta})=0, \tag{A5}$$

with $\mathbf{p}_1+\mathbf{p}_2=\mathbf{k}$ at $\mathbf{q}=0$.

The contribution of the diagrams a′–d′ can be obtained from (A5) by the interchange $\alpha\leftrightarrow\beta$ and therefore is also zero. Thus, the sums of the primed and unprimed diagrams in Fig. 1 are equal to zero separately. Note that at $E_1=E_2$ the sum $\Sigma_i L_i(q)$ is of order q, but allowing for such terms would lead to a correction $\sim\alpha^2Z^2$ to the total cross section of the process.

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

1)In this paper we use the relativistic system of units, in which $\hbar=c=1$.

2)Primed letters indicate diagrams obtained from the corresponding diagrams in Fig. 1 by an interchange of the initial states (a→a′, etc.).

3)The subscript α on the wave function should not be confused with the fine-structure constant α.

4)Integrals of type (21) have been considered by Gorshkov and Polikanov.[16,18]

5)The normalization constants $1/\sqrt{2E_i}$ from the electron wave functions are incorporated into the respective bispinors u_{p_i}.

6)The same value as in Eq. (82) was obtained in Ref. 17 for the ratio of the cross sections of double and single ionization of an atom in one-photon annihilation of an ultrarelativistic positron and *K*-electron.

7)Drukarev and Karpeshin[14] arrived at the following values for β and R in the ultrarelativistic limit: $\beta(\infty)=0.5$ and $R(\infty)=0.59/Z^2$. Their value of $\beta(\infty)$ is almost twice as large as ours, probably for reasons related to the loss of the factor 1/2 in averaging over photon polarizations. We cannot indicate the reason more exactly since Ref. 14 does not contain the intermediate calculations.

[1] T. A. Carlson, Phys. Rev. **156**, 142 (1967).
[2] F. W. Byron, Jr. and C. J. Joachain, Phys. Rev. **164**, 1 (1967).
[3] R. L. Brown, Phys. Rev. A **1**, 586 (1970).
[4] T. Aberg, Phys. Rev. A **2**, 1726 (1970).
[5] M. Ya. Amusia, E. G. Drukarev, V. G. Gorshkov, and M. P. Kazachkov, J. Phys. B **18**, 1248 (1975).
[6] S. L. Carter and H. P. Kelly, Phys. Rev. A **24**, 170 (1981).
[7] T. Ishihara, K. Hino, and J. H. McGuire, Phys. Rev. A **44**, R6980 (1991).
[8] A. Dalgarno and H. R. Sadeghpour, Phys. Rev. A **46**, R3591 (1992).
[9] J. C. Levin, I. A. Sellin, B. M. Johnson, D. W. Lindle, R. D. Miller, N. Berrah, Y. Azuma, H. G. Berry, and D.-H. Lee, Phys. Rev. A **47**, R16 (1993).
[10] M. A. Kornberg and J. E. Miraglia, Phys. Rev. A **49**, 5120 (1994).
[11] R. C. Forrey, H. R. Sadeghpour, J. D. Baker, J. D. Morgan, and A. Dalgarno, Phys. Rev. A **51**, 2112 (1995).
[12] J. C. Levin, G. B. Armen, and I. A. Sellin, Phys. Rev. Lett. **76**, 1220 (1996).
[13] E. G. Drukarev, Phys. Rev. A **51**, R2684 (1995).
[14] E. G. Drukarev and F. F. Karpeshin, J. Phys. B **9**, 399 (1976).
[15] W. H. Furry, Phys. Rev. **81**, 15 (1951).
[16] V. G. Gorshkov, Zh. Éksp. Teor. Fiz. **47**, 352 (1964) [Sov. Phys. JETP **20**, 234 (1965)].
[17] A. I. Mikhaĭlov and I. A. Mikhaĭlov, Zh. Éksp. Teor. Fiz. **113**, 786 (1998) [JETP **86**, 429 (1998)].
[18] V. G. Gorshkov and V. S. Polikanov, JETP Lett. **9**, 279 (1969).
[19] V. G. Gorshkov, A. I. Mikhaĭlov, and V. S. Polikanov, Nucl. Phys. **55**, 273 (1964).

Translated by Eugene Yankovsky

The problem of two electrons in an external field and the method of integral equations in optics

O. N. Gadomsky*) and K. K. Altunin

Ul'yanovsk State University, 432700 Ul'yanovsk, Russia
(Submitted 19 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1555–1577 (November 1998)

This paper solves the problem of the interaction, via the field of virtual photon field with the emission or absorption of a real photon, of two atomic electrons located at arbitrary distances from one another. The interaction is interpreted as a third-order QED effect in the coordinate representation. The role of intermediate states with positive and negative frequencies is studied. A general expression is derived for the matrix elements of the operator of the effective electron–electron interaction energy for different types of quantum transitions. The expression makes it possible to calculate the probabilities of the corresponding transitions and to examine various patterns of induction of polarizing fields by one atom at the point occupied by the other atom. The exchange of virtual photons between the atoms located at arbitrary distances from one another is shown to lead to additional terms in the operators of spin–orbit and spin–spin coupling of the atomic electrons, over and above those in the corresponding Breit operators. It is shown that there is an important difference between the induction of polarizing fields and the transfer of optical photons. In particular, it is found that when polarizing fields are induced, a situation may arise in which the disappearance (production) of a photon takes place at the point occupied by one atom, while absorption (emission) of the same photon occurs at the place occupied by the other atom. A block diagram of an experimental device that could be used to study this property of polarizing fields is presented. Finally, a method of deriving integral field equations is proposed. The method is based on allowing for polarizing fields, and its effectiveness is demonstrated by the example of electric dipole and spin transitions in the spectrum of interacting atomic electrons. © *1998 American Institute of Physics.* [S1063-7761(98)00211-X]

1. INTRODUCTION

In classical electrodynamics, a system of interacting charges in motion can be examined by introducing retarded potentials. Following Darwin[1] (see also Ref. 2), one can introduce a Lagrangian function for this system to within terms of order $(v/c)^2$, where v is the velocity of the moving charges, and c is the speed of light in vacuum. It is in this approximation that the system of charges can be examined without allowing for a light field. The passage to a system of interacting neutral atoms followed by the passage to a Hamiltonian operator was carried out in Refs. 3 and 4. The effect of a light field was taken into account in such a system phenomenologically by simply replacing the momenta $\mathbf{p}_i$ of the atomic electrons with $\mathbf{p}_i-(e/c)\mathbf{A}_i$, where i is the particle number, e is the electron charge, $\mathbf{A}_i$ is the vector potential of the light field at the point occupied by the ith atom. An interaction Hamiltonian for an N-atom system including the light field was derived and then successfully used in explaining the anomalous density dependence of the intensity of light echo[5] by the presence in the intensity of light echo of terms proportional to N^3 and N^4 (in addition to terms proportional to N^2).

After Darwin, the theoretical study of the two-electron problem was continued by Breit,[6] Landau,[7] and Bethe and Fermi,[8] who used the Dirac equation and the QED perturbation method. An operator describing the interaction of two electrons via the field of virtual photons was derived as a second-order perturbative effect. However, the use of the Breit operator is limited to moderate distances between the electrons. Hence research into the problem of two electrons belonging to two atoms located at arbitrary distances from one another was revived in the early 1970s in connection with intensive studies of the behavior of multiatomic systems in a light field. For instance, Chang and Stehle[9] examined the resonant interaction of two neutral atoms separated by distances greater than $2\pi c/\omega_0$, where ω_0 is the resonance frequency in the spectrum of the interacting atoms. This effect was considered a second-order QED effect without emission or absorption of real photons in the electric dipole approximation. The resonant interaction of two hydrogenlike atoms located at arbitrary distances from one another was examined in Ref. 10 as a second-order QED effect with allowance for orbital and spin degrees of freedom. There a relativistic operator describing the interaction of two electrons to within terms of order $1/c^2$ was derived, but in the passage to the nonrelativistic limit the focus was also only on electric dipole transitions.

Lifshitz,[11] Fedyushin,[12] and Akhiezer and Berestetskiĭ[13] studied the interaction of two electrons as a third-order QED effect with emission or absorption of a real photon under the

very stringent condition that the interacting particles have well-defined momenta before and after collision. This is the case of the interaction of two free particles. However, the situation is inappropriate for studying quantum transitions between states of atomic electrons, where the constants of motion are the square of the total angular momentum and the projection of the total angular momentum on the quantization axis. In the present paper we examine an entirely new problem: the interaction of two atomic electrons located at arbitrary distances from one another, with emission (absorption) of a real photon in the coordinate representation and with allowance for various quantum transitions and intermediate states.

Drake[14] examined the interaction of two atomic electrons in a heliumlike atom as a QED third-order effect and allowed for intermediate states with positive and negative energies. The role of intermediate positron states in the probability of quantum transitions accompanied by emission (absorption) of a real photon was estimated.

The next step in studying the problem of two electrons was taken in Refs. 5, 3, and 10 for the case of two hydrogenlike atoms located at arbitrary distances from one another. There formulas for electron and positron polarizing fields were derived in the electric dipole approximation.

As shown in Ref. 10, an electron polarizing field, interpreted as a third-order QED effect, is the field of the surrounding dipoles in the *N*-atom system. By summing the polarization fields in such a system we are able to derive an integral equation describing the propagation of light in a dielectric medium.[4] We have successfully used this equation (as have other researchers) to solve various boundary-value problems of classical,[15] nonlinear,[4] and quantum optics.[16,17] The near-field effect was studied in Ref. 16.

The present paper is a step in the theoretical studies of the problem of two electrons belonging to two hydrogenlike atoms located at arbitrary distances from one another. We derive an expression for the operator of the electron–electron interaction via the exchange of virtual photons. The expression allows for additional retardation in the spin–orbit and spin–spin interactions of the two electrons. In our research into the problem we examine the various properties of polarizing fields in the two-electron system.

A detailed investigation of the two-electron problem is important if we want to derive new integral equations that describe the propagation of photons in a medium in accordance with the different types of quantum transitions in the spectrum of the interacting atoms. In the present paper we use a specific example to derive integral equations for the

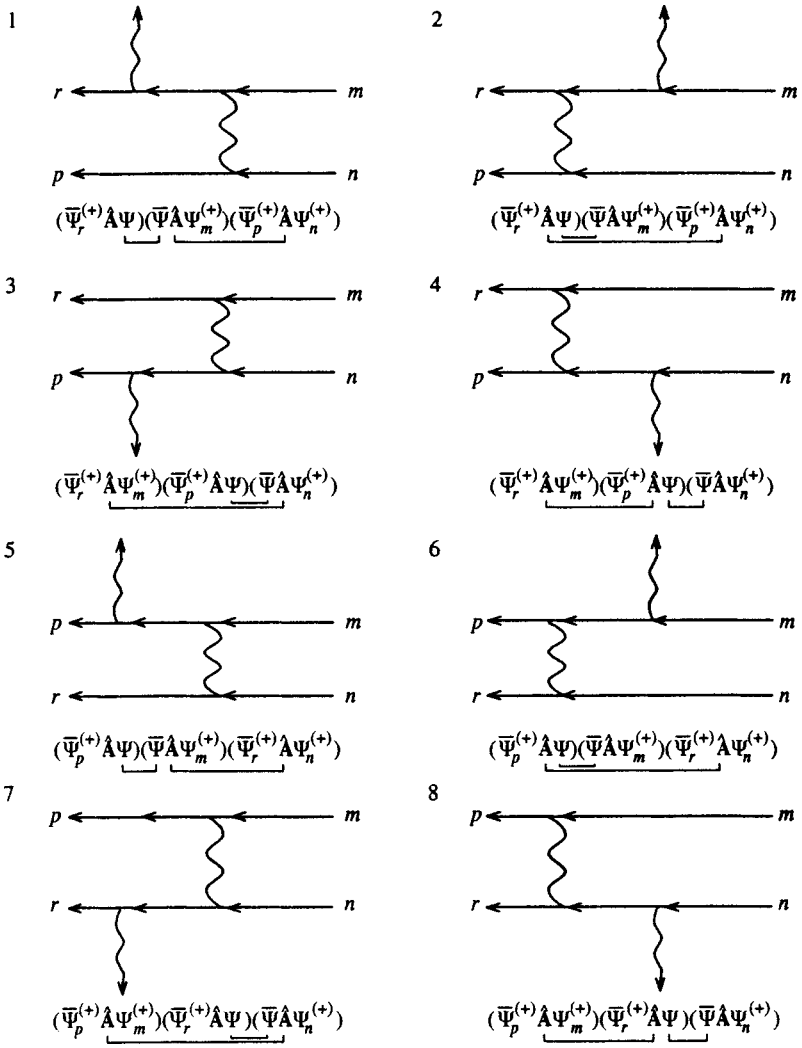


FIG. 1. Feynman diagrams for the electron–electron interaction of two atoms with emission or absorption of a photon.

electric and magnetic fields in dipole and spin media.

2. THE EFFECTIVE INTERACTION ENERGY MATRIX FOR TWO HYDROGENLIKE ATOMS LOCATED AT ARBITRARY DISTANCES FROM ONE ANOTHER

The Feynman diagrams for the electron–electron interaction are depicted in Fig. 1. Integrating in the S-matrix with respect to time, frequencies, and wave vectors of the virtual photons, we arrive at an expression for the matrix of the effective interaction energy of two atomic electrons (here $\hbar=c=1$):

$$
\begin{aligned}
U^{(3)}_{i\to f}=e^3\int\Bigg(d\mathbf{r}'\,d\mathbf{r}''\,d\mathbf{r}'''\Bigg\{&-\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}\\
&\times\sum_{l_\pm}\frac{\exp(i|\omega_n^{(+)}-\omega_p^{(+)}||\mathbf{r}''-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_r^{(+)}}\overline{\Psi}_r^{(+)}(\mathbf{r}')\hat{\mathbf{A}}(\mathbf{r}')\Psi_l\\
&\times(\mathbf{r}')\overline{\Psi}_l(\mathbf{r}'')\gamma_\mu''\Psi_m^{(+)}(\mathbf{r}'')\overline{\Psi}_p^{(+)}(\mathbf{r}''')\gamma_\mu'''\Psi_n^{(+)}(\mathbf{r}''')\\
&-\frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\sum_{l_\pm}\frac{\exp(i|\omega_n^{(+)}-\omega_p^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)-\omega-\omega_m^{(+)}}\overline{\Psi}_r^{(+)}\\
&\times(\mathbf{r}')\gamma_\mu'\Psi_l(\mathbf{r}')\overline{\Psi}_l(\mathbf{r}'')\hat{\mathbf{A}}(\mathbf{r}'')\Psi_m^{(+)}(\mathbf{r}'')\overline{\Psi}_m^{(+)}\\
&\times(\mathbf{r}''')\gamma_\mu'''\Psi_n^{(+)}(\mathbf{r}''')-\frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\\
&\times\sum_{l_\pm}\frac{\exp(i|\omega_r^{(+)}-\omega_m^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_p^{(+)}}\\
&\times\overline{\Psi}_r^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_m^{(+)}(\mathbf{r}')\overline{\Psi}_p^{(+)}\\
&\times(\mathbf{r}'')\hat{\mathbf{A}}(\mathbf{r}'')\Psi_l(\mathbf{r}'')\overline{\Psi}_l(\mathbf{r}''')\\
&\times\gamma_\mu'''\Psi_n^{(+)}(\mathbf{r}''')-\frac{1}{|\mathbf{r}'-\mathbf{r}''|}\\
&\times\sum_{l_\pm}\frac{\exp(i|\omega_r^{(+)}-\omega_m^{(+)}||\mathbf{r}'-\mathbf{r}''|)}{\omega_l(1-i0)-\omega-\omega_n^{(+)}}\overline{\Psi}_r^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_m^{(+)}\\
&\times\ (\mathbf{r}')\overline{\Psi}_p(\mathbf{r}'')\gamma_\mu''\Psi_l(\mathbf{r}'')\overline{\Psi}_l(\mathbf{r}''')\hat{\mathbf{A}}(\mathbf{r}''')\Psi_n^{(+)}(\mathbf{r}''')\\
&+\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}\sum_{l_\pm}\frac{\exp(i|\omega_r^{(+)}-\omega_n^{(+)}||\mathbf{r}''-\mathbf{r}'''|)}{\omega_l(1+i0)+\omega-\omega_p^{(+)}}\overline{\Psi}_p^{(+)}(\mathbf{r}')\\
&\times\hat{\mathbf{A}}(\mathbf{r}')\Psi_l(\mathbf{r}')\overline{\Psi}_l(\mathbf{r}'')\gamma_\mu''\Psi_m^{(+)}(\mathbf{r}'')\overline{\Psi}_r^{(+)}(\mathbf{r}''')\gamma_\mu'''\Psi_n^{(+)}\\
&\times(\mathbf{r}''')+\frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\sum_{l_\pm}\frac{\exp(i|\omega_n^{(+)}-\omega_r^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)-\omega-\omega_m^{(+)}}\\
&\times\overline{\Psi}_p^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_l(\mathbf{r}')\overline{\Psi}_l(\mathbf{r}'')\hat{\mathbf{A}}(\mathbf{r}'')\Psi_m^{(+)}(\mathbf{r}'')\overline{\Psi}_r^{(+)}(\mathbf{r}''')\\
&\times\gamma_\mu'''\Psi_n^{(+)}(\mathbf{r}''')+\frac{1}{|\mathbf{r}'-\mathbf{r}'''|}\\
&\times\sum_{l_\pm}\frac{\exp(i|\omega_p^{(+)}-\omega_m^{(+)}||\mathbf{r}'-\mathbf{r}'''|)}{\omega_l(1-i0)+\omega-\omega_r^{(+)}}\overline{\Psi}_p^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_m^{(+)}\\
&\times(\mathbf{r}')\overline{\Psi}_r^{(+)}(\mathbf{r}'')\hat{\mathbf{A}}(\mathbf{r}'')\Psi_l(\mathbf{r}'')\overline{\Psi}_l(\mathbf{r}''')\gamma_\mu'''\Psi_n^{(+)}(\mathbf{r}''')\\
&+\frac{1}{|\mathbf{r}'-\mathbf{r}''|}\sum_{l_\pm}\frac{\exp(i|\omega_p^{(+)}-\omega_m^{(+)}||\mathbf{r}'-\mathbf{r}''|)}{\omega_l(1-i0)+\omega-\omega_r^{(+)}}\\
&\times\overline{\Psi}_p^{(+)}(\mathbf{r}')\gamma_\mu'\Psi_m^{(+)}(\mathbf{r}')\overline{\Psi}_r^{(+)}\\
&\times(\mathbf{r}'')\gamma_\mu''\Psi_l(\mathbf{r}'')\overline{\Psi}_l(\mathbf{r}''')\hat{A}(\mathbf{r}''')\Psi_n^{(+)}(\mathbf{r}''')\Bigg\}\Bigg),
\end{aligned}
\tag{2.1}
$$

where $\omega_n^{(+)}$ and $\omega_m^{(+)}$ are the frequencies of the initial states of the electrons, $\omega_p^{(+)}$ and $\omega_r^{(+)}$ are the frequencies of the final electron states, $\Psi^{(+)}_{m(n)}$ are the solutions of the Dirac equation for a positive-frequency electron, $\Psi^{(+)}_{p(r)}=\Psi^{(+)*}_{p(r)}\gamma_4$, $\Psi^{(+)*}_{p(r)}$ is the conjugate wave function, $\gamma_4=\boldsymbol{\beta}$, $\gamma_j=-i\boldsymbol{\beta}\boldsymbol{\alpha}_j$ $(j=1,2,3)$,

$$
\boldsymbol{\alpha}=\begin{pmatrix}0 & \boldsymbol{\sigma}\\ \boldsymbol{\sigma} & 0\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},
$$

and $\boldsymbol{\sigma}$ are the Pauli matrices. The primes on the radius vectors $\mathbf{r}$ and the γ matrices correspond to different wave functions of the interacting particles, and the γ matrices with different numbers of primes commute. Summation in (2.1) is over all intermediate states $l_\pm$ with positive and negative frequencies. Electron states with negative frequencies are interpreted as positron states, and we do not introduce the positron wave function, which contains the charge conjugation transformation. The reason is that a linear combination of states with opposite signs of charge cannot be a general solution of the Dirac equation.[13] In this paper we use the solution

$$
\Psi=\sum_r a_r\psi_r^{(+)}+\sum_r b_r^\dagger\psi_r^{(-)},
$$

$$
\overline{\Psi}=\sum_r a_r^\dagger\overline{\psi}_r^{(+)}+\sum_r b_r\overline{\psi}_r^{(-)},
\tag{2.1a}
$$

where $a_r, a_r^\dagger$, $b_r, b_r^\dagger$ are operators of second quantization of the electron–positron field. The passage from $S^{(3)}_{i\to f}$ to the matrix of the effective interaction energy (2.1) is done using the relation

$$
S^{(3)}_{i\to f}=-2\pi i U^{(3)}_{i\to f}\delta(\omega_r^{(+)}-\omega_m^{(+)}-\omega+\omega_p^{(+)}-\omega_n^{(+)}),
\tag{2.2}
$$

where the sign of the frequency ω of the optical photon indicates that in the vector potential operator $\hat{\mathbf{A}}=\Sigma\gamma_\mu\mathbf{A}_\mu$ we have specified the negative-frequency part, which is proportional to the annihilation operator for a photon of the given mode.

2.1 Exchange of virtual photons

Let us start with the first term on the right-hand side of Eq. (2.1), in which we identify the factor responsible for the exchange of virtual photons between the electrons. We write the expression for the distance between the electrons as follows:

$$|\mathbf{r}''-\mathbf{r}'''|\approx a\left(1+\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a^2}+\frac{M_1}{a}\right), \tag{2.3}$$

here $a=|\mathbf{a}''-\mathbf{a}'''|$ is the distance between the atoms, $\Delta\boldsymbol{\xi}=\boldsymbol{\xi}''-\boldsymbol{\xi}'''$, where $\boldsymbol{\xi}''$ and $\boldsymbol{\xi}'''$ are the displacements of electrons in the atoms with respect to their nuclei, and $M_1=M_1(a,\Delta\boldsymbol{\xi})$ are corrections terms incorporating higher powers of the ratio $\Delta\xi/a$.

We introduce a system of units in which $c\neq 1$ and assume that

$$\frac{1}{c}|\omega_n^{(+)}-\omega_p^{(+)}|\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a}\ll 1. \tag{2.4}$$

Here the distance between the nuclei may vary within broad limits: $|\Delta\xi|\leqslant a\leqslant\infty$. We then find that

$$\begin{aligned}\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}&\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|\,|\mathbf{r}''-\mathbf{r}'''|\right)\\&=\frac{1}{a}\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\left[1+\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a^2}+\frac{M_1}{a}\right]^{-1}\\&\times\left\{1+\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|\left[\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a}+M_1\right]\right.\\&\left.-\frac{1}{2c^2}(\omega_n^{(+)}-\omega_p^{(+)})^2\left[\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a}+\frac{M_1}{a}\right]^2\right\}.\end{aligned} \tag{2.3′}$$

We remove the frequencies from (2.3′) by using the equation

$$H'''\psi_n^{(+)}(\mathbf{r}''')=\omega_n^{(+)}\psi_n^{(+)}(\mathbf{r}''').$$

Then we find that the following transformation holds:

$$\begin{aligned}\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}&\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|\,|\mathbf{r}''-\mathbf{r}'''|\right)\\&\to\frac{1}{a}\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\\&\times\left\{\left[1+\frac{\mathbf{a}\cdot\Delta\boldsymbol{\xi}}{a^2}+\frac{M_1}{a}\right]^{-1}+\frac{i}{c}[f_1,H''']\right.\\&\left.+\frac{1}{2c^2}[H'',[H''',f_2]]\frac{\omega_n^{(+)}-\omega_p^{(+)}}{\omega_l-\omega_m^{(+)}}\right\},\end{aligned} \tag{2.5}$$

where we have introduced the notation

$$\begin{aligned}f_1&\equiv a-\frac{a^2}{|\mathbf{r}''-\mathbf{r}'''|},\\f_2&\equiv a|\mathbf{r}''-\mathbf{r}'''|+\frac{a^3}{|\mathbf{r}''-\mathbf{r}'''|}-2a^2.\end{aligned} \tag{2.6}$$

Equation (2.3′) shows that, in addition to a series expansion in powers of $1/c$, we have a series expansion in powers of the small parameter $\Delta\xi/a$. Here in the $1/c$-expansion we restrict ourselves to terms quadratic in $1/c$, while in the $\Delta\xi/a$-expansion there are no limits, since the function M_1 contains all higher-order corrections. This can easily be verified by substituting the expansion (2.3) in (2.6) and comparing (2.5) with (2.3′). Thus, in our study we allow for the interaction of two atomic electrons of arbitrary multipole order.

The Hamiltonian operators for separate atoms with fixed nuclei have the form

$$\begin{aligned}H''&=c\boldsymbol{\alpha}''\cdot\mathbf{p}''+\gamma_4''mc^2-\frac{Z_1e^2}{\xi''},\\H'''&=c\boldsymbol{\alpha}'''\cdot\mathbf{p}'''+\gamma_4'''mc^2-\frac{Z_2e^2}{\xi'''},\end{aligned} \tag{2.7}$$

where $\mathbf{p}''$ and $\mathbf{p}'''$ are the electron momentum operators, and Z_1 and Z_2 are the charges of the point nuclei. Obviously, other terms can be introduced into the operators (2.7), e.g., terms determined by the finite size of the nuclei (or core) of an atom and by nuclear spin. We calculate the commutators in (2.5) assuming that the nuclei are immobile. Then with $\hbar\neq 1$ we have

$$\begin{aligned}[f_1,H''']&=-i\hbar ca^2\frac{\boldsymbol{\alpha}''\cdot\mathbf{n}}{|\mathbf{r}''-\mathbf{r}'''|^2},\\[H'',[H''',f_2]]&=ac^2[\boldsymbol{\alpha}''\cdot\mathbf{p}'',[\boldsymbol{\alpha}'''\cdot\mathbf{p}''',|\mathbf{r}''-\mathbf{r}'''|]]\\&\quad+a^3c^2\left[\boldsymbol{\alpha}''\cdot\mathbf{p}'',\left[\boldsymbol{\alpha}'''\cdot\mathbf{p}''',\frac{1}{|\mathbf{r}''-\mathbf{r}'''|}\right]\right].\end{aligned} \tag{2.8}$$

Thus, the operator describing the exchange of virtual photons in the matrix (2.1) takes the form

$$\begin{aligned}B_{1l}(\mathbf{r}'',\mathbf{r}''')&=e^2\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\\&\times\left\{\frac{1-\boldsymbol{\alpha}''\cdot\boldsymbol{\alpha}'''}{|\mathbf{r}''-\mathbf{r}'''|}+a\frac{\boldsymbol{\alpha}''\cdot\mathbf{n}}{|\mathbf{r}''-\mathbf{r}'''|^2}\right.\\&+\frac{1}{2}R_{1l}\left(\frac{\boldsymbol{\alpha}''\cdot\boldsymbol{\alpha}'''-(\boldsymbol{\alpha}''\cdot\mathbf{n})(\boldsymbol{\alpha}'''\cdot\mathbf{n})}{|\mathbf{r}''-\mathbf{r}'''|}\right.\\&\left.\left.-a^2\frac{\boldsymbol{\alpha}''\cdot\boldsymbol{\alpha}'''-3(\boldsymbol{\alpha}''\cdot\mathbf{n})(\boldsymbol{\alpha}'''\cdot\mathbf{n})}{|\mathbf{r}''-\mathbf{r}'''|^3}\right)\right\},\end{aligned} \tag{2.9}$$

where $R_{1l}=(\omega_n^{(+)}-\omega_p^{(+)})/(\omega_l-\omega_m^{(+)})$, and $\mathbf{n}=(\mathbf{r}''-\mathbf{r}''')/|\mathbf{r}''-\mathbf{r}'''|$. In the particular case of a resonant exchange of photons we have $R_{1l}=1$, and the operator (2.8) becomes the respective operator of Ref. 4. In the limit $a\to 0$ and $R_{1l}=1$, the operator (2.9) coincides with the Breit operator.[13]

Let us examine the physical meaning of the expansion (2.5). In deriving the Lagrangian function to within terms of order $(v/c)^2$ for a system of charges with a continuous spectrum,[2] we use a universal time scale R/c, where R is the distance between the charges. This is the time the interaction takes to propagate from charge to charge, and an expansion of the retarded potentials in powers of R/c is possible if we assume that the distribution of the charges does not change appreciably during time R/c. Obviously, this is an extremely

stringent condition if one is examining extended systems. In deriving the Breit operator[13] for a system of two electrons with a discrete energy spectrum, one can use $\omega_0 r/c \ll 1$ as the expansion parameter, where ω_0 is the characteristic frequency in the spectrum of the interacting electrons, and r is the distance between electrons. Thus, in addition to the time $T_e = r/c$ it takes the interaction to propagate, we have the characteristic time $T_a = 2\pi/\omega_0$. Here $2\pi T_e \ll T_a$, i.e., the time it takes the electron density in the system of two interacting (and moving) electrons to change appreciably is roughly the time it takes the interaction to propagate from charge to charge. Obviously, this is true only if the electrons are not too far apart, i.e., for interatomic distance in heliumlike atoms. The expansion (2.5) is valid for two atomic electrons located at arbitrary distances from one another, since we have chosen (2.4) as the expansion parameter. Here the time it takes the interaction to propagate, $T_e = a/c$, is much longer than the characteristic interatomic time scale T_a. This leads to additional retardation in the electron–electron interaction, with the retardation described by the additional terms in the operator (2.9).

As is known, the characteristic frequencies of atomic electrons range from several megahertz to 10^9 MHz, if one allows for fine and hyperfine structure of atomic levels and for optical transitions. This poses the problem of selecting the characteristic interatomic time scale T_a. Obviously, to fully account for the retardation effect in the electron–electron interaction we must take the highest frequencies in the spectrum of the interacting electrons, i.e., optical frequencies, as the characteristic frequencies. Below we show that the operator (2.9) corresponds to different types of electron–electron interaction, and the fact that there is a universal time scale T_a explains why, for instance, the spin–spin coupling of electrons contains additional retardation terms (additional with respect to the terms obtained in Ref. 13), although the spin transitions correspond to the microwave range.

Let us now consider the other terms in the matrix (2.1), which correspond to the other Feynman diagrams in Fig. 1. Using the same procedure of replacing frequencies with operators, we arrive at the operators B_{sl} ($s = 2,3,\ldots,8$), which are similar to the operator B_{1l}. Here the operators B_{sl} contain various coefficients determining the difference in initial, final, and intermediate frequencies and different retardation factors in accordance with the positions of the wave functions in (2.1).

Third-order effects for the Feynman diagrams in Fig. 1 take place provided that the conservation law (2.2) is obeyed. For emission of a real photon the sign of the frequency ω in (2.2) must be changed. Thus, for the electron–electron interaction energies we have

$$E_r^{(+)} - E_m^{(+)} + E_p^{(+)} - E_n^{(+)} \pm \hbar\omega = 0. \tag{2.10}$$

Several patterns of quantum transitions follow from this conservation law. To estimate them, we take two states, say p and n, in the spectrum of the interacting atoms and assume that the initial states of these atoms, n and m, coincide.

A. The exchange of virtual photons with a frequency $\omega'' = \omega_n^{(+)} - \omega_p^{(+)}$ results in a transition of one atom (e.g., the first) to an intermediate state ω_l and then back to the initial state $E_m^{(+)}$, i.e., $E_r^{(+)} = E_m^{(+)}$. The second atom changes its quantum state, $E_n^{(+)} \to E_p^{(+)}$, so that a photon is emitted or absorbed in the two-atom system. The retardation factor for this type of interaction is $\exp\{(i/c)|\omega_n^{(+)} - \omega_p^{(+)}|a\}$. The point where a real photon is emitted or absorbed may not coincide with the point occupied by the atom undergoing the quantum transition $E_n^{(+)} \to E_p^{(+)}$. This pattern of quantum transitions corresponds to diagrams 1, 2, 7, and 8 in Fig. 1.

B. For diagrams 1–8 there can be transitions in which a photon of double energy is emitted or absorbed. The two atoms change their quantum states, i.e., in this case we have $E_r^{(+)} \neq E_m^{(+)}$ and $E_p^{(+)} \neq E_n^{(+)}$.

C. For $E_r^{(+)} = E_m^{(+)}$, when only one atom changes its quantum state and a real photon with the energy $\hbar\omega = |E_p^{(+)} - E_n^{(+)}|$ is either emitted or absorbed, interaction with a retardation factor equal to unity is possible. Diagrams 3–6 possess such properties.

The nature of the quantum transitions in patterns A, B, and C is determined by the properties of the operators B_{sl} and $\hat{\mathbf{A}}$ and of the wave functions and energy values of the atomic electrons.

3. CONVERSION TO TWO-COMPONENT WAVE FUNCTIONS

3.1. Exchange of virtual photons

The conversion to two-component wave functions $\Phi_n^{(+)}$ is done by the following transformations:[13]

$$\Psi_n^{(+)} = \begin{pmatrix} \varphi_n^{(+)} \\ \dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{2mc}\varphi_n^{(+)} \end{pmatrix}, \quad \varphi_n^{(+)} = \left(1 - \frac{\mathbf{p}^2}{8m^2c^2}\right)\Phi_n^{(+)}. \tag{3.1}$$

Let us use these transformations to calculate the matrix element

$$\langle \Psi_l^*(\mathbf{r}'')\Psi_p^{(+)*}(\mathbf{r}''')|B_{1l}|\Psi_m^{(+)}(\mathbf{r}'')\Psi_n^{(+)}(\mathbf{r}''')\rangle, \tag{3.2}$$

in the first term on the right-hand side of Eq. (2.1).

We consider the matrix element of the operator $1/|\mathbf{r}'' - \mathbf{r}'''|$ in (2.9). For the intermediate states in (3.2) we select positive-energy states. Then, applying (3.1), we arrive at the expression

$$e^2 \exp\left(\frac{i}{c}|\omega_n^{(+)} - \omega_p^{(+)}|a\right)$$

$$\times \int \left\{ \Phi_l^{(+)*}(\mathbf{r}'')\Phi_p^{(+)*}(\mathbf{r}''')\Phi_m^{(+)}(\mathbf{r}'')\Phi_n^{(+)}(\mathbf{r}''') \right.$$

$$- \frac{1}{8m^2c^2}\Phi_l^{(+)*}(\mathbf{r}'')\Phi_p^{(+)*}(\mathbf{r}''')\Phi_m^{(+)}(\mathbf{r}'')$$

$$\times [\mathbf{p}'''^2\Phi_n^{(+)}(\mathbf{r}''')] - \frac{1}{8m^2c^2}\Phi_l^{(+)*}(\mathbf{r}'')\Phi_p^{(+)*}(\mathbf{r}''')\Phi_n^{(+)}$$

$$\times (\mathbf{r}'')[\mathbf{p}''^2\Phi_m^{(+)}(\mathbf{r}'')] + \frac{1}{8m^2c^2}\Phi_l^{(+)*}(\mathbf{r}'')[\mathbf{p}'''^2\Phi_p^{(+)}$$

$$\times(\mathbf{r}''')]^*\Phi_m^{(+)}(\mathbf{r}'')\Phi_n^{(+)}(\mathbf{r}''')+\frac{1}{8m^2c^2}[\mathbf{p}''^2\Phi_n^{(+)}(\mathbf{r}'')]^*$$

$$\times\Phi_p^{(+)}(\mathbf{r}''')\Phi_m^{(+)}(\mathbf{r}'')\Phi_n^{(+)}(\mathbf{r}''')\Bigg\}\frac{d\mathbf{r}''\,d\mathbf{r}'''}{|\mathbf{r}''-\mathbf{r}'''|}, \tag{3.3}$$

where we have discarded terms of orders higher than the second in $1/c$. We transform (3.3) so that it takes the form

$$\int \Phi_l^{(+)*}(\mathbf{r}'')\Phi_p^{(+)*}(\mathbf{r}''')V_1^{(+)}\Phi_m^{(+)}(\mathbf{r}'')\Phi_n^{(+)}(\mathbf{r}''')$$

$$=(V_1^{(+)})_{lp,mn}, \tag{3.4}$$

and find the explicit form of the operator $V_1^{(+)}$. Transforming (3.3) into (3.4) requires integrating by parts. We place the origin of the coordinate system at the point a'' and assume that the radius vectors $\mathbf{r}'$ and $\mathbf{r}''$ refer to the first atom and the radius vector $\mathbf{r}'''$, to the second. We also allow for the fact that the atomic wave functions do not overlap and vanish at infinity. This means that, in contrast to the case discussed in Ref. 13, the powers of the $1/|\mathbf{r}''-\mathbf{r}'''|$ do not become infinite when $\mathbf{r}''$ and $\mathbf{r}'''$ vary within the regions occupied by the interacting atoms. Taking all these remarks into account, we obtain the operator

$$V_1^{(+)}=e^2\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\frac{1}{r},\quad r=|\mathbf{r}''-\mathbf{r}'''|, \tag{3.5}$$

which is analogous to the Coulomb interaction of electrons located at arbitrary distances from one another. As $a\to 1$, this term becomes the ordinary Coulomb interaction of electrons.

Let us now examine the matrix element of the other terms in the operator B_{1l} by employing two-component wave functions $\Phi_n^{(+)}$. Substituting the wave functions (3.1) in the matrix (2.1) and multiplying the matrices in the integrand, we find that it is enough to replace φ with Φ in all the terms, since they already contain the factor $1/c^2$. We perform transformations similar to those that were used in deriving the operator (3.5). Here we can separate operators containing different powers of $1/r$.

The operator proportional to $1/r^3$ has the form

$$V_{2l}^{(+)}=\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\frac{e^2\hbar^2}{4m^2c^2}$$

$$\times\Bigg\{\frac{1}{r^3}[\boldsymbol{\sigma}\cdot\boldsymbol{\sigma}'''-3(\boldsymbol{\sigma}''\cdot\mathbf{n})(\boldsymbol{\sigma}'''\cdot\mathbf{n})]$$

$$+R_{1l}\frac{a^2}{r^5}[15(\boldsymbol{\sigma}''\cdot\mathbf{n})(\boldsymbol{\sigma}'''\cdot\mathbf{n})-9\boldsymbol{\sigma}''\cdot\boldsymbol{\sigma}''']\Bigg\}. \tag{3.6}$$

As $a\to 0$, (3.6) becomes the operator of the spin–orbit coupling of two electrons.[13] According to (2.4), for electrons located at arbitrary distances from one another there is addition retardation, determined by the retardation factor $\exp\{(i/c)|\omega_n^{(+)}-\omega_p^{(+)}|a\}$ and the additional terms in (3.6). In the particular case of resonant electron–electron interaction without photon emission or absorption we have $R_{1l}=1$.

After the necessary transformations have been done, the operator proportional to $1/r$ becomes

$$V_{3l}^{(+)}=e^2\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\frac{1}{m^2c^2}$$

$$\times\Bigg\{\left(\frac{R_{1l}}{2}-1\right)\frac{1}{\mathbf{r}}\mathbf{p}''\cdot\mathbf{p}'''-\frac{R_{1l}}{2}\frac{1}{r}\,\mathbf{n}\cdot(\mathbf{n}\cdot\mathbf{p}'')\mathbf{p}'''$$

$$-\frac{R_{1l}}{2}\frac{a^2}{r^3}[\mathbf{p}''\cdot\mathbf{p}'''-3\mathbf{n}\cdot(\mathbf{n}\cdot\mathbf{p}'')\mathbf{p}''']\Bigg\}$$

$$+e^2\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\frac{a}{mc}\frac{\mathbf{n}\cdot\mathbf{p}''}{r^2}. \tag{3.7}$$

In the limit $a\to 0$ and $R_{1l}=1$, this operator becomes the retarded interaction operator for the two electrons in a heliumlike atom.[13] Hence we call (3.7) the operator of the retarded interaction of two electrons located at arbitrary distances from one another. At $R_{1l}=1$ (resonant electron–electron interaction), the operator (3.7) can be derived from the classical Hamiltonian function for a system of atoms by replacing the electron momenta with the corresponding momentum operators.[4] The terms proportional to $1/r^2$ in the matrix elements of B_{1l} are

$$V_{4l}^{(+)}=\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\frac{e^2\hbar}{4m^2c^2}$$

$$\times\Bigg\{\left(\frac{R_{1l}}{2}-1\right)\frac{1}{r^2}[2\boldsymbol{\sigma}''\cdot(\mathbf{n}\times\mathbf{p}''')-2\boldsymbol{\sigma}'''\cdot(\mathbf{n}\times\mathbf{p}'')]$$

$$+\frac{R_{1l}}{2}\frac{1}{r^2}[2\boldsymbol{\sigma}''\cdot(\mathbf{n}\times\mathbf{p}'')+2\boldsymbol{\sigma}'''\cdot(\mathbf{n}\times\mathbf{p}''')$$

$$-2\boldsymbol{\sigma}'''\cdot(\mathbf{n}\times\mathbf{p}'')-2\boldsymbol{\sigma}'''\cdot(\mathbf{n}\times\mathbf{p}''')]$$

$$+3R_{1l}\frac{a^2}{r^4}\boldsymbol{\sigma}''\cdot(\mathbf{n}\times\mathbf{p}'')\Bigg\}. \tag{3.8}$$

In the limit $a\to 0$ and $R_{1l=1}$, this operator becomes the corresponding operator of spin–orbit coupling of two electrons in the Breit operator.[13] Hence we call (3.8) the operator of spin–orbit retardation of two atomic electrons located at arbitrary distances from one another.

3.2. Electric dipole–dipole interaction of atoms located at arbitrary distances from one another

We turn to the electron–electron interaction via the field of virtual photons and allow only for the orbital degrees of freedom. The corresponding operator is $V_{1l}^{(+)}+V_{3l}^{(+)}$. As noted earlier, in deriving the operator of the interaction of two atoms we allow for transitions of arbitrary multipole order in the atomic spectrum. Taking into account only the terms in the expansions of $1/r$ and $1/r^2$ that are linear in the displacements $\boldsymbol{\xi}''$ and $\boldsymbol{\xi}'''$, we arrive at the operator

$$V_{1l}^{(+)}+V_{3l}^{(+)}=\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\times\left\{\frac{\mathbf{d}''\cdot\mathbf{d}'''-3(\mathbf{n}\cdot\mathbf{d}'')(\mathbf{n}\cdot\mathbf{d}''')}{a^3}+\frac{e}{mc}\frac{\mathbf{d}'''\cdot\mathbf{p}''-3(\mathbf{n}\cdot\mathbf{d}''')(\mathbf{n}\cdot\mathbf{p}'')}{a^2}+\frac{e^2}{m^2c^2}\times\left[\left(\frac{R_{1l}}{2}-1\right)\frac{\mathbf{p}''\cdot\mathbf{p}'''}{a}-\frac{R_{1l}}{2}\frac{(\mathbf{n}\cdot\mathbf{p}'')(\mathbf{n}\cdot\mathbf{p}''')}{a}-\frac{R_{1l}}{2}\frac{\mathbf{p}''\cdot\mathbf{p}'''-3(\mathbf{n}\cdot\mathbf{p}'')(\mathbf{n}\cdot\mathbf{p}''')}{a}\right]\right\}, \tag{3.9}$$

where $\mathbf{n}=\mathbf{a}/a$, and $\mathbf{d}''=e\boldsymbol{\xi}''$ and $\mathbf{d}'''=e\boldsymbol{\xi}'''$ are the operators of the electric dipole moments of individual atoms. Here (3.9) is the operator of the electric dipole–dipole interaction of two neutral atoms located at arbitrary distances from one another in the transition of one of the atoms to an intermediate state, where the interaction (3.9) is the integral part of the process of emission or absorption of a real photon. In the particular case of the interaction of two atoms without emission or absorption of a real photon, the (3.9) corresponds to a second-order QED effect, for which the following energy-conservation law holds:

$$E_r^{(+)}-E_m^{(+)}+E_p^{(+)}-E_n^{(+)}=0.$$

In this case we must set R_{1l} to unity in (3.9), with the result that (3.9) becomes the corresponding operator of Ref. 13.

4. THE ROLE OF AN EXTERNAL FIELD IN THE INTERACTION OF TWO ATOMIC ELECTRONS

We use the transformation (3.1) for the passage from the wave functions Ψ to the two-component wave functions Φ in the matrix elements of type

$$\langle\overline{\Psi}_r^{(+)}(\mathbf{r}')|e\hat{\mathbf{A}}'(\mathbf{r}')|\Psi_l(\mathbf{r}')\rangle, \tag{4.1}$$

which enter into the matrix (2.1). The matrix elements (4.1) of the operator of the vector potential of an external field determine the interaction of two atomic electrons with the field of real photons. We start with the matrix elements (4.1) for the transitions of the atomic electrons through intermediate positive-energy states. After the necessary transformations have been carried out, we identify the following transition operator in the matrix element (4.1) for the transition of the atom from state $\Phi_l^{(+)}(\mathbf{r}')$ to state $\Phi_r^{(+)}(\mathbf{r}')$:

$$R_1^+=eA_4'(\mathbf{r}'). \tag{4.2}$$

The terms proportional to $1/c$ are

$$R_2^+=-\frac{ie}{2mc}(\mathbf{p}'\cdot\mathbf{A}')-\frac{ie}{2mc}(\mathbf{A}'\cdot\mathbf{p}')-\frac{i\hbar e}{2mc}(\boldsymbol{\sigma}'\cdot\mathbf{H}'), \tag{4.3}$$

where $\mathbf{A}'(\mathbf{r}')$ is the operator of the vector potential of the external field at the point occupied by an electron and specified by the radius vector $\mathbf{r}'$, and $\mathbf{H}'(\mathbf{r}')$ is the corresponding magnetic field operator, $([\nabla'\times\mathbf{A}']=\mathbf{H}')$.

The operator R_3^+ containing $1/c$ squared has the form

$$R_3^+=-\frac{e\hbar^2}{8m^2c^2}\Delta'A_4'-\frac{i\hbar e}{4m^2c^2}\nabla'A_4'\mathbf{p}'+\frac{e}{4m^2c^2}(\boldsymbol{\sigma}'\cdot\mathbf{p}')A_4'(\boldsymbol{\sigma}'\cdot\mathbf{p}'). \tag{4.4}$$

The terms proportional to $1/c^3$ in the transition operator have the form

$$R_4^+=\frac{e}{16m^3c^3}\left\{i(\boldsymbol{\sigma}'\cdot\mathbf{A}')(\boldsymbol{\sigma}'\cdot\mathbf{p}')\mathbf{p}'^2-i\mathbf{p}'^2(\boldsymbol{\sigma}'\cdot\mathbf{A}')(\boldsymbol{\sigma}'\cdot\mathbf{p}')+\hbar\sum_\alpha\sigma_\alpha'\left(\boldsymbol{\sigma}'\cdot\frac{\partial\mathbf{A}'}{\partial x_\alpha'}\right)\mathbf{p}'^2+i\boldsymbol{\sigma}'\cdot(\boldsymbol{\sigma}'\cdot\mathbf{A}')\mathbf{p}'^3-\hbar\sum_\alpha\sigma_\alpha'\times\mathbf{p}'^2\left(\boldsymbol{\sigma}'\cdot\frac{\partial\mathbf{A}'}{\partial x_\alpha'}\right)-i\boldsymbol{\sigma}'\cdot[\mathbf{p}'^2(\boldsymbol{\sigma}'\cdot\mathbf{A}')\mathbf{p}']\right\}, \tag{4.5}$$

where $\alpha=x,y,z$.

Each of the above operators can initiate a quantum transition of an atomic electron from an intermediate state $\Phi_l^{(+)}$ to the final state $\Phi_r^{(+)}$ if the atomic electron reached the intermediate state because of an exchange of virtual photons with an electron of another atom. Here the quantized external light field has a potential A_4' equal to zero. For a constant external field the term (4.2) is finite, and in this case the interaction of electrons belonging to two different atoms via the field of virtual photons takes place in accordance with the following conservation law:

$$E_r^{(+)}-E_m^{(+)}+E_p^{(+)}-E_n^{(+)}=0.$$

4.1. Allowing for intermediate state with negative energies

The matrix (2.1) of the effective interaction energy contains summation over the intermediate states of the interacting electrons with negative energies. This means that a part of the energy of the interaction of two electrons is due to the effect of positron states in the spectrum of the interacting electrons. The effect enters indirectly via the intermediate virtual states of the electrons. Note that the initial and final states are electron states with positive energies.

The intermediate positron states can be taken into account as we go over to two-component wave functions in (2.1) with wave functions of the form

$$\Psi_l^{(-)}=\begin{pmatrix}-\dfrac{\boldsymbol{\sigma}\cdot\mathbf{p}}{2mc}\chi_l^{(-)}\\ \chi_l^{(-)}\end{pmatrix},\quad \chi_l^{(-)}=\left(1-\frac{\mathbf{p}^2}{8m^2c^2}\right)\Phi_l^{(-)}. \tag{4.6}$$

Here, as in Ref. 17, we do not need to go over to the positron wave function, which contains the charge conjugation transformation.[13]

However, we will choose another method for allowing for intermediate positron states, a method based on the following obvious conditions. First, the photon energy is much lower than the energy of the electron field, and the electron energy differs little from the electron rest energy. Hence in

(2.1) we put

$$\hbar[\omega_l(1-i0)+\omega-\omega_r^{(+)}]\approx -2mc^2, \tag{4.7}$$

etc. Second, we introduce the projection operators

$$\Lambda'_-=\frac{mc^2-H'}{2mc^2}, \quad \Lambda''_-=\frac{mc^2-H''}{2mc^2} \tag{4.8}$$

with the properties

$$\Lambda'_-\Psi_l(\mathbf{r}')=\Psi_l^{(-)}(\mathbf{r}'), \quad \Lambda''_-\Psi_l(\mathbf{r}'')=\Psi_l^{(-)}(\mathbf{r}''). \tag{4.9}$$

Then, applying the transformation of Sec. 2, we separate the following operator in the first and second term on the right-hand side of (2.1):

$$P_{1l}^+=\frac{e}{2mc^2}(\gamma'_4\gamma'_\delta A'_\delta\Lambda'_-B_{1l}+B_{2l}\Lambda''_-\gamma''_4\gamma''_\delta A''_\delta), \tag{4.10}$$

where $\delta=1,2,3,4$, and a'_g is the four-operator of the vector potential of the external field at the point occupied by the electron with the radius vector $\mathbf{r}'$. The other terms of the matrix (2.1) can also be combined into pairs.

Now, using the wave functions (3.1) for the initial and final electron states, we transform the operator (4.10). Both R_{1l} and R_{2l} already have the factor $1/2mc^2$, so that the operators B_{1l} and B_{2l} have a truncated form. We multiply the operators in (4.10) in matrix form and the two-component wave functions of the interacting electrons and keep only the terms proportional to $1/c$. Integrating by parts where necessary, we arrive at the following formula for the operator in (4.10):

$$\begin{aligned}
P_1^+=&\frac{e^3}{2mc^2}\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\left\{-\frac{i\hbar}{2mc}\frac{a}{r^2}(\boldsymbol{\sigma}''\cdot\mathbf{n})\right.\\
&\times\sum_\alpha\sigma'_\alpha\frac{\partial A'_4}{\partial x'_\alpha}-\frac{1}{2mc}(\boldsymbol{\sigma}'\cdot\mathbf{p}')(\boldsymbol{\sigma}''\cdot\mathbf{n})\frac{a}{r^2}A'_4+\frac{i\hbar}{4m^2c^2}\\
&\times(\boldsymbol{\sigma}'\cdot\mathbf{A}')(\boldsymbol{\sigma}''\cdot\mathbf{p}'')\frac{1}{r}(\boldsymbol{\sigma}'\cdot\mathbf{A}')\boldsymbol{\sigma}''\cdot\sum_\alpha\sigma'''_\alpha\boldsymbol{\sigma}'''\\
&\times\left(\frac{1}{r}\frac{\partial}{\partial x'''_\alpha}+\frac{n_\alpha}{r^2}\right)-i(\boldsymbol{\sigma}'\cdot\mathbf{A}')(\boldsymbol{\sigma}''\cdot\mathbf{n})\frac{a}{r^2}+\frac{i}{2mc}\\
&\times(\boldsymbol{\sigma}'\cdot\mathbf{A}')(\boldsymbol{\sigma}''\cdot\boldsymbol{\sigma}''')(\boldsymbol{\sigma}'''\cdot\mathbf{p}''')\frac{1}{r}+\frac{i}{2mc}(\boldsymbol{\sigma}'\cdot\mathbf{A}')\\
&\times(\boldsymbol{\sigma}'\cdot\mathbf{p}')\frac{1}{r}+\frac{1}{2mc}(\boldsymbol{\sigma}'\cdot\mathbf{n})(\boldsymbol{\sigma}''\cdot\mathbf{p}'')\frac{a}{r^2}A''_4-\frac{1}{2mc}\frac{a}{r^2}\\
&\times(\boldsymbol{\sigma}'\cdot\mathbf{n})(\boldsymbol{\sigma}''\cdot\mathbf{p}'')A''_4+\frac{\hbar}{2mc}\sum_\alpha(\boldsymbol{\sigma}''\cdot\mathbf{A}'')\sigma'_\alpha\frac{n_\alpha}{r^2}\\
&+\frac{i}{2mc}\frac{1}{r}(\boldsymbol{\sigma}''\cdot\mathbf{p}'')(\boldsymbol{\sigma}''\cdot\mathbf{A}'')+\frac{i}{2mc}\frac{1}{r}\boldsymbol{\sigma}'\cdot(\boldsymbol{\sigma}''\cdot\mathbf{A}'')\\
&\left.\times\boldsymbol{\sigma}'''(\boldsymbol{\sigma}'''\cdot\mathbf{p}''')-i\frac{a}{r^2}(\boldsymbol{\sigma}'\cdot\mathbf{n})(\boldsymbol{\sigma}''\cdot\mathbf{A}'')\right\},
\end{aligned} \tag{4.11}$$

where the operators $\boldsymbol{\sigma}'$ and $\boldsymbol{\sigma}''$ act on the spin wave functions of the electron of the first atom and the operator $\boldsymbol{\sigma}'''$, on the spin wave functions of the electron of the second atom.

Next we use the identities

$$\begin{aligned}
&\boldsymbol{\sigma}'''(\boldsymbol{\sigma}'''\cdot\mathbf{p}''')=-(\boldsymbol{\sigma}'''\cdot\mathbf{p}''')\boldsymbol{\sigma}'''+\mathbf{p}''',\\
&\boldsymbol{\sigma}'''\cdot\mathbf{p}'''\frac{1}{r}=\frac{1}{r}\boldsymbol{\sigma}'''\cdot\mathbf{p}'''-i\hbar\frac{\mathbf{n}\cdot\boldsymbol{\sigma}'''}{r^2},\\
&(\boldsymbol{\sigma}''\cdot\mathbf{A}'')(\boldsymbol{\sigma}''\cdot\boldsymbol{\sigma}''')(\boldsymbol{\sigma}'''\cdot\mathbf{p}''')=\mathbf{A}''\cdot\mathbf{p}'''+i\boldsymbol{\sigma}'''\cdot(\mathbf{A}''\times\mathbf{p}''')\\
&\qquad-i\,\mathbf{p}'''(\mathbf{A}''\times\boldsymbol{\sigma}'')\\
&\qquad+\sum_{\alpha\neq\beta}\sigma''_\alpha\sigma'''_\alpha A''_\beta p'''_\beta\\
&\qquad-\sum_{\alpha\neq\beta}\sigma''_\alpha p'''_\alpha A''_\beta\sigma'''_\beta.
\end{aligned} \tag{4.12}$$

Suppose that the interaction of the two atomic electrons takes place in a variable external field and that the vector potential $\mathbf{A}$ satisfies the Lorentz condition $\partial A_\mu/\partial x_\mu=0$ ($\mu=1,2,3,4$). In this case we can put[13]

$$A'_4=0, \quad \frac{\partial A'_4}{\partial x_\alpha}=0. \tag{4.13}$$

Then the operator (4.11) takes the form

$$\begin{aligned}
P_1^+=&\frac{e^3}{2mc^2}\exp\left(\frac{i}{c}|\omega_n^{(+)}-\omega_p^{(+)}|a\right)\\
&\times\left\{-2i\mathbf{A}''\cdot\mathbf{n}\frac{a}{r^2}+2\boldsymbol{\sigma}''\cdot(\mathbf{A}''\times\mathbf{n})\frac{a}{r^2}+\frac{\hbar}{mc}\mathbf{A}''\cdot\mathbf{n}\frac{1}{r^2}\right.\\
&+\frac{i\hbar}{mc}\boldsymbol{\sigma}''\cdot(\mathbf{A}''\times\mathbf{n})\frac{1}{r^2}+\frac{i}{2mc}\frac{1}{r}\mathbf{p}''\cdot\mathbf{A}''\\
&-\frac{1}{2mc}\frac{1}{r}\boldsymbol{\sigma}''\cdot(\mathbf{p}''\times\mathbf{A}'')+\frac{i}{mc}\frac{1}{r}\mathbf{A}''\cdot\mathbf{p}'''-\frac{i}{2mc}\frac{1}{r}\\
&\left.\times\left(\sum_{\alpha\neq\beta}\sigma''_\alpha p'''_\alpha A''_\beta p'''_\beta-\sum_{\alpha\neq\beta}\sigma''_\alpha p'''_\alpha A''_\beta\sigma'''_\beta\right)\right\}.
\end{aligned} \tag{4.14}$$

The other terms in the matrix (2.1) and the corresponding operators for the intermediate state with negative energies can be transformed similarly; we denote these operators by P_2^+, P_3^+, and P_4^+. The physical meaning of these operators is discussed below.

4.2. Intermediate states with positive energies

We will now discuss the problem of summing over the intermediate states with positive energies in the matrix of the effective interaction energy, i.e., over the electron states l_+. We apply the same transformations of matrix elements that were carried out in Secs. 3 and 4 that involve the approximate wave functions (3.1). Then the first two terms of the matrix (2.1), corresponding to the diagrams 1 and 2 in Fig. 1 with the same retardation factor $\exp\{(i/c)|\omega_n^{(+)}-\omega_p^{(+)}|a\}$, take the form

$$A^{(3)}_{i\to f}$$
$$=\sum_{s,s'}\sum_{l_+}\left\{\frac{\langle\Phi_r^{(+)}|R_s^+|\Phi_l^{(+)}\rangle\langle\Phi_l^{(+)}\Phi_p^{(+)}|V_{s'l1}^{(+)}|\Phi_m^{(+)}\Phi_n^{(+)}\rangle}{\hbar[\omega_l(1-i0)+\omega-\omega_r^{(+)}]}\right.$$
$$\left.+\frac{\langle\Phi_r^{(+)}\Phi_p^{(+)}|V_{s'l2}^{(+)}|\Phi_l^{(+)}\Phi_n^{(+)}\rangle\langle\Phi_l^{(+)}|R_s^+|\Phi_m^{(+)}\rangle}{\hbar[\omega_l(1-i0)-\omega-\omega_m^{(+)}]}\right\}, \quad (4.15)$$

where the operators $V_{s'l1}^{(+)}$ and $V_{s'l2}^{(+)}$ are derived from the corresponding operators B_{1l} and B_{2l}. In multiplying the different matrix elements in (4.15), we discard terms containing $1/c$ raised to a power higher than the third.

The other terms in (2.1) take a similar form. We denote these terms by $B^{(3)}_{i\to f}$, $C^{(3)}_{i\to f}$, and $D^{(3)}_{i\to f}$. Below we show that the various terms distinguished in (2.1) have different physical meaning.

5. POLARIZING FIELDS IN A SYSTEM OF HYDROGENLIKE ATOMS EMITTING AND ABSORBING PHOTONS

The operators (4.3), (4.4), and (4.5), which correspond to first-order effects, contain (as factors) atomic and field operators that give rise to the transitions between atoms and photon states at the same point under observation. The situation is different for the interaction of two atoms located at arbitrary distance from one another, with the operators (4.14) or the matrix elements (4.15). In the different terms of (4.14) we can identify an operator that ''acts'' at a certain point under observation and an operator that ''acts'' at the point occupied by the other atom, which generates the polarizing field. Here the polarizing field is a field of virtual photons rather than a field of real photons. Similarly, matrix elements of type (4.15) contain a dependence on the coordinates of the two atoms, one of which occupies the point under observation.

Let us examine the polarizing fields that form according to pattern C of Sec. 2, when as a result of exchange of virtual photons and emission or absorption of one real photon the quantum state of only one of the atoms changes. We assume that the point occupied by the first atom is specified by the radius vector $\mathbf{r}_1$ (coordinates $\mathbf{r}'$ and $\mathbf{r}''$) and that occupied by the other, by the radius vector $\mathbf{r}_2$ (coordinates $\mathbf{r}'''$). The initial state of the first atom is labeled m and has the energy $E_m^{(+)}$, while the initial state of the second atom is labeled n and has the energy $E_n^{(+)}$. As a result of an exchange of virtual photons the first atom arrives at an intermediate state $E_l^{(+)}$ or $E_l^{(-)}$ and then goes back to the initial state, i.e., $E_r^{(+)}=E_m^{(+)}$. One real photon is absorbed at point occupied by the first atom, while the second atom changes its quantum state and reaches the level $E_p^{(+)}>E_n^{(+)}$. This pattern of quantum transitions corresponds to the first term of the matrix (2.1) and the diagram 1 in Fig. 1. A similar situation occurs for the second term in (2.1) with the Feynman diagram 2 (Fig. 1), where the absorption of a photon takes place not at the point occupied by the atom undergoing the $E_n^{(+)}\to E_p^{(+)}$ transition but at the point occupied by the other atom, the one that forms the polarizing field. The other terms in (2.1) with the corresponding diagrams 3–8 do not participate in the formation of the polarizing field in this pattern. Indeed, when $r=m$, each of these terms contains zero matrix element of type $\langle\Psi_r^{(+)}|\gamma'_\mu|\Psi_m^{(+)}\rangle$.

We now use (4.14) and (4.15) to write down the vector potentials of the polarizing fields. To this end we first specify a certain type of quantum transitions between the states of the interacting atoms, e.g., orbital quantum transitions. Then, in accordance with (4.3), for first-order effects we arrive at the formula for the Hamiltonian operator of each atom:

$$H'''_1=-\frac{e}{mc}\mathbf{p}'''\cdot\mathbf{A}''', \quad (5.1)$$

where $\mathbf{A}'''$ is the operator of the vector potential of the external field at the point occupied by the second atom. We use (4.14) to write the following Hamiltonian operator:

$$H'''_2=-\frac{e}{mc}\mathbf{p}'''\cdot\mathbf{A}^{(p)},$$
$$\mathbf{A}^{(p)}=-\frac{e^2}{2mc^2}\frac{1}{n}\exp\left(\frac{i}{c}\omega_0 a\right)\mathbf{A}'', \quad (5.2)$$

where $\omega_0=\omega_p^{(+)}-\omega_n^{(+)}$ is the frequency of the $p\to n$ transition. We call $\mathbf{A}^{(p)}$ the positron polarizing field. The polarizing field in (5.2) is due to the disappearance of a photon at the point occupied by the second atom, while the absorption of the photon occurs at the point occupied by the first atom. The other terms in the operator (4.14) corresponding to orbital quantum transitions differ from (5.2) in their physical properties. For instance, a term of the form $\mathbf{p}''\cdot\mathbf{A}''$ corresponds to a positron polarizing field, but here the disappearance of a photon and its absorption occurs at the same point.

Thus, allowing for intermediate positron states in the interaction of two atomic electrons gives rise to an additional Hamiltonian for the interaction of the atomic electron with the external field. As a result, in examining the interaction of a system of atomic electrons and an external field we must add to the external field $\mathbf{A}'''$ the positron polarizing field $\mathbf{A}^{(p)}$.

Let us examine the role of the interaction (4.15) via intermediate states with positive energies, allowing only for orbital quantum transitions. Substituting the operators (4.3) and (3.9) in (4.15), we arrive at the following interaction Hamiltonian:

$$H'''_3=-\frac{e}{mc}\mathbf{p}'''\cdot\mathbf{A}^{(e)}, \quad (5.3)$$

where $\mathbf{A}^{(e)}$ is the vector potential of the electron polarizing field,

$$\mathbf{A}^{(e)}=\exp\left(\frac{i}{c}\omega_0 a\right)\sum_{l_\pm}\left(\frac{\mathbf{p}'_{rl}\cdot\mathbf{A}'}{\hbar[\omega_l(1-i0)+\omega-\omega_r^{(+)}]}\right.$$
$$\times\left\{\frac{e}{im\omega_0}\frac{\mathbf{d}''_{lm}-3(\mathbf{d}''_{lm}\cdot\mathbf{n})\mathbf{n}}{a^3}+\frac{e^2}{im^2c\omega_0}\right.$$
$$\left.\times\frac{\mathbf{p}''_{lm}-3(\mathbf{p}''_{lm}\cdot\mathbf{n})\mathbf{n}}{a^2}+\frac{e^2}{m^2c^2}\left[\left(\frac{R_{1l}}{2}-1\right)\frac{\mathbf{p}''_{lm}}{a}\right.\right.$$

$$-\frac{R_{1l}}{2}\frac{(\mathbf{n}\cdot\mathbf{p}''_{lm})\mathbf{n}}{a}-\frac{R_{1l}}{2}\frac{\mathbf{p}''_{lm}-3(\mathbf{n}\cdot\mathbf{p}''_{lm})\mathbf{n}}{a}\Bigg]\Bigg\}$$
$$+\frac{1}{\hbar[\omega_l(1-i0)-\omega-\omega_m^{(+)}]}\Bigg\{\frac{e}{im\omega_0}\frac{\mathbf{d}'_{rl}-3(\mathbf{d}'_{rl}\cdot\mathbf{n})\mathbf{n}}{a^3}$$
$$+\frac{e^2}{im^2c\omega_0}\frac{\mathbf{p}'_{rl}-3(\mathbf{p}'_{rl}\cdot\mathbf{n})\mathbf{n}}{a^2}+\frac{e^2}{m^2c^2}\Bigg[\left(\frac{R_{2l}}{2}-1\right)\frac{\mathbf{p}'_{rl}}{a}$$
$$-\frac{R_{2l}}{2}\frac{(\mathbf{n}\cdot\mathbf{p}'_{rl})\mathbf{n}}{a}-\frac{R_{2l}}{2}\frac{\mathbf{p}'_{rl}-3(\mathbf{n}\cdot\mathbf{p}'_{rl})\mathbf{n}}{a}\Bigg]\Bigg\}\mathbf{p}''_{rl}\cdot\mathbf{A}''\Bigg), \tag{5.4}$$

with

$$R_{1l}=\frac{\omega_n^{(+)}-\omega_p^{(+)}}{\omega_l-\omega_m^{(+)}},\quad R_{2l}=\frac{\omega_n^{(+)}-\omega_p^{(+)}}{\omega_l-\omega_r^{(+)}}.$$

The electron polarizing field (5.4) is written in the electric dipole approximation, which uses the operator (3.9) to account for the exchange of virtual photons. In (5.4) the operators $\mathbf{A}'$ and $\mathbf{A}''$ have been taken outside the matrix-element sign in this approximation.

The electron polarizing field (5.4) is formed by two atoms, and one of these atoms reach the initial quantum state as a result of the sequence of quantum transitions considered above. This means that in (5.4) we can identify the average value of the electric dipole moment of that atom in a state $r=m$. We write the expression for the average value of the dipole moment $\mathbf{d}'_m$ calculated in first-order perturbation theory:[18]

$$\mathbf{d}'_m=\frac{e}{mc\hbar}\exp(-i\omega t)\sum_{l_+}\Bigg\{\frac{\mathbf{d}'_{ml}(\mathbf{p}'_{lm}\cdot\mathbf{A}'_0)}{\omega_{lm}-\omega+i(\Gamma_l+\Gamma_m)/2}$$
$$+\frac{(\mathbf{p}'_{ml}\cdot\mathbf{A}'_0)\mathbf{d}'_{lm}}{\omega_{lm}+\omega-i(\Gamma_l+\Gamma_m)/2}\Bigg\}, \tag{5.5}$$

where ω_{lm} is the transition frequency, $\Gamma^{-1}_{l(m)}$ is the lifetime of the state $l\,(m)$, and $\mathbf{A}'=\mathbf{A}_0\exp(-i\omega t)$. In (5.5) we have kept only the negative-frequency part, corresponding to the photon absorption process, in which the polarizing field (5.4) is formed. Using (5.5), instead of (5.4) we arrive at the formula (the label m has been dropped)

$$\mathbf{A}^{(e)}=\frac{c}{i\omega_0}\frac{[\mathbf{d}']-3([\mathbf{d}']\cdot\mathbf{n})\mathbf{n}}{a^3}$$
$$+\frac{e}{im\omega_0}\frac{[\mathbf{p}']-3([\mathbf{p}']\cdot\mathbf{n})\mathbf{n}}{a^2}$$
$$-\frac{i\omega_0}{c}\frac{([\mathbf{d}']\cdot\mathbf{n})\mathbf{n}}{a}-\frac{e}{mc}\frac{[\mathbf{p}']}{a}, \tag{5.6}$$

where $[\cdots]$ indicates that the average value of the quantity is taken at time $t'=t-\omega_0 a/\omega c$. Here we have used the fact that atoms are isotropic, according to the quantities $(\mathbf{p}'_{rl}\cdot\mathbf{A}')\mathbf{d}'_{lm}$ and $\mathbf{p}'_{rl}(\mathbf{d}'_{lm}\cdot\mathbf{A}')$ are equal.

In (5.2) and (5.3) we go from the vector potentials of the polarizing fields to the corresponding electric and magnetic fields via the Lorentz condition $\partial A_\mu/\partial x_\mu=0$ (Ref. 13). For the a field proportional to $\exp(-i\omega t)$ we have

$$\mathbf{E}^{(e)}_-=\frac{i\omega}{c}\mathbf{A}^{(e)}_-,\quad \mathbf{E}^{(p)}_-=\frac{i\omega}{c}\mathbf{A}^{(p)}_-, \tag{5.7}$$

where the vector potentials $\mathbf{A}^{(e)}_-$ and $\mathbf{A}^{(p)}_-$ are proportional to the operator $c^v_{\mathbf{k}\lambda}$ of annihilation of a photon with wave vector $\mathbf{k}$ and polarization $\lambda=1,2$ at the point occupied by the atom that generates the polarizing field (the polarizer atom). Here the quantum transition $p\to n$ takes place at point occupied by the other atom (the observer atom) located at an arbitrary distance from the polarizer atom.

Consider the case where the frequency ω of the external field is close to one of the frequencies $\omega_{lm}>0$. The average value of the momentum is $[\mathbf{p}']=(im/e)\omega_0[\mathbf{d}']$, according to (5.6), and the operator of the electron polarizing field is

$$\mathbf{E}^{(e)}_-=\nabla\times\nabla\times\frac{[\mathbf{d}]}{a}, \tag{5.8}$$

where differentiation is carried out with respect to the coordinates $\mathbf{a}'''$ of the point under observation. We define the magnetic fields $\mathbf{H}^{(e)}_-$ and $\mathbf{H}^{(p)}_-$ in similar way using the common relationship between the vector potential and the magnetic field.[13]

5.1. Integral equations for photon propagation in an electric-dipole optical medium

We introduce the dipole (α_d) and momentum (α_p) polarizabilities of an atom by the following relations:

$$\mathbf{d}'=\alpha_d\mathbf{A}',\quad \mathbf{p}'=\alpha_p\mathbf{A}', \tag{5.9}$$

where $\mathbf{d}'$ and $\mathbf{p}'$ are the average values of the dipole moment and the momentum in a certain state m, values calculated in first-order perturbation theory. We can determine the dipole polarizability of an isotropic atom in accordance with (5.5). To determine the momentum polarizability α_p, we must replace the matrix elements $\mathbf{d}'_{lm}$ with the matrix elements $\mathbf{p}'_{lm}$ of the momentum operator in (5.5). Then the vector potential of the electron polarizing field takes the form

$$\mathbf{A}^{(e)}=\Bigg\{\alpha_d\frac{c}{i\omega_0}\frac{\mathbf{e}-3(\mathbf{e}\cdot\mathbf{n})\mathbf{n}}{a^3}+\alpha_p\frac{c}{im\omega_0}\frac{\mathbf{e}-3(\mathbf{e}\cdot\mathbf{n})\mathbf{n}}{a^2}$$
$$-\alpha_d\frac{i\omega_0}{c}\frac{(\mathbf{e}\cdot\mathbf{n})\mathbf{n}}{a}-\frac{e}{mc}\alpha_p\frac{\mathbf{e}}{a}\Bigg\}[\mathbf{A}']\equiv K_e(\mathbf{a}',\mathbf{a}''')[\mathbf{A}'], \tag{5.10}$$

where $\mathbf{e}$ is the unit vector along the direction of field $\mathbf{A}'$.

The transition to an N-atom system can be carried out by summing the vector potentials (5.10) and (5.2) of the polarizing fields generated by $N-1$ atoms at the point occupied by the atom with the radius vector $\mathbf{a}''$. We can estimate the role of the electron and positron polarizing fields in such a system of atoms by comparing the terms in (5.10) proportional to $1/a$ with the vector potential (5.2). A fact worth noting in this connection is that the electron polarizing field contains the polarizabilities α_d and α_p, which depend on the

random distribution of the natural frequencies due to inhomogeneous broadening, while the positron polarizing field is independent of this broadening. This leads to a situation in which under certain conditions the electron and positron polarizing fields in an N-atom system may be comparable in absolute value.

Now let us turn to the case of an optical medium, which we assume to be continuous. To this end we introduce the density N/V of the distribution of atoms in the medium. More than that, we assume that the polarizing fields (5.10) and (5.2) are proportional to the field inside the medium rather than to external field. This makes it possible to write the following integral equation for the operator the electric field strength:

$$\mathbf{E}(\mathbf{r},t)=\mathbf{E}_{\mathrm{I}}(\mathbf{r},t)+\int \frac{N}{V}K_e(\mathbf{r},\mathbf{r}')\mathbf{E}\left(\mathbf{r}',t-\frac{R}{c}\right)dV' + \int \frac{N}{V}K_p(R)\mathbf{E}\left(\mathbf{r}',t-\frac{R}{c}\right)dV', \tag{5.11}$$

where $\mathbf{r}$ is the radius vector of the point under observation, $\mathbf{r}'$ is the radius vector of a point inside the medium or at its surface, $\mathbf{E}_I(\mathbf{r},t)$ is the operator of the external electric field represented in the form of a linear combination of plane waves with amplitudes $c^{v}_{\mathbf{k}\lambda}$ (Ref. 13) that are coordinate-independent, $R=|\mathbf{r}-\mathbf{r}'|$, and

$$K_p(\mathbf{R})=-\frac{e^2}{2mc^2}\frac{1}{R}\mathbf{e}. \tag{5.12}$$

An integral equation for the operator of the magnetic field can be derived in a similar manner by applying the curl operator to the vector potentials (5.10) and (5.2).

If the point under observation, $\mathbf{r}$, lies outside the medium, the integral in (5.11) is over the entire medium. If this point is inside the medium, we must first exclude a small region occupied by the atom and surrounded by a sphere of a small radius L_0. Equation (5.11) in the particular case of zero positron polarization of the medium and only one natural frequency in the spectrum of the atoms coincides with the integro-differential equation of classical optics[19] if the operators in (5.10) are replaced by the corresponding classical fields.

We intend to use Eq. (5.11) to solve boundary-value problems of quantum optics, since this equation incorporates the boundary conditions.

5.2. Integral equations of propagation of photons in a system of electron spins

Here is another example of an integral equation that can be derived by the proposed method. In contrast to the previous case, we consider only the spin degrees of freedom of interacting electrons belonging to different one-electron atoms located at arbitrary distances from one another. Such a situations is of interest, say, in magneto-optics, in the design of inversionless lasers, and in processes of laser cooling of atoms. Let us assume that spin transitions take place between atomic states separated by the optical frequency ω_0. Such transitions may occur independently of electric dipole transition, which were examined in Sec. 5.1.

We use the operator (3.6) and then the operator

$$(R_2^{(+)})_\sigma=-\frac{i\hbar e}{2mc}\boldsymbol{\sigma}'\cdot\mathbf{H}' \tag{5.13}$$

to describe the polarizing fields in the system of electron spins. Here we consider the same pattern of quantum transitions as in Sec. 5.1.

We insert the operators (3.6) and (5.13) into (4.15). This leads us to the interaction operator

$$H_3'''=-\frac{\hbar e}{2mc}\boldsymbol{\sigma}'''\cdot\mathbf{H}^{(e)}, \tag{5.14}$$

where the magnetic field

$$\mathbf{H}^{(e)}=\exp\left(\frac{i}{c}\omega_0 a\right)\frac{e^2\hbar^2}{4m^2c^2}\sum_{l_+}\left(\frac{\boldsymbol{\sigma}'_{rl}\cdot\mathbf{H}'}{\hbar[\omega_l(1-i0)+\omega-\omega_r^{(+)}]}\times\left\{\frac{\boldsymbol{\sigma}''_{lm}-3(\boldsymbol{\sigma}''_{lm}\cdot\mathbf{n})\mathbf{n}}{r^3}+15\frac{a^2}{r^5}R_{1l}(\boldsymbol{\sigma}''_{lm}\cdot\mathbf{n})\mathbf{n} - 9\frac{a^2}{r^5}R_{1l}\boldsymbol{\sigma}''_{lm}\right\}+\frac{1}{\hbar[\omega_l(1-i0)-\omega-\omega_m^{(+)}]}\times\left\{\frac{\boldsymbol{\sigma}'_{rl}-3(\boldsymbol{\sigma}'_{rl}\cdot\mathbf{n})\mathbf{n}}{r^3}+15\frac{a^2}{r^5}R_{2l}(\boldsymbol{\sigma}'_{rl}\cdot\mathbf{n})\mathbf{n} - 9\frac{a^2}{r^5}R_{2l}\boldsymbol{\sigma}'_{rl}\right\}\right). \tag{5.15}$$

This is an electron polarizing field, i.e., it is realized via only intermediate electron states. Equation (4.14) implies that can be no positron polarizing field in this pattern of quantum transitions.

In (5.15) we distinguish the average values of the spin magnetic moments of the polarizer atom by a formula similar to (5.5) in first-order perturbation theory. We introduce the following notation for the average values of the spin variables in state m:

$$\boldsymbol{\sigma}'_m=(\alpha_\sigma)_m\mathbf{H}', \quad \mathbf{q}'_m=(\alpha_q)_m\mathbf{H}', \tag{5.16}$$

where $(\alpha_\sigma)_m$ is the spin polarizability of the state in state m, $\mathbf{q}'_m$ is the average value in state m of the operator $\mathbf{q}'$ with the matrix elements $\mathbf{q}'_{lm}=\boldsymbol{\sigma}'_{lm}/\omega_{lm}$, and $(\alpha_q)_m$ is the corresponding polarizability. Theexpression for the polarizing field (5.15) becomes

$$\mathbf{H}^{(e)}=\mu_B\left\{\alpha_\sigma\frac{\mathbf{h}-3(\mathbf{h}\cdot\mathbf{n})\mathbf{n}}{r^3}-15\omega_0\alpha_q\frac{a^2}{r^5}(\mathbf{h}\cdot\mathbf{n})\mathbf{n} + 9\omega_0\alpha_q\frac{a^2}{r^5}\right\}[\mathbf{H}']\equiv K_\sigma(\mathbf{a}',\mathbf{a}''')[\mathbf{H}'], \tag{5.17}$$

where μ_B is the Bohr magneton, and **h** is the unit vector directed long the magnetic field vector $\mathbf{H}'$.

The passage to the integral equation of photon propagation in a continuous optical medium can be achieved in the same way as in Sec. 5.1, i.e., by replacing the external field in (5.17) with the field inside the medium. Then, for spin quantum transitions, the operator of the magnetic field at a certain point **r** under observation at time t inside or outside the medium becomes

$$\mathbf{H}(\mathbf{r},t)=\mathbf{H}_{\mathrm{I}}(\mathbf{r},t)+\int \frac{N}{V}K_{\sigma}(\mathbf{r},\mathbf{r}')\mathbf{H}\left(\mathbf{r}',t-\frac{R}{c}\right)dV', \qquad (5.18)$$

where $\mathbf{H}_{\mathrm{I}}(\mathbf{r},t)$ is the operator of the external magnetic field.

6. DISCUSSION

Our results can by divided into two main categories. The first deals with the problem of solving an important QED problem, the problem of two interacting electrons. The second is devoted to using the solution in deriving integral field equations in optics.

The interaction of two atomic electrons belonging to two immobile atoms located at arbitrary distances from one another is a considered a third-order QED effect, where an integral part is the exchange of virtual photons involving various intermediate states with positive and negative energies. It appears that here the various patterns of quantum transitions leading to emission or absorption of a single photon in the system of two interacting hydrogenlike atoms can be classified. In the present paper we discuss in detail the pattern of quantum transitions in which one atom is a polarizer atom and the other, an emitter or absorber atom.

In contrast to the earlier papers to the two-electron problem, here for the first time we account for all the intermediate states in the spectrum of the interacting atoms rather than only for the ω_l-states, in which the transition frequency ω_{lm} is close to the frequency ω of a real photon. Our allowance for these states is related to the role that the factors R_{sl} play in the operators B_{sl}, which describe virtual photon exchange.

The earlier papers focused on the interaction of two hydrogenlike atoms in the electric dipole approximation. In this paper we also allow for other types of quantum transition. We show that the spin–orbit and spin–spin coupling operators acquire additional terms in comparison to the Breit operator. These terms arise because of additional retardation in the interaction of two atoms located at arbitrary distances from one another. Here we introduce a universal time scale, the period of optical oscillations, for different types of quantum transition.

In the present paper we have shown that three types of field act on each atom: an external field (including the vacuum field), an electron polarizing field, and a positron polarizing field. We have established that the disappearance of a photon may occur at the point occupied by one of the atoms of the system, while the absorption of that same photon takes place at the point occupied by the other atom, which is at an arbitrary distances from the polarizer atom. A similar situation occurs for the emission of a photon: emission takes place at the point occupied by the emitter atom, which is at an arbitrary distance from the polarizer atom, where the photon was created.

The results of studies of the processes of interaction of two atom in a light field can be used to derive new integral equations that describe the propagation of photons in a continuous optical medium. To demonstrate the effectiveness of the method of integral equations, we have derived to integral equations corresponding to electric and spin transitions in the spectrum of the interacting atoms. We have found that in the particular case where the atomic spectrum has only one transition with frequency $\omega_{lm}=0$, the integral equation of photon propagation in an electric-dipole medium coincides with the corresponding integral equations of classical optics.

Let us examine the feasibility of replacing differential equations by integral equations in this problem. Actually, the integral equations for the electric and magnetic fields of classical optics are equivalent to the Maxwell equations. However, by using integral equations one can derive in a rigorous manner the Lorenz–Lorentz formula and the extinction theorem and solve a number of important optical problems. And yet this powerful method has rarely been used. This is because the integral equations known from classical optics have meaning only for isotropic nonmagnetic dielectric media. Hence it would be interesting to derive integral equations that would describe a much broader class of phenomena. The present paper is an attempt to do this by studying the two-electron problem in the QED setting.

By studying the interaction of two electron in a light field we can distinguish the various type of quantum transition. In particular, as shown in this paper, we can separate the positron polarizing field, which in some cases is not negligible in comparison to the electron polarizing field. Allowing for the correspondence between Eq. (5.11) and the classical integral equation, we state that the positron polarizing field can be interpreted by introducing an additional current into the Maxwell equations. Indeed, the derivation of the integral equation of classical optics[19] is based in Ref. 4 on the expansion of the retarded potentials. Here the external field can be accounted for by two methods. The first is based on the idea of polarizability of atoms, which make it possible to allow for what is known in optics as the dipole field. The second method is based on the idea of replacing the electron momenta **p** by $\mathbf{p}-(e/c)\mathbf{A}$. In the present we have shown that, from the viewpoint of quantum electrodynamics, these two methods correspond to electric and positron polarizing fields. The electron polarizing field in the electric-dipole approximation is the dipole field known from optics, a field generated only by electron states in the spectrum of the interacting atoms. The positron polarizing field is generated only by positron states. Here we are dealing not with real positron states but with virtual states, and the creation of such states is not guided by the law of energy conservation. Thus, the study of the two-electron problem in the QED setting makes it possible not only to develop the method of integral equations but also to establish the new mechanism of emission and absorption of real photons in a system of interacting atoms.

Two types of atom–field interaction can be identified

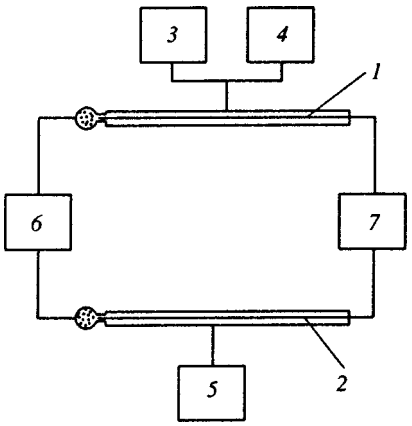


FIG. 2. Block diagram of an experimental devices for studying the polarizing fields of atoms in an external light field: *1*—beam of atoms in ground state, *2*—beam of atoms in excited state, *3*—spectrophotometer, *4* and *5*—devices for sorting the atoms in the ground and excited states, *6*—synchronizer, and *7*—reading device.

here. The first corresponds to the transfer of real photons from excited atoms to unexcited. Here the time of radiation transfer from one atom to another is determined by the atomic separation. This type of interaction can serve as a base for deriving the corresponding integral equations for radiation transfer, which are used, e.g., in the optics of turbulent media.[20] The other type based on the idea of polarizing fields, interpreted in the present paper as third-order QED effects. In this paper we have shown that a photon disappears at the point occupied by one of the interacting atoms, while the absorption of that photon takes place at the point occupied by the other atom. And, according to the meaning of the expansion in Sec. 2, the distance between the atoms may be arbitrary. Obviously, a similar situation occurs for the process of emission of a photon in a system of two interacting atoms.

We note an important feature of this type of interaction. Specifically, the disappearance of a photon and its absorption occur at the same moment in time, i.e., we can speak of instantaneous formation of a polarizing field due to continuous exchange of virtual photons between the atomic electrons. If this was not so, the law of energy conservation (Eq. (2.10)) would be violated. The retardation factor $\exp[(i/c)\omega_0 a]$ in the various terms of the matrix (2.1) points only to a periodic dependence of the effective interaction energy on the interatomic distance. In other words, we can speak of instantaneous energy transfer between different atoms located at arbitrary distances from one another provided that there is continuous exchange of virtual photons.

Figure 2 depicts the block diagram of an experimental device that could be used to confirm or reject this property of polarizing fields. For the experiment we select hydrogenlike atoms, say Cs and Rb^{85}, or hydrogen atoms, whose spectrum has been thoroughly studied. In this paper we examine polarizing fields generated by electric dipole or spin transitions. However, by using the above method, polarizing fields for other types of quantum transition can also be considered. Hence it is advisable in our experiment to select quantum transitions in the spectrum of atoms with large wavelengths (e.g., the wavelength of the emission line in a hydrogen laser), which would substantially reduce the effect of retardation. The experimental device consists of two types of atom-ray tubes. Tube *1* is used for the atom beam in which the atoms are in their ground state. This beam is irradiated by light (photons). Spectrophotometer *3* registers the evvents in which photons disappear, events that are due either to induction of polarizing fields or to direct absorption by resonant atoms. In the latter case the atoms become excited and can be sorted in the atomic beam by separator *4*. In tube *2* synchronized with tube *1* by synchronizer *6* there is a beam of atoms in their ground state. The polarizer atoms in beam *1* act on the atoms in beam *2* in such a way that resonant photons are absorbed, the absorber atoms become excited, and are sorted by device *5*. The absorption probability can be calculated to the first order via the operator (5.3) with the electron polarizing field (5.10). This makes it possible to calculate the number of atoms participating in the formation of the polarizing field. However, in addition to this process, atoms in tube *2* become excited due to re-emission of photons by the atoms in tube *1*, which is located at a distance L from tune *2*. This process has a time lag equal to L/c and can be separated from the process of formation of polarizing field by using the reading device *7*.

The present work was supported by grants from the Russian Fund for Fundamental Research (Grant 98-02-16035) and the Federal ''Integration'' Program (Grant K0179).

*)E-mail: gadomsky@quant.univ.simbirsk.su

[1] C. G. Darwin, Trans. Cambridge Philos. Soc. **23**, 137 (1924).
[2] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, Oxford (1975).
[3] O. N. Gadomskiĭ, V. R. Nagibarov, and N. K. Solovarov, Zh. Éksp. Teor. Fiz. **63**, 813 (1972) [Sov. Phys. JETP **36**, 426 (1973)].
[4] O. N. Gadomskiĭ and R. A. Vlasov, *Optical Echo Spectroscopy of Surfaces* [in Russian], Nauka i Tekhnika, Minsk (1990).
[5] O. N. Gadomsky, V. R. Nagibarov, and N. K. Solovarov, Phys. Lett. A **42**, 219 (1972).
[6] G. Breit, Phys. Rev. **34**, 553 (1929).
[7] L. D. Landau, Sov. Phys. **8**, 487 (1932).
[8] H. Bethe and E. Fermi, Z. Phys. **77**, 296 (1932).
[9] C. S. Chang and P. Stehle, Phys. Rev. A **4**, 630 (1971).
[10] O. N. Gadomskiĭ, V. P. Nagibarov, and N. K. Solovarov, Zh. Éksp. Teor. Fiz. **70**, 435 (1976) [Sov. Phys. JETP **43**, 225 (1976)].
[11] E. M. Lifshitz, Zh. Éksp. Teor. Fiz. **18**, 562 (1948).
[12] B. D. Fedyushin, Zh. Éksp. Teor. Fiz. **22**, 140 (1952).
[13] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, Wiley, New York (1974).
[14] G. W. F. Drake, Phys. Rev. A **5**, 1979 (1972).
[15] O. N. Gadomsky and K. V. Krutitsky, J. Opt. Soc. Am. B **13**, 1679 (1996).
[16] O. N. Gadomskiĭ and K. V. Krutitskiĭ, Zh. Éksp. Teor. Fiz. **106**, 936 (1994) [JETP **79**, 513 (1994)].
[17] O. N. Gadomskiĭ and S. G. Moiseev, Zh. Éksp. Teor. Fiz. **113**, 471 (1998) [JETP **86**, 259 (1998)].
[18] A. S. Davydov, *Quantum Mechanics*, 2nd ed., Pergamon Press, Oxford (1976).
[19] M. Born and E. Wolf, *Principles of Optics*, 6th ed., Pergamon Press, Oxford (1980).
[20] A. Ishimaru, *Wave Propagation and Scattering in Random Media*, 2 vols., Academic Press, New York (1978).

Translated by Eugene Yankovsky

Dynamic features of interaction between a sequence of ultrashort laser pulses and planar, thin-film microcavities

V. A. Goryachev*) and S. M. Zakharov

Institute of High-Performance Computing Systems, Russian Academy of Sciences, 117872 Moscow, Russia
(Submitted 30 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1578–1594 (November 1998)

A theory of coherent interaction between a sequence of ultrashort laser pulses and planar, thin-film microcavities in the form of Fabry–Perot resonators containing resonating atoms has been developed. The dynamics of transmission of single laser pulses has been analyzed numerically. Analytic solutions of the problem of four-wave mixing of optical fields separated in time have been obtained in the small-area approximation. The dynamic efficiency of conversion of incident waves to the coherent response field (photon echo) generated in a microcavity can be higher than in the case of a bulk resonant structure. Specific features of photon echo generation in a microcavity for arbitrary ''areas'' of laser pump pulses are discussed.
 [S1063-7761(98)00311-4]

1. INTRODUCTION

Interest in studies of nonlinear optical properties of thin-film microcavities has been aroused recently.[1–10] In the context of basic research, planar microcavities present an example of a model physical system for studying various physical processes, such as optical bi- and multistabilities, self-pulsations of light, solitonic transmission regime of ultrashort optical pulses, generation of coherent light, and dynamical chaos. On the other hand, there is a pragmatic interest in developing such structures since they can be used as components of electronic optical devices.

Previously interaction between isolated ultrashort laser pulses and thin-film microcavities filled with resonant atoms has been studied in detail, and it turned out that bi- and multistable regimes of transmission of ultrashort pulses can occur in a nonstationary regime when the pulse duration is less than the irreversible polarization time (transverse relaxation) and the lifetime of resonating atoms in an excited state (longitudinal relaxation).[11,12] An important parameter describing multivalued nonlinear optical characteristics of microcavities was the laser pulse ''area.'' At the same time, it was shown that the anomalous transmission of ''soliton-like'' pulses through a thin layer is due to the existence of an integral of the motion determined by the form of the underlying equations of motion.[11]

Effects of bistability in states of such systems occur at sufficiently high atom densities inside their cavities or sufficient reflectivities of their mirrors. Otherwise a peculiar mechanism of field energy loss inside the cavity or due to reflection from its boundaries is active. These properties of the systems (the existence of bistable states and peculiar dissipative property of the medium) are notably different from those of long, bulk resonant media. Therefore, it seems worthwhile to investigate various coherent processes typical of bulk active media in the context of resonant cavity excitation, including photon echo phenomena. An important component of this investigation is a consistent account of conditions at the boundaries of a resonant medium and formation of standing waves.

The first investigation of photon echo generation by optical pulses in the form of standing waves was undertaken by Le Gouet and Berman[13] and tackled gaseous resonant media. The main feature of the two-pulse photon echo generated in a resonant gaseous medium was the presence of multiple signals, which was, in the long run, a specific manifestation of inhomogeneous broadening. It is widely known that the inhomogeneous broadening in gases is caused by the Doppler effect, since a moving atom ''sees'' an optical field at a frequency shifted by $\mathbf{k}\cdot\mathbf{v}$, where $\mathbf{k}$ is the optical wave vector and $\mathbf{v}$ is the thermal velocity of the atom.[14,15] As a result of this atomic motion in space, atoms offset from resonance by ε go to regions with different phase relations in the standing wave field. Similar effects occur in spatially separated light beams.[16] It is also important that in Doppler-induced inhomogeneous broadening, a particular asymmetry emerges in averaging over the detuning parameter ε, since atoms with detuning parameters ε of opposite sign travel in opposite directions. This is the fundamental difference between gaseous media and solids, where inhomogeneous broadening is usually due to irregularities in the crystal field and is local in nature. As a result, such features of the photon echo typical of gaseous media as the multiple photon echo can take place in solids only at considerably high nonlinearity parameters.

In this paper, we theoretically investigate features of interaction between the sequence of ultrashort laser pulses and thin-film microcavities of finite widths and containing resonant atoms inside Fabry–Perot resonators. Computer simulations of dynamic processes indicate that the following effects can be observed in thin microcavities: a) self-induced transparency, which leads to splitting of a 4π-pulse into a set of two 2π-pulses with differing temporal dynamics and delay times (this splitting is controlled by optical cavity parameters and duration of input pulses); b) population inversion in the ensemble of atoms at frequencies close to the atomic reso-

nance; c) generation of a multiple photon echo with an excess delay close to the photon lifetime in the cavity. Dynamic properties of photon echo signals are also determined by transient processes associated with formation of the optical field configuration inside the cavity and finite photon lifetime in the cavity. Moreover, the dynamic efficiency during holographic shaping of nonstationary patterns created in microcavities at the instant that photon echos emerge can be substantially greater than in the case of bulk excitation of resonant media. We discuss a number of feasible practical applications of these features.

2. BASIC EQUATIONS AND SOLUTION TECHNIQUES

As in our previous publications,[11,12] we define ultrashort laser pulses interacting with planar, thin-film microcavities as optical pulses whose duration δ satisfies

$$T_2^* \ll 2Ln_0/c \ll \delta \ll T_2 \approx \gamma_\perp^{-1} \ll T_1 \approx \gamma_\parallel^{-1}, \tag{1}$$

where T_2^* and T_2 are the times of reversible and irreversible polarization relaxation in the resonant medium, respectively, T_1 is the population relaxation time, and $2Ln_0/c$ is the round-trip time of an optical pulse in the cavity.

Note also that the relation between δ and the photon lifetime in the cavity,

$$\tau_c = -\frac{2Ln_0}{c\ln(R_1R_2)} \approx \frac{2Ln_0}{c(1-R_1R_2)},$$

can be arbitrary by virtue of the inequality $2Ln_0/c \ll \delta$ [Eq. (1)] and the condition

$$1 - R_1R_2 \ll 1,$$

where $R_{1,2}$ are mirror reflectivities. Another important point is that the inequality $T_2^* \ll 2Ln_0$ [Eq. (1)] means that an arbitrary number of longitudinal cavity modes can be contained within the spectral width of an inhomogeneously broadened spectral line.

The basic equations describing interaction between ultrashort laser pulses and microcavities containing two-level resonant atoms were derived in a previous study.[11] In the approximation defined above, in the absence of phase modulation of the external field, and given a ''tuned'' cavity, these equations have the form

$$\begin{aligned}
&\frac{\tilde{\tau}_c}{2}\frac{d}{dt}E_+ + E_+ - C'\langle V_\varepsilon^+\rangle = E_0'(t),\\
&\frac{dU_\varepsilon^+}{dt} + \varepsilon V_\varepsilon^+ = 0,\\
&\frac{dV_\varepsilon^+}{dt} - \varepsilon U_\varepsilon^+ = \frac{|d_{12}|^2}{\hbar} n_\varepsilon E_+,\\
&\frac{dn_\varepsilon}{dt} = -\frac{1}{\hbar}(R_2+1)E_+V_\varepsilon^+,
\end{aligned} \tag{2}$$

where E_+ is the slowly changing electric field amplitude of the optical wave traveling through the cavity in the same direction as the incident wave, U_ε^+ and V_ε^+ are the real and imaginary parts of the slowly changing dipole moment amplitude P_ε^+, n_ε is the population inversion of resonant atoms, with

$$\begin{aligned}
&P_\varepsilon^+ = (U_\varepsilon^+ + iV_\varepsilon^+)\exp(i\varphi),\\
&\tilde{\tau}_c = (1+\sqrt{R_1R_2})^2\tau_c, \quad E_0'(t) = \frac{\sqrt{1-R_1}}{1-\sqrt{R_1R_2}}E_0(t),\\
&C' = \frac{2\pi\omega L N_0(1+\sqrt{R_1R_2})}{cn_0(1-\sqrt{R_1R_2})},
\end{aligned} \tag{3}$$

d_{12} is the reduced matrix element of the resonant transition dipole moment, N_0 is the density of resonant atoms inside the cavity, n_0 is the nonresonant refraction index, c is the speed of light, and angular brackets denote averaging over the offset parameter $\varepsilon = \omega_{21} - \omega$ with a weight function $G(\varepsilon)$, which is the shape of the inhomogeneously broadened line:

$$\langle V_\varepsilon^+\rangle = \int_{-\infty}^{\infty} d\varepsilon\, G(\varepsilon) V_\varepsilon^+.$$

Certain features of solutions describing isolated ultrashort pulses were discussed previously.[11,12] In general, Eqs. (2) describe an interaction problem between an ultrashort pulse and a microcavity that is nonstationary and nonlinear in the field, and that can probably only be solved numerically.

Now let us transform Eqs. (2) to a form convenient for numerical integration. To this end, we introduce dimensionless variables[12] that are functions of the dimensionless time $\tau = t/T_2^*$:

$$\begin{aligned}
&e(\tau) = \frac{|d_{12}|T_2^*\sqrt{1+R_2}\,E_+}{\hbar},\\
&e_0(\tau) = \frac{|d_{12}|T_2^*\sqrt{1+R_2}\,E_0'}{\hbar},\\
&v(\tau,x) = \frac{\sqrt{1+R_2}\,V_\varepsilon^+}{|d_{12}|},\\
&u(\tau,x) = \frac{\sqrt{1+R_2}\,U_\varepsilon^+}{|d_{12}|},\\
&n(\tau,x) = n_\varepsilon, \quad x = \varepsilon T_2^*, \quad \tau_0 = \tau_c/T_2^*.
\end{aligned} \tag{4}$$

Then the basic system of equations (2) take the form

$$\begin{aligned}
&\tau_0\dot{e} + e(\tau) = e_0 - CY(\tau),\\
&\dot{v} - xu(\tau,x) = n(\tau,x)e(\tau),\\
&\dot{u} + xv(\tau,x) = 0,\\
&\dot{n} = -v(\tau,x)e(\tau),
\end{aligned} \tag{5}$$

where the parameter C is defined by the expressions

$$C = \frac{\alpha_0 L(1+\sqrt{R_1R_2})}{2(1-\sqrt{R_1R_2})}, \tag{6a}$$

$$\alpha_0=\frac{4\pi^2\omega|d_{12}|^2N_0G(0)}{\hbar cn_0}, \tag{6b}$$

α_0 is the absorption factor in the low-intensity limit, function $Y(t)$ describes inhomogeneous broadening of energy levels:

$$Y(\tau)=\frac{2}{\pi}\int_0^\infty g(x)v(\tau,x)dx, \tag{7}$$

and the dot over variables denotes differentiation with respect to time.

Note that the spectral line profile plotted against the offset in our numerical model is chosen to be symmetrical—in particular, $g(x)=g(0)\exp(-x^2)$—and the electromagnetic field frequency is set at the spectral line peak.

Equations (5) and (7) with initial conditions $e(-\infty)=v(-\infty,x)=u(-\infty,x)=0$, $n(-\infty,x)=-1$ describe evolution of the electromagnetic field of the optical wave interacting with a microcavity, given an incident wave form defined by function $e_0(\tau)$. Note that in the absence of reflection at the cavity boundaries ($R_1\to0$, $R_2\to0$), $\tau_0\to0$ and Eqs. (5) describe transmission of an ultrashort pulse through a thin film of resonant atoms.

The choice of the specific technique used in numerical integration of Eqs. (5) and the algorithm based on this technique are discussed in the Appendix.

An important method used in solving the problem of photon echo generated in a microcavity by a sequence of several ultrashort laser pulses is the so-called ''small-areas'' approximation. In this case, the resonant medium inside the cavity can act as a dynamic hologram, and at the moment when the photon echo is generated, it can produce patterns that are functions of the original configuration of the exciting fields.

We now consider solutions of Eqs. (2) under these conditions. The basic tenet of this approximation is that the change in the population inversion of the resonant medium due to a pulse with an ''area'' $\theta<1$,

$$\theta=\left|\frac{d_{12}}{\hbar}\int_{-\infty}^{\infty}E(t)dt\right|,$$

is negligible.

One can derive from Eq. (2) a general solution for the active component of the dipole moment:

$$V_\varepsilon^+(t)=\frac{|d_{12}|^2}{\hbar}\int_{-\infty}^{\infty}E_+(t')n_\varepsilon(t')\cos[\varepsilon(t-t')]dt'. \tag{8}$$

At $n_\varepsilon(t')\approx-1$ and under the condition $\delta\gg T_2^*$, one can substitute in Eq. (8)

$$\int_0^\infty\cos(\varepsilon\tau)d\tau=\frac{1}{2}\int_{-\infty}^{\infty}\exp(i\varepsilon\tau)d\tau=\pi\delta(\varepsilon).$$

Then

$$\langle V_\varepsilon^+(t)\rangle=-\frac{|d_{12}|^2}{\hbar}\pi G(0)E_+(t), \tag{9}$$

and the equation for the field inside the cavity takes the simple linear form

$$\frac{\tilde\tau_c}{2}\dot E_+ +(1+C)E_+=E_0', \tag{10}$$

where the parameter C is still defined by Eq. (6).

Calculation of the solution to Eq. (10) is straightforward:

$$E_+(t)=\frac{2}{\tilde\tau_c}\int_0^\infty d\tau E_0'(t-\tau)\exp\left[-\frac{2(1+C)\tau}{\tilde\tau_c}\right]. \tag{11}$$

If $\delta\gg\tilde\tau_c/2(1+C)$, the solution (11) simplifies considerably:

$$E_+(t)=E_0'(t)/(1+C). \tag{12}$$

In this case, the field inside the cavity ''tracks'' the external field in a quasistationary manner, since the latter condition allows us to neglect the dependence of the electric field amplitude on τ in Eq. (11).

Note that Eq. (12) yields the steady-state electric field amplitude of the forward wave in the cavity,

$$E_+=\frac{\sqrt{1-R_1}\,E_0}{(1-\sqrt{R_1R_2})(1+C)},$$

and for $R_1\sim R_2\sim1$ that amplitude can be considerably greater than the incident wave amplitude (even when $C>1$). In this case, even though $\alpha_0L<1$,

$$E_+\sim\frac{E_0}{\sqrt{1-R}\,(1+C)},\quad C\sim\frac{\alpha_0L}{1-R}. \tag{13}$$

This result depends exclusively on cavity parameters, and can be of great significance in solving the problem of photon echo generation by such structures.

3. SPLITTING OF SINGLE ULTRASHORT PULSES IN A MICROCAVITY

Now let us consider the dynamics of interaction between a single ultrashort pulse and a microcavity taking as an example transmission of optical pulses through a thin planar layer of resonant atoms. The splitting of an optical pulse with area θ which is a multiple of 2π ($\theta=2\pi m$, $m=1,2,\cdots$) into isolated 2π-pulses in a large volume containing resonant atoms is a familiar phenomenon.[17] The shapes of such pulses do not change when they are transmitted through a resonant medium, as in a transparent material, and their delay Δt over distance L is proportional to their widths δ ($\Delta t=\alpha_0L\delta/2$), where α_0 is the absorption coefficient. Under the conditions of the problem under discussion, observation of this effect is questionable for two reasons: first, because L is small [by virtue of the condition $2Ln_0/c\ll\delta$ (Eq. (1)], and second, because of transformation of an isolated 2π-pulse to a pair of subpulses of approximately equal areas with delay $\Delta\tau$ determined by cavity parameters (the photon lifetime or film parameter[12,18]).

Figure 1 shows as an example the shapes of a soliton-like pulse, $e(\tau)$, whose initial profile is $e_0(\tau)=(2/\tilde\delta)\mathrm{sech}(\tau/\tilde\delta)$, with an area 2π and $\tilde\delta=0.5$ transmitted through a resonant medium at different values of the nonlinearity parameter C of the microcavity and $\tau_0=0.1$.

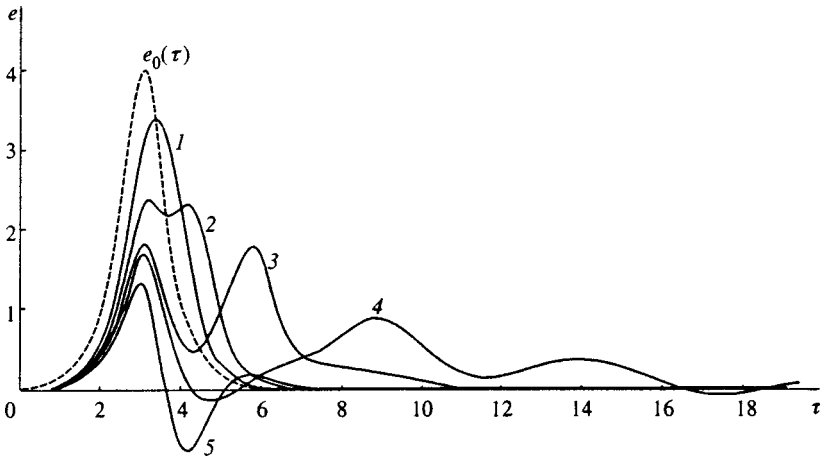


FIG. 1. Time dependencies of the microcavity response amplitude $e(\tau)$ to a single-soliton 2π-pulse $e_0(\tau)$ (dashed line) of duration $\tilde{\delta}=0.5$ with cavity parameters $\tau_0=0.1$, $C=1.0$, 2.0, 2.8, 3.0, and 4.0 (solid lines labeled by *1, 2, 3, 4,* and *5,* respectively).

Figures 2 and 3 show calculations of $e(\tau)$ at different C and τ_0 for a pulse with area $\theta_0=4\pi$.

These calculations show that as the cavity parameters τ_0 and C increase, the incident 4π-pulse splits first into two sequential pulses, of which the second has a smaller area, then the leading pulse is also split. By comparing the amplitudes and delays of these subpulses with those of pulses resulting from transformation of a single 2π-pulse, we conclude that at smaller τ_0 and C, a delayed π-pulse produced by splitting one of the 2π-pulses separates out. As τ_0 increases, this delayed pulse moves toward ''larger'' times and is no longer visible in the range of delays under consideration. In the case of a 4π-pulse, its separation into two subpulses becomes appreciable on the leading section of the curve of the field amplitude versus time. The total area of these pulses is approximately 3π. The second subpulse at higher C runs ahead of a negative ''half-wave'' typical of π-pulses.

We can derive the delays of separate subpulses from calculations of time dependencies of the microcavity ''response.'' For comparison, we can take estimates of delay times $\Delta\tau$ obtained by generalizing to the case $\tau_0\neq 0$ an analytic expression for $\Delta\tau$ in a thin film for an optical pulse tuned to the spectral line peak:[18]

$$\Delta\tau=\tau_0+(\tau_0+\tilde{\delta})\ln[(\tilde{\delta}^{-1}+1+C)/(\tilde{\delta}^{-1}+1-C)]. \qquad (14)$$

The parameter τ_0 is introduced because the delayed pulse is additionally shifted and broadened in the cavity by comparison with the soliton-like response of the film. Note that an unshifted, inhomogeneously broadened spectral line shape was used in deriving Eq. (14). However, no logarithmic divergence like that in Eq. (14) as $C\to\tilde{\delta}^{-1}+1$ can be seen in our calculations. The divergence in Eq. (14) is eliminated when the optical frequency is offset from the peak of the inhomogeneously broadened line. Therefore, it is probable that a cavity, unlike a thin film, leads to a distinctive shift of an inhomogeneously broadened atomic spectral line. This feature can occur because of coupling between two resonant systems (the atoms and cavity) via the optical pulse field, and it shows up in the population function $n(\tau,x)$ of atomic
levels.

Calculations of this characteristics of the system for various cases are shown in the insets to Figs. 2 and 3. The curves clearly indicate that inversion population spectra in this problem have effective minima at $x\sim 1$. This feature, with due account of the symmetry of atomic levels with respect to variable x, is a manifestation of the ''hole burning'' effect in spectra of inhomogeneously broadened resonant levels. The typical spectral hole amplitude depends on both the photon lifetime τ_0 in the cavity (Fig. 2) and the incident pulse duration $\tilde{\delta}$ (Fig. 3), and is of order of $\tilde{\delta}^{-1}$.

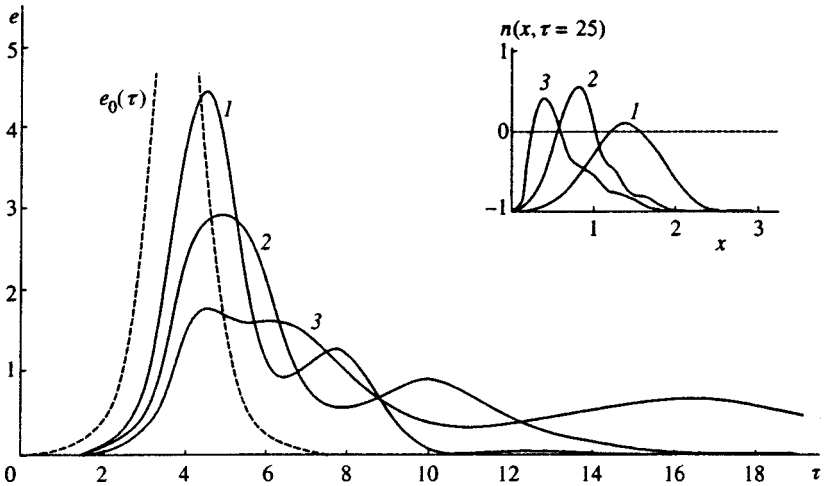


FIG. 2. Time dependencies of the microcavity response amplitude $e(\tau)$ to an incident 4π-pulse (dashed curve) of duration $\tilde{\delta}=0.5$ with cavity parameters $C=2$, $\tau_0=1$, 2, 4, and population distributions $n(x,\tau=25)$ of resonant atoms (solid lines labeled by *1, 2,* and *3*).

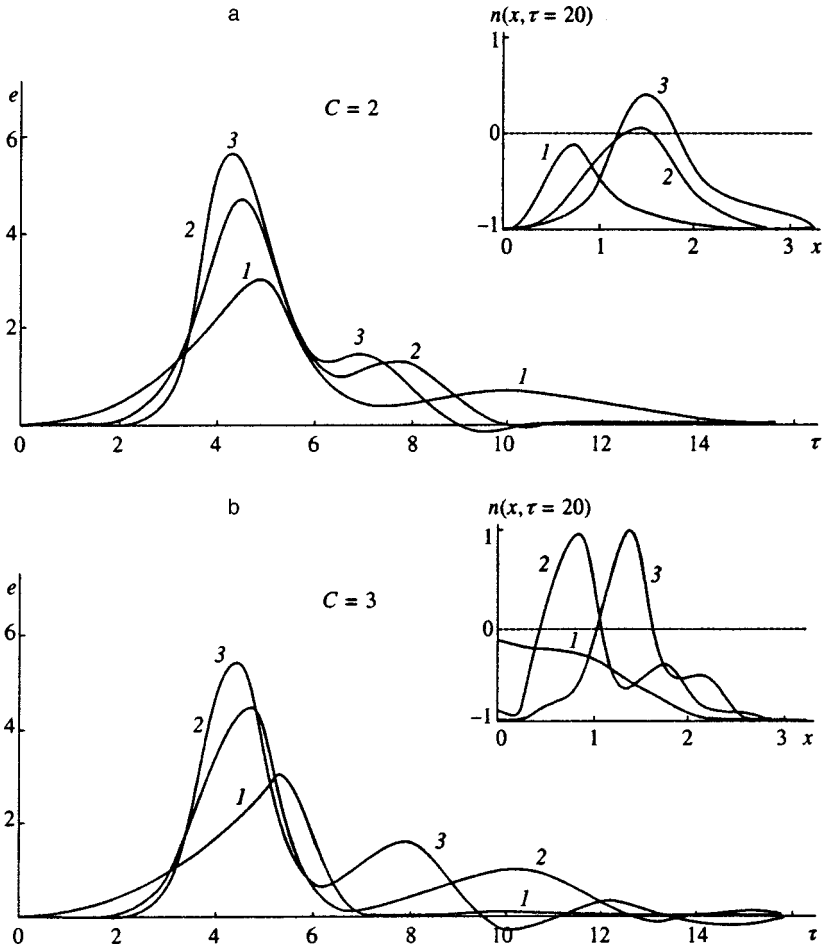


FIG. 3. Time dependencies of the microcavity response amplitude $e(\tau)$ to 4π-pulses of various lengths $\tilde{\delta}=1.0$, 0.5, and 0.33 (curves *1*, *2*, and *3*) with cavity parameters $\tau_0=1.0$ and (a) $C=2$, (b) $C=3$, and population distributions $n(x,\tau=20)$ of resonant atoms.

At certain values of the cavity and incident pulse parameters, positive population inversions $n(\tau,x)$ of order unity takes place (see curves for $C=3$, $\tau_0=1$, and $\tilde{\delta}=0.5$, 0.33 in Fig. 3), which indicate atomic transitions to a new excited state. However, at larger optical pulse duration ($\tilde{\delta}=1$ in Fig. 3) a de-excitation occurs, but the peak on the initial section of the outgoing signal is sharpened. These features of temporal and spectral characteristics of the microcavity indicate the presence of the superradiation effect in the cavity.

Another example of nonlinear optical effects in the cavity is self-induced transparency accompanied by splitting of the leading part of a transmitted pulse into 4π-pulses. The characteristic splitting of pulses takes place with significantly augmented photon lifetimes in the cavity (Fig. 2).

This effect is somewhat nebulous, because the shift and broadening of the pulse with time are small. Nonetheless, our calculations allow us to identify the linear part of the delay time $\Delta\tau$ as a function of τ_0 in the cavity and the incident pulse duration δ, which is similar to Δt in an extended medium.

Thus, the splitting of soliton-like pulses with area $\theta=\pi m$, $m=2,3,\ldots$, into a sequence of 2π- and π-pulses after transmission through a thin-film microcavity is caused by two effects of self-induced transparency typical of thin films and extended media. The main specific feature of these processes in microcavities is the possibility of changing their characteristic times through variation of the photon lifetime in the cavity.

4. PHOTON ECHO GENERATED BY A SEQUENCE OF ULTRASHORT LASER PULSES IN A MICROCAVITY

In analyzing the interaction between sequences of several ultrashort laser pulses and microcavities with a view to retaining arbitrary phase relations for incident waves, it is more convenient to use, instead of the representation of the resonant dipole moment of an isolated atom in the form of the real and imaginary parts, the complex representation

$$P_\varepsilon^+=(U_\varepsilon^+ + iV_\varepsilon^+)\exp(i\varphi),$$

where $\varphi=\varphi_0$ in the absence of phase modulation and φ_0 is the incident wave phase.

Note that in calculating the dipole moment of resonant atoms averaged over frequencies within the inhomogeneously broadened line, since T_2^* is the shortest of all characteristic times of the problem [see Eq. (1)], we can bring the function $G(\varepsilon)$ outside the integrand:

$$\langle P_\varepsilon(t)\rangle=\int_{-\infty}^{\infty} d\varepsilon\, G(\varepsilon)p_\varepsilon(t)\approx G(0)\int_{-\infty}^{\infty} d\varepsilon\, p_\varepsilon(t). \quad (15)$$

Note also that for simplicity of notation, the superscript + will be omitted in all subsequent formulas.

The symmetry of the spectral functions with respect to the offset leads to the identity

$$\int_{-\infty}^{\infty} U_{\varepsilon}(t)dt \equiv 0, \tag{16}$$

and, finally, to the formula

$$\langle P_{\varepsilon}(t)\rangle = iG(0)\int_{-\infty}^{\infty} d\varepsilon\, V_{\varepsilon}(t). \tag{17}$$

The latter property allows us to neglect the phase modulation even when the spectral function $G(\varepsilon)$ is arbitrary.

The solution for the resonant dipole moment of an isolated atom has the form

$$P_{\varepsilon}(t) = i\frac{|d_{12}|^2}{\hbar}\int_{-\infty}^{t} E_{+}(t')n_{\varepsilon}(t')\cos[\varepsilon(t-t')]dt', \tag{18}$$

and in the small-area approximation for the first ultrashort pulse of duration δ_1 we obtain

$$P_{\varepsilon}(\delta_1) = -i\frac{|d_{12}|^2}{\hbar}E^{F}_{+1}(\varepsilon), \tag{19}$$

where $E^{F}_{+1}(\varepsilon)$ is the Fourier transform of the function describing the slow envelope of the first incident optical pulse.

The subsequent procedure for calculating the polarization in the microcavity is similar to that in the case of volume excitation of a resonant medium.[19] At the onset t_2 of the second pulse, one can in fact derive from Eqs. (18) and (19) in the approximation $\theta < 1$ (Ref. 19)

$$P_{\varepsilon}(t_2) = -i\frac{|d_{12}|^2}{\hbar}E^{F}_{+1}(\varepsilon)\exp(i\varepsilon\tau_1), \tag{20}$$

which takes into account the fact that the optical pulse duration δ_1 is much less than the separation τ_1 between incident laser pulses.

Similarly, the expression for the dipole moment of a resonant atom upon termination of the second incident pulse has the form

$$P_{\varepsilon}(t_2+\delta_2) = i\frac{|d_{12}|^4}{4\hbar^3}(R_2+1)E^{F2}_{+2}(\varepsilon)E^{F*}_{+1}(\varepsilon)\exp(-i\varepsilon\tau_1), \tag{21}$$

and we obtain the dipole moment of the resonant atom averaged over the inhomogeneously broadened spectral line:

$$\langle P_{\varepsilon}(t)\rangle = i\frac{|d_{12}|^4}{4\hbar^3}(R_2+1)\cdot 2\pi G(0) \times \int_{-\infty}^{\infty}\frac{d\varepsilon}{2\pi}E^{F2}_{+2}(\varepsilon)E^{F*}_{+1}(\varepsilon)\exp[i\varepsilon(t-2\tau_1)]. \tag{22}$$

Then the following equation applies to the photon echo field amplitude at the corresponding time of onset:

$$\frac{\tilde{\tau}_c}{2}\dot{E}_{+} + E_{+} = CF(t), \tag{23}$$

where C is still determined by Eq. (6) and $F(t)$ is the temporal behavior of the response signal (two-pulse photon echo):

$$F(t) = \frac{|d_{12}|^2}{2\hbar^3}(R_2+1)\int_{-\infty}^{\infty}\frac{d\varepsilon}{2\pi}E^{F2}_{+2}(\varepsilon) \times E^{F*}_{+1}(\varepsilon)\exp[i\varepsilon(t-2\tau_1)]. \tag{24}$$

5. COMPARISON WITH A BULK RESONANT MEDIUM

It follows from the expression for $F(t)$, which describes the temporal behavior of a signal generated in a microcavity, that the shapes of photon echo signals in the approximation $\delta \gg \tilde{\tau}_c/2(1+C)$ are very similar to those generated in bulk resonant media. Depending on the ratio of the durations of incident laser pulses in Eq. (24), function $F(t)$ is, in the general case, either a convolution of the original laser pump fields ($\delta_1 \sim \delta_2$) or a time-reversed image of the first pump pulse ($\delta_1 \ll \delta_2$).

We now define the dynamic efficiency to be the ratio of the intensity of the resonant response at the moment it is generated, $t \sim 2\tau_1$, to the pump laser pulse intensity. From Eqs. (13) and (24) in the approximation $\delta \gg \tilde{\tau}_c/2(1+C)$ we obtain the following expressions for the dynamic efficiency of microcavities operating in the transmission (d) and reflection (r) modes:

$$\eta_d \cong \frac{(1-R_1)^3(1-R_2)(1+R_2)^2}{4(1-\sqrt{R_1R_2})^6(1+C)^6}C^2\theta^4, \tag{25}$$

$$\eta_r \cong \frac{(1-R_1)^4R_2(1+R_2)^2}{4(1-\sqrt{R_1R_2})^6(1+C)^6}C^2\theta^4, \tag{26}$$

where the characteristic parameter θ is, to order of magnitude, the ''area'' of the second pump laser pulse.

It is interesting to compare these results with calculations for a bulk resonant medium. The photon echo amplitude in that case is given by[19]

$$E_e \simeq \frac{\alpha_0 L}{4}\frac{|d_{12}|^2}{\hbar^2}\int_{-\infty}^{\infty}\frac{d\varepsilon}{2\pi}E^2_2(\varepsilon) \times E^{*}_{1}(\varepsilon)\exp[-i\varepsilon(t-2\tau_1)]. \tag{27}$$

Then the conversion efficiency in the bulk medium is, to order of magnitude,

$$\eta_V \simeq \frac{(\alpha_0 L)^2}{16}\theta^4, \tag{28}$$

where θ also characterizes the ''area'' of the second pump pulse. The relative excitation efficiency is then given by

$$\frac{\eta_d}{\eta_V} \simeq \frac{4(1-R_1)^3(1-R_2)(1+R_2)^2C^2}{(1-\sqrt{R_1R_2})^6(1+C)^6(\alpha_0 L)^2}, \qquad \frac{\eta_r}{\eta_V} \simeq \frac{4(1-R_1)^4R_2(1+R_2)^2C^2}{(1-\sqrt{R_1R_2})^6(1+C)^6(\alpha_0 L)^2}. \tag{29}$$

Note that Eq. (28) can be derived from Eqs. (25) and (26) by taking $R_{1,2} \ll 1$. Obviously, the microcavity then in no

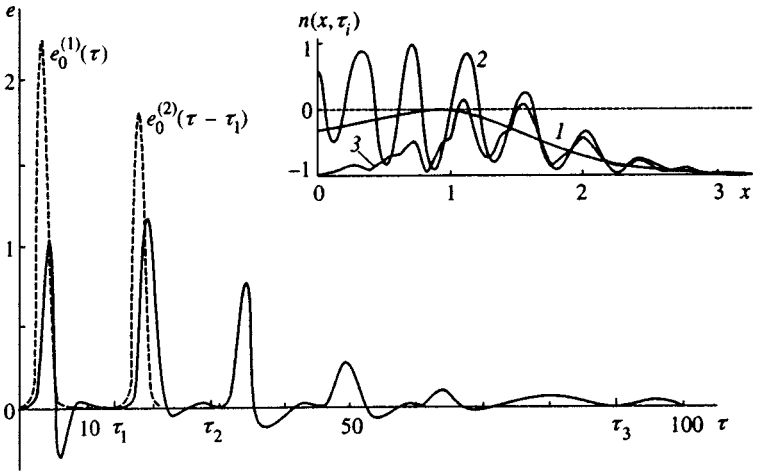


FIG. 4. Response of a microcavity $e(\tau)$ and population distributions $n(x,\tau_i)$ at various times $\tau_1=14$, $\tau_2=28$, and $\tau_3=90$ (solid curves *1*, *2*, and *3*, respectively) generated by two incident pulses $e_0(\tau)=e_0^{(1)}(\tau)+e_0^{(2)}(\tau-\tau_1)$ (dashed line) in a microcavity with parameters $C=2.0$ and $\tau_0=0.5$.

way differs from a bulk resonant medium of length L less than the inverse absorption factor α_0 in the limit of low intensity ($\alpha_0 L \ll 1$).

Evaluating the conversion efficiencies for $R_1=R_2\approx 1$, we have

$$\frac{\eta_d}{\eta_V}=\frac{\eta_r}{\eta_V}=\eta_0\approx\frac{16C^2}{(1-R^2)(1+C)^6(\alpha_0 L)^2}.$$

Remarkably, the relative efficiency for $C<1$ is

$$\eta_0\approx\frac{16}{(1-R)^4}\gg 1,$$

and thus a microcavity is a much more efficient device for generating photon echos than the usual bulk resonant medium. Similarly, for $C>1$, the relative efficiency is

$$\eta_0\approx\frac{16(1-R)^2}{(\alpha_0 L)^6}\approx\frac{16}{C^6(1-R)^4}$$

which can also exceed unity.

6. PHOTON ECHO SIGNALS GENERATED BY A SEQUENCE OF ULTRASHORT PULSES WITH ARBITRARY AREAS

For arbitrary areas of pump laser pulses, the basic equations can be solved only by numerical integration. Previously, the interaction of isolated ultrashort laser pulses and thin microcavities was analyzed under conditions that permit the observation of bistable transmission of optical pulses of area θ_0 close to the critical value θ_0'.[12] The difference between the temporal behavior of signals generated on different branches of the output characteristic is the presence (for $\theta_0>\theta_0'$) or absence (for $\theta_0<\theta_0'$) of a delayed pulse, whose delay $\Delta\tau$ depends on both the photon lifetime τ_0 in the cavity and the nonlinearity parameter C. Under these conditions, it would be interesting to consider two ultrashort pulses transmitted through a microcavity when the photon echo signal is affected by the difference in the resonant medium state before and after the transmission of the first laser pulse, which has some area θ_0.

An important feature of the two-pulse photon echo generated in a microcavity is, in our opinion, the distinctive mechanism of inhomogeneous broadening of atomic levels in the cavity when the total area of the incident laser pulses is relatively large ($\theta_0^{(1)}+\theta_0^{(2)}>\theta_0$). The first pulse perturbs the resonant medium and produces a population inversion, $n(\tau,x)>0$, at offsets comparable to the inverse time of inhomogeneous broadening ($x=\varepsilon T_2^*\leqslant 1$). This perturbation can enhance the system response to the second laser pulse.

Solutions of Eq. (5) illustrating these features of the photon echo generated in resonant structures for $\theta_0=1.4\pi$, $\theta_0=0.9\pi$, pulse duration $\tilde{\delta}=0.57$, and delay $\tau_1=15.0$ at resonant structure parameters $\tau_0=0.5$ and $C=2.0$ are plotted in Fig. 4. The lower part of the graph shows the input intensity $e_0(\tau)$ and response signal as functions of time, and the upper part shows spectral distributions of population inversion, $n(\tau,x)$, at times $\tau_{1,2,3}=14$, 28, and 90 after the onset of the first pulse.

These curves clearly show that the first transmitted pulse perturbs atoms in the cavity over a broad spectral range about the resonant frequency. After the second pulse, distinctive oscillations in the spectrum of the population inversion with period Δx equal to the reciprocal time interval τ_1 between the pulses ($\Delta x\sim\tau^{-1}$) are observed. This effect is due to a distinctive ''spectral grating,'' and is important in photon echo generation by three incident optical pulses. In the present case, the first and strongest photon echo pulse eliminates most of the population inversion created by the first incident pulse. The subsequent damped pulses of the photon echo continue the relaxation of the perturbed resonant structure.

7. CONCLUSION

The reported results suggest that dynamic characteristics of photon echo signals generated in microcavities can be similar to those of bulk resonant media. The temporal behavior of generated optical signals can be direct replicas (stimulated photon echo) or time-reversed versions (common two-pulse photon echo) of one of the pump pulses, or they can be described by convolutions of slow envelope functions of incident pulses. The effect of multiple photon echo, which is inherent to gaseous media under excitation by small-area pulses in the form of standing waves, does not occur in mi-

crocavities containing resonant solid-state materials. This effect is typical of microcavities excited by laser pulses with arbitrary area $\theta \sim 1$. On the other hand, resonant microcavities can be more efficient than bulk media. This property is especially important for holographic systems that produce and transform two-dimensional images. It thus becomes possible to increase the dynamic range of dynamic holography, which can be used in both analog and digital image processing.

We are grateful to É. A. Manykin and V. N. Beloborodov for useful discussions of some topics discussed here. We also thank V. P. Zagonov, who suggested the approach to the numerical solution of the problem.

APPENDIX A: DETAILS OF THE NUMERICAL INTEGRATION TECHNIQUE

In selecting the technique for numerical integration of Eq. (5) with given initial conditions and an arbitrary source function $e_0(\tau)$, two significant conditions were immediately imposed on the numerical algorithms, stemming essentially from the physical model under investigation. On the one hand, the existence of regions where solutions of the original nonlinear equations are unstable requires a stable difference scheme for solving these equations over a wide time interval. On the other, the integral nature of the calculated response signal $e(\tau)$ and parameters θ (θ_0) requires that solutions $e(\tau)$, $v(\tau,x)$, $u(\tau,x)$, and $n(\tau,x)$ be highly accurate. Moreover, application of numerical techniques to the investigation of physical characteristics of the system imposed significant constraints on the computational complexity and CPU time.

A preliminary analysis of available numerical techniques led us to select the quasianalytic interpolation method,[20] which results in a simple mixed scheme for development of a computational procedure with guaranteed accuracy. This method allowed us to take into account specific features of the equations due to the integral nature of its nonlinearity. The original set of equations is, in a sense, quasilinear. This property of the problem, which could be solved by the quasianalytic interpolation technique, enabled us to carry out analytic calculations to the greatest possible extent, and thereby to reduce the number of operations typical of standard algorithms for solving systems of nonlinear differential equations.

The essence of the present technique is that we seek a solution of the Cauchy problem by linearly interpolating the functions that enter into the set of equations on the mesh width. Solving Eq. (5) in this way, the equation derived from the first line of Eq. (5) to recalculate the field e_{i-1} at the $(i-1)$th level in terms of the values at the ith level has the form

$$e_i = e_{i-1}p + (h/\tau_0)(C_1F_{i-1} + C_2F_i),$$

where $i = 1,2,3,\ldots,I$; h is the time increment, $F_i = e_{0i} - CY_i$, $e_0 = e(\tau_0) = 0$, $p = \exp(h/\tau_0)$; $C_1 = (\Delta - p)\tau_0/h$, $C_2 = \Delta - C_1$, and $\Delta = (1-p)\tau_0/h$ are the constants of the problem.

It follows from the above description that the numerical scheme is Euler's combined scheme. The absolute values of uncertainties in the explicit and implicit parts are commensurate [proportional to $(h/\tau_0)^2$], but have opposite signs. Therefore the errors of this scheme for $h < \tau_0$, $C_1 \sim C_2$ can be smaller than those of the Runge–Kutta method with a similar approximation scheme. The equations relating $(v,u,n)_{i,m}$ to $(v,u,n)_{i-1,m}$ (for the mth spectral component $x = mh_x$, where $m = 1,2,\ldots,M$ and h_x is the frequency increment) can be derived from the last three equations in (5). After integrating piece-wise interpolated functions and solving the system of algebraic equations analytically, one has

$$v_{i+1}(x) = \left\{ v_i(x) - \frac{hx}{2}u_i(x) - \frac{n_i(x)h}{6}(2e_i + e_{i+1}) \right\} \frac{1}{Z},$$

$$v_{i+1}(x) = v_{i+1}(x) + \frac{hx}{2}[v_{i+1}(x) + v_i(x)],$$

$$n_{i+1}(x) = n_i(x) + \frac{h}{6}[v_{i+1}(x)(2e_{i+1} + e_i) + v_i(x)(2e_i + e_{i+1})],$$

$$Z = 1 + \left(\frac{hx}{2}\right)^2 + \frac{h}{6}(2e_i + e_{i+1})^2.$$

The integral $Y(\tau)$, based on the resulting values of $v_{i,m} = v(\tau_i, x_m)$ (see Eq. (7)), can be calculated using simple methods like the trapezoidal rule, since oscillations in the spectrum of function $v(\tau,x)$ are effectively ''damped'' in the integral by the exponential factor $g(x)$.

In solving the resulting difference problem, we must calculate $Y_i = Y(\tau_i)$ at the upper ith level of each layer. To this end, the algorithm iteratively improves the calculated value of Y_i^l ($l = 1,2,\ldots,L$), which was initialized with the value $Y_i^0 = Y_{i-1}^l$. Convergence of subsequent calculations of Y_i^{l+1} based on Y_i^l is determined by the smallness of the mesh width h in comparison with the dimensionless parameter τ_0. As follows from computer simulations, given a mesh width that ensures acceptable accuracy of calculated functions over a several-decade range of the dimensionless variable τ, $L = 2$ or 3 iterations typically suffice.

The technique for calculating the response $e(\tau)$ of a microcavity to external field was tested by analyzing several cases when analytic solutions could also be found. In particular, calculations were performed for an empty Fabry–Perot cavity ($C = 0$), a small area of the input pulse, and transmission of a ''soliton'' through a thin film when the photon lifetime is vanishingly small ($\tau_0 \to 0$).[18] Numerical calculations on embedded uniform meshes demonstrated uniform convergence of the numerical errors in $e(\tau)$ with decreasing step size h, and stability of calculations against errors in the initial values due to a finite shift in the initial time from which the equations were integrated.

Note that in a typical calculation for a ''soliton-like'' 2π-pulse $e_0(\tau)$ ($\tilde{\delta} = 0.5$) in a cavity with $\tau_0 = 0.1$ and nonlinearity parameter $C = 1$ over time interval $\tau = [-3,10]$, the absolute error of calculated $e(\tau)$ with step size $h = 0.05$ and $h_x = 0.1$ for $L = 3$ is estimated to be below 0.03 (less than 3% of the peak function value). The CPU time required by the code developed for the PC AT-286 was several tens of sec-

onds. The accuracy of a similar technique developed for solving differential equations is described in detail elsewhere.[21]

*)E-mail: administrator@mail.ivvs.ru

[1] L. A. Lugiato and C. Oldano, Phys. Rev. A **37**, 3897 (1988).
[2] L. A. Lugiato and L. M. Narducci, Z. Phys. B **71**, 129 (1988).
[3] M. Haelterman, G. Vitrant, and R. Reinisch, J. Opt. Soc. Am. B **7**, 1309; 1319 (1990).
[4] K. C. Ho and G. Indebetouw, Appl. Opt. **30**, 2437 (1991).
[5] Yu. A. Logvin and A. M. Samson, Zh. Éksp. Teor. Fiz. **102**, 472 (1992) [Sov. Phys. JETP **75**, 250 (1992)]; Yu. A. Logvin and A. M. Samson, Opt. Commun. **84**, 99 (1991).
[6] T. Perchel and F. Lederer, Phys. Rev. B **46**, 7632 (1992).
[7] N. A. Ivanova, A. S. Rubanov, and A. L. Tolstik, Izv. Ross. Akad. Nauk, Ser. Fiz. **57**, 149 (1993); N. A. Ivanova, A. S. Rubanov, and A. L. Tolstik, Proc. SPIE **1807**, 157 (1993).
[8] É. Vanagas, Pis'ma Zh. Tekh. Fiz. **19**(7), 25 (1993) [Tech. Phys. Lett. **19**, 408 (1993)].
[9] L. A. Lugiato, Phys. Rev. A **49**, 2049 (1994).
[10] V. Pierro and I. M. Pinto, Phys. Lett. A **185**, 14 (1994).
[11] S. M. Zakharov, Zh. Éksp. Teor. Fiz. **108**, 829 (1995) [Sov. Phys. JETP **81**, 452 (1995)].
[12] V. A. Goryachev and S. M. Zakharov, Kvantovaya Élektronika **24**, 251 (1997).
[13] J.-L. Le Gouet and P. R. Berman, Phys. Rev. A **20**, 1105 (1979).
[14] A. I. Alekseev, A. M. Basharov, and V. N. Beloborodov, Zh. Éksp. Teor. Fiz. **79**, 787 (1980) [Sov. Phys. JETP **52**, 401 (1980)].
[15] A. I. Alekseev and A. M. Basharov, Zh. Eksp. Teor. Fiz. **80**, 1361 (1981) [Sov. Phys. JETP **53**, 694 (1981)]; Kvantovaya Élektronika **8**, 182 (1981) [Sov. J. Quantum Electron. **11**, 106 (1981)].
[16] S. N. Bagaev, A. S. Dychkov, and V. P. Chebotaev, JETP Lett. **26**, 442 (1977).
[17] S. L. McCall and E. L. Hahn, Phys. Rev. **183**, 457 (1969).
[18] V. I. Rupasov and V. I. Yudson, Zh. Éksp. Teor. Fiz. **93**, 494 (1987) [Sov. Phys. JETP **66**, 282 (1987)].
[19] S. M. Zakharov and É. A. Manykin, Zh. Éksp. Teor. Fiz. **91**, 1289 (1986) [Sov. Phys. JETP **64**, 761 (1986)].
[20] R. P. Fedorenko, *Introduction to Computational Physics* [in Russian], Nauka, Moscow (1994).
[21] G. D. Vasil'kov and V. P. Zagonov, *Differentsial'nye Uravneniya* **19**, 1146 (1983).

Translation provided by the Russian Editorial office

Self-induced transparency for ultrashort pulses propagating in a multilevel quantum medium

A. Yu. Parkhomenko

Tomsk State University, 634050 Tomsk, Russia

S. V. Sazonov

Kaliningrad State Technical University, 236000 Kaliningrad, Russia

(Submitted 31 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1595–1617 (November 1998)

We study theoretically the phenomenon of self-induced transparency for ultrashort pulses (videopulses) propagating in a multilevel quantum medium under conditions in which the method of slowly varying amplitudes and phases breaks down and the pulse spectrum substantially overlaps the quantum transitions. A special class of transitions with one common level is considered. We find that the dynamics of the videopulses in such a medium is described by a double sine-Gordon equation. We establish the conditions under which steady-state traveling 0π-, 2π-, and 4π-videosolitons are formed. It is found that 4π-videosolitons can propagate both in equilibrium media and in some nonequilibrium media, while 0π-videosolitons can propagate only in nonequilibrium media. We study amplification processes in highly nonequilibrium systems and show that, depending on the initial state of the medium, 2π- and $q\pi$-pulses $(0<q<1)$ with increasing amplitudes may be formed. We conclude that the amplification of an ultrashort pulse occurs due to an increase in the photon number density and to an increase in the frequency of each photon. Finally, we study the possibility of an electromagnetic autosoliton being formed in a nonequilibrium dissipative medium.
 [S1063-7761(98)00411-9]

1. INTRODUCTION

The phenomenon of self-induced transparency, discovered by McCall and Hahn[1] in 1967, served as a powerful impetus to develop nonlinear resonance optics for nano- and picosecond pulses. In view of the resonant nature of the interaction, the approximation of a two-level medium is commonly used in theoretical studies of self-induced transparency. Moreover, in solving the wave and material equations the approximation of slowly varying amplitudes and phases became very popular.[2] In all fairness it should be said that in the 1970s several theoretical papers appeared that did not use this approximation.[3–7] But at the time such papers were more of a purely theoretical nature, since the femtosecond barrier for pulse lengths had yet to be overcome. The possibility of actually generating ultrashort (femtosecond) pulses in laboratories, each pulse ''holding'' about one period of optical oscillations,[8–10] stimulated further theoretical studies of self-induced transparency that did not use the approximation of slowly varying amplitudes and phases.[11–13]

Since an ultrashort pulse is not a monochromatic signal, we cannot use the idea of a pulse envelope, so that there is no way in which the approximation of slowly varying amplitudes and phases can be used. In some cases the approximation of a two-level medium is still valid. For instance, this is the case if the pair of levels under investigation is far from the other quantum levels of the medium. Theoretical studies of the propagation of ultrashort pulses in two-level quantum media have shown that electromagnetic videosolitons, i.e., i.e. solitons with no high-frequency component, can form in such media.

In particular, when the spectrum of an ultrashort pulse substantially overlaps the quantum transition under consideration, the electric field of the pulse (and not of the envelope) obeys the sine-Gordon equation.[14–16] It is a well-known fact that in the case of resonant self-induced transparency the envelope of the electric field of a monochromatic pulse also obeys this equation. In view of this it can be said that the soliton modes of propagation of broadband ultrashort pulses in highly excited absorbing media belong to the phenomenon of self-induced transparency by analogy with the corresponding effect for monochromatic signals. The spectrum of an ultrashort pulse is extremely broad, so that more than often several quantum transitions interact with it simultaneously.

The present paper is devoted to a theoretical study of self-induced transparency when an ultrashort pulse propagates in multilevel quantum media and the pulse spectrum overlaps all allowed transitions.

2. SOLUTION OF THE SYSTEM OF MATERIAL EQUATIONS

We write the system of material equations for the elements L_{mk} $(m,k=1,2,\ldots,N)$ of the density matrix $\hat{L}$ of an N-level system in the form

$$\frac{\partial L_{mk}}{\partial t}=i\omega_{mk}L_{mk}+i[\hat{L},\hat{A}]_{mk}, \tag{1}$$

where $\omega_{mk}=-\omega_{km}$ is the frequency of the transition between the mth and kth levels.

The transition matrix $\hat{A}$ is given by the formula

$$\hat{A}=\frac{1}{\hbar}\begin{pmatrix} 0 & \mathbf{d}_{12}\cdot\mathbf{E} & \mathbf{d}_{13}\cdot\mathbf{E} & \cdots & \mathbf{d}_{1N}\cdot\mathbf{E} \\ \mathbf{d}_{21}\cdot\mathbf{E} & 0 & \mathbf{d}_{23}\cdot\mathbf{E} & \cdots & \mathbf{d}_{2N}\cdot\mathbf{E} \\ \vdots & \ddots & & \vdots & \\ \vdots & & & 0 & \mathbf{d}_{N-1,N}\cdot\mathbf{E} \\ \mathbf{d}_{N1}\cdot\mathbf{E} & \mathbf{d}_{N2}\cdot\mathbf{E} & \cdots & \mathbf{d}_{N,N-1}\cdot\mathbf{E} & 0 \end{pmatrix}. \tag{2}$$

Here $\hbar$ is Planck's constant, $\mathbf{E}$ is the electric field of the ultrashort pulse, $\mathbf{d}_{mk}$ $(k,m=1,\ldots,N)$ are the matrix elements of the dipole moments of the allowed quantum transitions. In (1) there is no summation over the repeated indices m and k.

We will now assume that the initial pulse applied to the medium is linearly polarized. Obviously, in the course of the propagation of the pulse in the medium, the pulse's polarization plane does not change orientation. In this case the vector of the induced dipole moment is collinear with the electric field of the ultrashort pulse. Then the matrix elements $\mathbf{d}_{mk}$ and their projections on the direction of $\mathbf{E}$ are time−independent. This situation corresponds to $\pi-$transitions, in which the magnetic quantum number does not change, the matrix elements $\mathbf{d}_{mk}$ are real, and due to its hermiticity the transition matrix $\hat{A}$ is symmetric $(A_{mk}=A_{km})$. Bearing all this in mind, we can go from the vector quantities $\mathbf{E}$ and $\mathbf{d}_{mk}$ to the scalar quantities E and d_{mk}.

To solve the self-consistent problem of the interaction of an ultrashort pulse and the medium we combine the system of material equations (1) with the Maxwell equation, which in view of the collinearity of the electric field of this pulse and the induced dipole moment can be written in scalar form:

$$\Delta E-\frac{1}{c^2}\frac{\partial^2 E}{\partial t^2}-\frac{\gamma}{c^2}\frac{\partial E}{\partial t}=\frac{4\pi}{c^2}\frac{\partial^2 P}{\partial t^2}. \tag{3}$$

Here Δ is the Laplacian operator, c is the speed of light, and P is the polarization of the medium consisting of N-level atoms interacting with the ultrashort pulse and is given by the formula

$$P=\sum_{m\neq k} d_{mk}L_{mk}. \tag{4}$$

The phenomenological constant γ allows for dissipation related to losses due to the effective conductivity of the medium and scattering by quantum levels located far from those considered here.[17,18]

The condition that the spectrum of the ultrashort pulse overlaps all the allowed transitions can formally be written as an inequality,[14–16]

$$\omega_{mk}\tau_p\ll 1, \tag{5}$$

valid for all m and k. Here τ_p is the characteristic time scale associated with the ultrashort pulse (1) by the method of successive approximations in the parameters ω_{mk} $(m,k=1,\ldots,N)$. Assuming that in the zeroth approximation $\omega_{mk}=0$ for all m and k, we write the system (1) as follows:

$$\frac{\partial\hat{L}^{(0)}}{\partial t}=i[\hat{L}^{(0)},\hat{A}]. \tag{6}$$

The upper index in parentheses corresponds to the order of approximation in the parameter $\omega_{mk}\tau_p$.

Generally speaking, the analysis of (6) is extremely complicated.[19] However, under the conditions adopted in this paper (time independence and the fact that the matrix elements d_{mk} are real), the matrix $\hat{A}$, obviously, commutes with the integral of $\hat{A}$ with respect to time, i.e.,

$$\left[\hat{A}(t),\int_{-\infty}^{t}\hat{A}(t')\,dt'\right]=0. \tag{7}$$

In fact, we can write (7) as

$$\sum_k\left(\frac{\partial\theta_{mk}}{\partial t}\theta_{kn}-\theta_{mk}\frac{\partial\theta_{kn}}{\partial t}\right)=0,$$

where

$$\theta_{mk}=2\int_{-\infty}^{t}A_{mk}\,dt'=\frac{2d_{mk}}{\hbar}\int_{-\infty}^{t}E\,dt'.$$

This clearly shows that each term in the above sum is zero.

Since condition (7) is sure to be met (see Ref. 19), the solution of (6) can be written in symbolic form:

$$\hat{L}(t)=\exp\left(-\frac{i\hat{\theta}}{2}\right)\hat{L}(-\infty)\exp\frac{i\hat{\theta}}{2}, \tag{8}$$

where $\hat{\theta}=2\int_{-\infty}^{t}\hat{A}(t')\,dt'$, and $\hat{L}(-\infty)$ is the density matrix of the medium before the ultrashort pulse was applied,

$$\hat{L}(-\infty)=\begin{pmatrix} \rho_{11} & 0 & \cdots & 0 \\ 0 & \rho_{22} & & \\ \vdots & & \ddots & \\ 0 & & & \rho_{NN} \end{pmatrix}. \tag{9}$$

To make our reasoning more specific and simple, we assume that the only allowed transitions are those that pass through a common level (Fig. 1). Suppose that the number of this level is j. Then the only elements in the symmetric matrices $\hat{A}$ and $\hat{\theta}$ that are finite are those of the jth row and jth column, with the exception of the diagonal elements A_{jj} and θ_{jj}. This model is realized in many physical problems. At $N=3$ and $N=4$ (N is the overall number of quantum levels involved in the problem), the model describes the optical properties of wide-gap insulators ($j=2$ and $j=1$ for $N=3$, and $j=3$ for $N=4$) (Refs. 20–22). For instance, for $N=4$ the first level corresponds to the electron states of the valence band, the second level corresponds to the same electron states and the excited states of optical oscillations of the crystal lattice, and the third and fourth levels correspond to the subbands of the conduction band, the subbands being strongly and weakly coupled to the valence band, respectively.[22] For $j=1$ the model describes indirect band-to-band transitions in multivalley semiconductors[23] and electric dipole transitions between the Zeeman sublevels in the

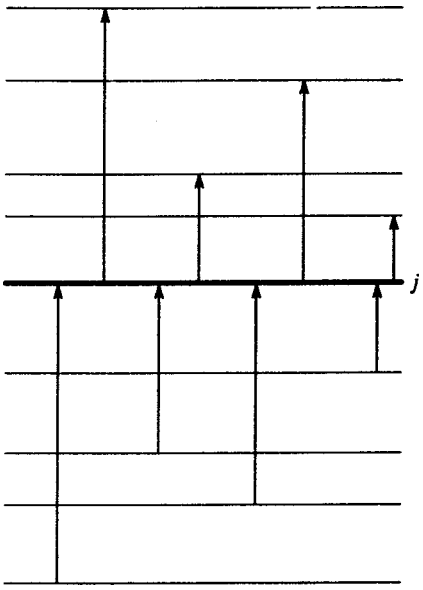


FIG. 1. Diagram of allowed quantum transitions have one common level j (the heavy horizontal line). The overall number of levels is arbitrary.

Voigt geometry, where the polarization plane of the pulse and the direction of pulse propagation are perpendicular the external magnetic field.[24] Another well-known example is the Cu_2O crystal, whose excitonic spectrum corresponds to a distinct series of transitions with $j=1$ (Ref. 25). If $j=N$ holds we have, e.g., models of media capable of phosphorescing.[26] In this case the ground state is thermodynamically stable, while all intermediate states are metastable. Phosphorescence is observed at frequencies corresponding to the transitions $N-1\to 1$, $N-2\to 1$, $\ldots$, $2\to 1$.

Moreover, an N-level quantum system with transitions involving a common level is of interest by itself as one of the simplest generalizations of a two-level system under conditions of spectral overlap of the quantum levels by the field of the ultrashort pulse.

Clearly, the matrix $\hat{\theta}$, in which only the elements of the jth row and the jth column, with the exception of θ_{jj}, are nonzero, obeys the following relation:

$$\hat{\theta}^{2n}=\hat{\theta}^2\left(\sum_{i=1}^{N}\theta_{ij}^2\right)^{n-1},\quad \hat{\theta}^{2n-1}=\hat{\theta}\left(\sum_{i=1}^{N}\theta_{ij}^2\right)^{n-1}, \tag{10}$$

where $n=1,2,3,\ldots$.

Expanding $\exp(\pm i\hat{\theta}/2)$ in a Maclaurin series in powers of $(\pm i\hat{\theta})$ and using (10) to sum the series, we find

$$\exp\left(\pm\frac{i\hat{\theta}}{2}\right)=\hat{I}-\frac{\hat{\theta}^2}{\theta^2}\left(1-\cos\frac{\theta}{2}\right)\pm i\frac{\hat{\theta}}{\theta}\sin\frac{\theta}{2}, \tag{11}$$

where

$$\theta=2\left(\sum_{i=1}^{N}\theta_{ij}^2\right)^{1/2}=\frac{2D_j}{\hbar}\int_{-\infty}^{t}E\,dt',$$

$$D_j=\left(\sum_{i=1}^{N}d_{ij}^2\right)^{1/2},$$

and $\hat{I}$ is the identity matrix.

The structure of the matrix $\hat{\theta}^2$ is as follows (no summation over j):

$$(\hat{\theta}^2)_{mn}=\theta_{mj}\theta_{jn}\quad \text{for}\quad m,n\neq j,$$

$$(\hat{\theta}^2)_{mj}=(\hat{\theta}^2)_{jm}=0\quad \text{for}\quad m\neq j,$$

$$(\hat{\theta}^2)_{jj}=\sum_{k=1}^{N}\theta_{jk}^2.$$

Substituting (11) into (8), we find the following elements of the matrix $L(t)$:

$$L_{jk}^{(0)}=i\frac{d_{jk}}{2D_j}\left(\left(\rho_{jj}-\frac{D_{j\rho}^2}{D_j^2}\right)\sin\theta+2\left(\frac{D_{j\rho}^2}{D_j^2}-\rho_{kk}\right)\sin\frac{\theta}{2}\right),\quad k\neq j, \tag{12}$$

$$L_{jj}^{(0)}=\rho_{jj}\cos^2\frac{\theta}{2}+\frac{D_{j\rho}^2}{D_j^2}\sin^2\frac{\theta}{2}, \tag{13}$$

$$L_{mk}^{(0)}=\rho_{mk}+\frac{d_{jm}d_{jk}}{D_j^2}\left(\rho_{jj}\sin^2\frac{\theta}{2}-2(\rho_{mm}+\rho_{kk})\sin^2\frac{\theta}{4}+4\frac{D_{j\rho}^2}{D_j^2}\sin^4\frac{\theta}{4}\right), \tag{14}$$

where $m,k\neq j$, and $D_{j\rho}^2=\Sigma_{k=1}^{N}d_{jk}^2\rho_{kk}$.

Setting $m=k$ in (14), we arrive at an expression for the population of the mth quantum level. As expected, Eqs. (13) and (14) obey the condition $\Sigma_{m=1}^{N}L_{mm}^{0}=1$.

In the first approximation in the parameter $\omega_{mk}\tau_p$, Eq. (1) becomes

$$\frac{\partial L_{mk}^{(1)}}{\partial t}=i\omega_{mk}L_{mk}^{(0)}+i[\hat{L}^{(0)},\hat{A}]_{mk}. \tag{15}$$

In the case at hand the polarization P can be written as follows:

$$P=n\sum_{j=1}^{N}d_{jk}(L_{jk}+L_{kj}), \tag{16}$$

where n is the density of N-level atoms.

Using (15), we write

$$\frac{\partial}{\partial t}(L_{jk}^{(1)}+L_{kj}^{(1)})=i\omega_{jk}L_{jk}^{(0)}+i\omega_{kj}L_{kj}^{(0)}+i\sum_{m=1}^{N}\{A_{jm}(L_{km}^{(0)}-L_{mk}^{(0)})+A_{mk}(L_{mj}^{(0)}-L_{jm}^{(0)})\}. \tag{17}$$

In view of the hermiticity of the density matrix, $L_{km}^{(0)}=L_{mk}^{(0)*}$ for all m and k. According to (12) and (14), the elements $L_{jm}^{(0)}$ are purely imaginary, while the elements L_{km} for $k,n\neq j$ are real. Moreover, due to the structure of the transitions, $A_{mk}=0$ for $m,k\neq j$. Hence the expression in braces in (17) vanishes and we have

$$\frac{\partial}{\partial t}(L_{jk}^{(1)}+L_{kj}^{(1)})=2i\omega_{kj}L_{jk}^{(0)}.$$

Combining this with (16) and (12), we find that

$$\frac{\partial P}{\partial t}=2in\sum_{k=1}^{N} d_{kj}\omega_{kj}L_{kj}^{(0)}$$

$$=\frac{n}{D_j}\left(D_{j\omega}(\rho_{jj}-R_j)\sin\theta\right.$$

$$\left.+2(D_{j\omega}R_j-D_{j\omega\rho})\sin\frac{\theta}{2}\right), \tag{18}$$

where

$$D_{j\omega}=\sum_{k=1}^{N} d_{kj}^2\omega_{kj}, \quad D_{j\omega\rho}=\sum_{k=1}^{N} d_{kj}^2\rho_{kk}\omega_{kj},$$

$$R_j=\frac{D_{j\rho}^2}{D_j^2}. \tag{19}$$

Substituting (18) into (3), we arrive at a double sine-Gordon equation with damping for the electric field of the ultrashort pulse:

$$\Delta\theta-\frac{1}{c^2}\frac{\partial^2\theta}{\partial t^2}-\frac{\gamma}{c^2}\frac{\partial\theta}{\partial t}=\alpha\sin\theta+\beta\sin\frac{\theta}{2}, \tag{20}$$

where

$$\alpha=\frac{8\pi n}{\hbar c^2}D_{j\omega}(\rho_{jj}-R_j), \quad \beta=\frac{16\pi n}{\hbar c^2}(D_{j\omega}R_j-D_{j\omega\rho}). \tag{21}$$

Equations (19) and (21) clearly show that in the case of a two-level system with $j=1$ we have

$$\beta=0, \quad \alpha=-\frac{16\pi n d_{21}^2\omega_{21}W_\infty}{\hbar c^2},$$

where $W_\infty=(\rho_{22}-\rho_{11})/2$ is the initial population inversion of the medium. Then at $\gamma=0$ Eq. (20) becomes the sine-Gordon equation found for this case in Refs. 14–16. We also note that $\beta=0$ also for the case of an N-level medium, provided that $j=1$ and that before the light pulse was applied to the medium only the ground level was populated, $(\rho_{11}=1)$.

If in (21) we set $N=3$ for $j=2$ and $j=3$, and $N=4$ at $j=3$, we arrive at the expressions for α and β derived in Ref. 27. We also note that the double sine-Gordon equation appears in the theory of self-induced transparency for a two-level medium with a fivefold degenerate second level when a resonant monochromatic pulse propagates in that medium (here $\beta/\alpha=1/2$) (Ref. 28). Below we use Eq. (20) to study the various modes of propagation of ultrashort pulses in a multilevel quantum medium.

3. PROPAGATION IN THE STEADY-STATE PULSE MODE

Let us suppose that a pulse is propagating along the z axis with a constant velocity v. Depending on the initial state of the medium, there are three different cases that allow for the propagation of an ultrashort pulse in steady-state modes in the absence of dissipation $(\gamma=0)$.

1. $\alpha>0$ and $\beta>0$. The solution of Eq. (20) has the form

$$\theta=4\arctan(\sqrt{1+\kappa^2}\,\mathrm{csch}\,\xi), \tag{22}$$

$$E=\frac{2\hbar}{D\tau_p}\frac{\kappa\,\mathrm{sech}\,\xi}{1+\kappa^2\mathrm{sech}^2\xi}, \tag{23}$$

where

$$\xi=\frac{t-z/v}{\tau_p}, \quad \kappa=\left(\frac{2\alpha}{\beta}\right)^{1/2}.$$

Here the relationship between the velocity v of the pulse propagation and the pulse length τ_p is given by the formula

$$\frac{1}{v^2}=\frac{1}{c^2}+\frac{\beta}{2}(1+\kappa^2)\tau_p^2. \tag{24}$$

Combining (13) and (14), we find an expression for the population of the medium:

$$L_{jj}=\rho_{jj}\left(\frac{1+\kappa^2-\sinh^2\xi}{\kappa^2+\cosh^2\xi}\right)^2+4R_j\frac{(1+\kappa^2)\sinh^2\xi}{(\kappa^2+\cosh^2\xi)^2}, \tag{25}$$

$$L_{mm}=\rho_{mm}\left(1-\frac{4d_{mj}^2}{D_j^2}\frac{1+\kappa^2}{\kappa^2+\cosh^2\xi}\right)$$

$$+\frac{4d_{mj}^2}{D_j^2}\frac{((1+\kappa^2)R_j+\rho_{jj}\sinh^2\xi)(1+\kappa^2)}{(\kappa^2+\cosh^2\xi)^2}, \quad m\neq j. \tag{26}$$

In accordance with (23), the profile of the electric field of a traveling ultrashort pulse is double-humped (Fig. 2). The hump separation $\tilde{\Delta}=2\xi^*$ is determined by the condition $\mathrm{sech}\,\xi^*=1/\kappa$, $\kappa>1$. The relative depth of the dip, $E_{\max}/E_{\min}$, is equal to $(1+\kappa^2)/2\kappa$.

Note that in contrast to classical self-induced transparency, the expression (23) is written for the electric field of the pulse rather than for the pulse's envelope. In accordance with this, here and in what follows we sometimes call an ultrashort pulse a videopulse, and solitonlike solutions of self-induced transparency for an ultrashort pulse, $2\pi n-$videopulses $(n=0,1,2)$. From (23) it follows that the area of the signal (23) is

$$\Phi\equiv\frac{2D}{\hbar}\int_{-\infty}^{\infty}E\,dt=4\pi,$$

so that we call this signal a 4π-videopulse.

The conditions $\alpha>0$ and $\beta>0$ are met, say, for a thermodynamically equilibrium medium in transitions between the levels considered here, which form a series with $j=1$. Let us analyze this case in detail. Before the pulse is applied to the medium, the quantum levels of the medium are populated according to the Boltzmann distribution

$$\rho_{kk}=\frac{1}{Z}\exp\left(-\frac{\hbar\omega_{k1}}{k_BT}\right), \quad k=1,2,\ldots,N, \tag{27}$$

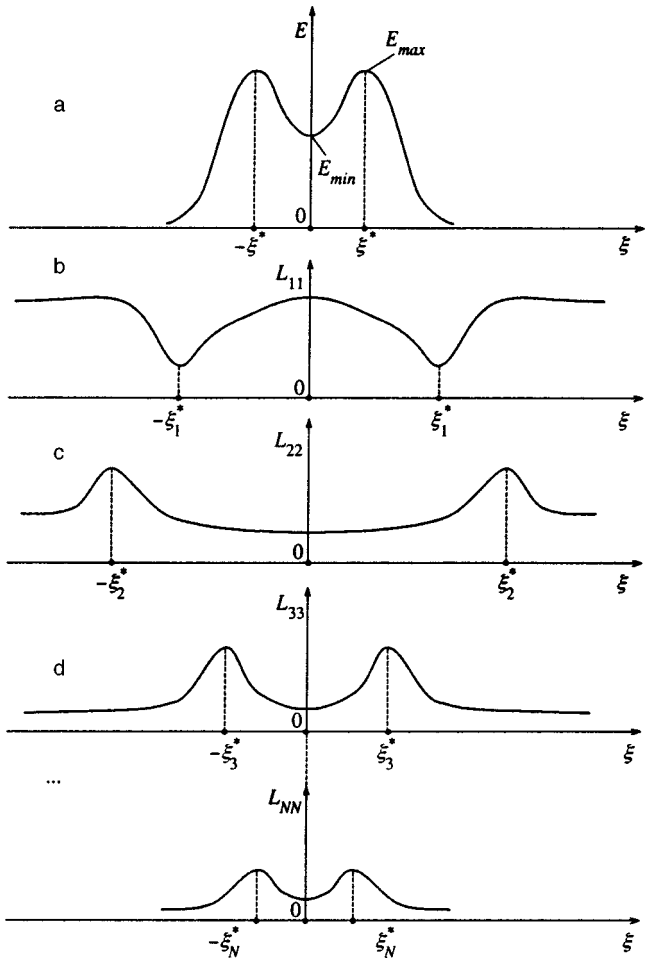


FIG. 2. Dynamics of the electric field of ultrashort pulses when they propagate in a medium in thermodynamic equilibrium for $a>0$ and $\beta>0$ (a) and the corresponding population dynamics in the quantum level (b, c, and d) with $j=1$.

where $Z=\Sigma_{m=1}^{N}\exp(-\hbar\omega_{m1}/k_BT)$ is the partition function, k_B is Boltzmann's constant, and T is the absolute temperature. Here we have the obvious inequalities $\rho_{11}>\rho_{22}>\rho_{33}>\cdots>\rho_{NN}$. As T drops to absolute zero, $\kappa\to\infty$. The humps in the pulse in this case move infinitely far from one another, and the field in the dip vanishes. Thus, the 4π-videopulse splits into two 2π-videopulses infinitely distant from each other, with each corresponding to the case where $\beta=0$. Then Eq. (20) becomes a sine-Gordon equation whose 2π-videopulse and the corresponding expressions for the medium populations are

$$\theta=4\arctan\exp\frac{t-z/v}{\tau_p}, \tag{28}$$

$$\frac{1}{v^2}=\frac{1}{c^2}+\alpha\tau_p^2, \tag{29}$$

$$E=\frac{\hbar}{D\tau_p}\operatorname{sech}\frac{t-z/v}{\tau_p}, \tag{30}$$

$$L_{11}=\tanh^2\frac{t-z/v}{\tau_p}, \tag{31}$$

$$L_{mm}=\frac{d_{mj}^2}{D_j^2}\operatorname{sech}^2\frac{t-z/v}{\tau_p}, \quad m=2,3,\ldots,N, \tag{32}$$

where τ_p is the length of the ultrashort pulse (this quantity is a free parameter of the solution).

As the pulse (30) propagates, the atoms first go from the ground state to excited states and then return to the ground state (see Eqs. (31) and (32)). In this case the dynamics of the ultrashort pulse is the simplest and differs little from that for a two-level system.[14,16] The main difference is that in our case the ultrashort pulse acts on the medium in such a way that many excited states become populated (see Eq. (32)), while in a two-level system only one excited state becomes populated.

When the upper levels are initially populated ($T\neq0$), $\beta\neq0$. This, as Eq. (30) implies, reduces the velocity of propagation of the ultrashort pulse. Our analysis has shown that the thermodynamically equilibrium population of the excited states couples the two 2π-videopulses into one 4π-videopulse, which in view of its high ''inertia'' propagates slower than a 2π-videopulse (see Eq. (19)), which is capable of forming in a medium at ultralow temperatures, when $k_BR\ll\hbar\omega_{21}$ holds.

Figure 2 depicts the dynamics of the populations of quantum levels in an equilibrium medium at $T\neq0$ that accompanies the propagation of an ultrashort pulse in such a medium. The distance between the dynamical minima of L_{11} exceeds the distance between the humps of the ultrashort pulse. The coordinates ξ_1^* of the minima of L_{11} can be found from the condition that $\sinh\xi_1^*=\pm(1+\kappa^2)^{1/2}$. Combining this with (25) yields the minimum value of the ground-level population, $L_{11\min}=R_1$. Thus, in the process of propagation of the 4π-videopulse the ground state does not become completely depleted. The dynamical maximum of L_{11} corresponds to the dip in the electric-field profile of the ultrashort pulse ($\xi=0$). Equation (25) clearly shows that at this point $L_{11}=\rho_{11}$. Thus, in the process of propagation of the 4π-videopulse (23), the medium undergoes excitation and de-excitation twice. The transition of atoms from the ground state as a result of the interaction of the medium and the pulse (23) is accompanied by a rise in the populations of excited states. The positions of the corresponding maxima ξ_m^* ($m=2,3,\ldots,N$) are determined by the equations $\sinh\xi_m^*=\pm(f_m(1+\kappa^2))^{1/2}$, where

$$f_m=\frac{\rho_{11}+\rho_{mm}-2R_1}{\rho_{11}-\rho_{mm}}.$$

Since $R_1=D_\rho^2/D^2<\rho_{22}$ holds, we have

$$f_m>1-2\frac{\rho_{22}-\rho_{mm}}{\rho_{11}-\rho_{mm}}, \quad m>2,$$

whence $f_2>1$. At the same time, the maxima of L_{mm} ($m>2$) are closer to each other and to the center of the dip in the electric-field profile of the ultrashort pulse. Allowing for the fact that $\Sigma_{m=2}^{N}L_{mm}=1-L_{11}$, we conclude that as number of the level increases, the maxima of the corresponding populations move closer to one another monotonically (see Fig. 2). As the absolute temperature rises, the humps in the

ultrashort pulse (23) move closer and the dip becomes more shallow. If the temperature of the medium is so high that the initial populations of all the levels become essentially equal, $\kappa^2 \to 1$. Then the humps of the ultrashort pulse merge and and the central dip disappears. In this case, as Eqs. (28) and (29) imply, the dynamics of the populations of all the levels ceases and we arrive at the hole-burning effect well known from optics.[2]

The conditions $\alpha>0$, $\beta>0$, and $j=1$ can be realized not only in the case of a thermodynamically equilibrium medium satisfying $\kappa>1$ ($\kappa\to\infty$ at $T=0$ and $\kappa\to 1$ as $T\to\infty$). The case with $\kappa<1$ and yet $\alpha>0$ and $\beta>0$ corresponds to a weakly nonequilibrium medium. In this situation the central dip in the electric-field profile disappears (as it does when $T\to\infty$), and the ultrashort pulse is a bell-shaped videopulse with its peak at the middle.

2. $\alpha<0$ and $\beta>0$. In this case the solutions for the field and populations can be found from Eqs. (23)–(26) by replacing κ^2 with $-\kappa_1^2=-2|\alpha|/\beta$, while the parameter κ_1 is restricted by the condition $\kappa_1^2<1$. The electric field of the ultrashort pulse is a 4π-videopulse with a tapered crest (Fig. 3). From the inequalities $\alpha<0$, $\beta>0$, and $\kappa_1<1$ and from (21) we find the conditions that the medium parameters must meet before the ultrashort pulse is applied to the medium:

$$\frac{D_{j\omega\rho}}{D_{j\omega}}<\rho_{jj}<R_j. \tag{33}$$

Analysis of (33) is the simplest and the most graphic if we are dealing with a three-level system. We also assume that $j=1$. Then the condition $D_{j\omega\rho}/D_{j\omega}<R_1$ is equivalent to the inequality $\rho_{22}>\rho_{33}$. Assuming for the sake of simplicity that $\rho_{33}=0$, we find that

$$\frac{d_{21}^2\omega_{21}}{2d_{21}^2+d_{31}^2\omega_{31}}<\rho_{11}<\frac{d_{21}^2}{2d_{21}^2+d_{31}^2}<\frac{1}{2}.$$

Since $\rho_{22}=1-\rho_{11}$ at $\rho_{33}=0$, we have

$$\frac{1}{2}<\frac{d_{21}^2+d_{31}^2}{2d_{21}^2+d_{31}^2}<\rho_{22}<\frac{d_{21}^2\omega_{21}+d_{31}^2\omega_{31}}{2d_{21}^2\omega_{21}+d_{31}^2\omega_{31}}.$$

Thus, $\rho_{22}>\rho_{11}$ holds and the medium is nonequilibrium. If, in addition, we have $d_{31}^2 \gg d_{21}^2$, then the nonequilibrium condition becomes stronger: $\rho_{22}\gg\rho_{11}$. Hence, a steady-state 4π-videopulse can form in a highly nonequilibrium medium with populations of the first and second levels inverted. Notwithstanding this fact, the medium remains absorbing, since according to the condition $d_{31}^2\gg d_{21}^2$ transitions from the first level to the third are more intensive than from the second to the first. This is the reason why, in the given situation, formation is possible of steady-state 4π-videopulses with a steepened crest that travel with speeds smaller than c (see Eq. (24) with κ^2 replaced by $-\kappa_1^2$, $\kappa_1<1$ and $\alpha<0$).

A steady-state 4π-videopulse can also form in an equilibrium medium if the level common to all quantum transitions is not the ground state, i.e., $j\neq 1$. Indeed, in this case at $T=0$ ($\rho_{11}=1$) the condition $\kappa_1^2>1$ can be written as follows:

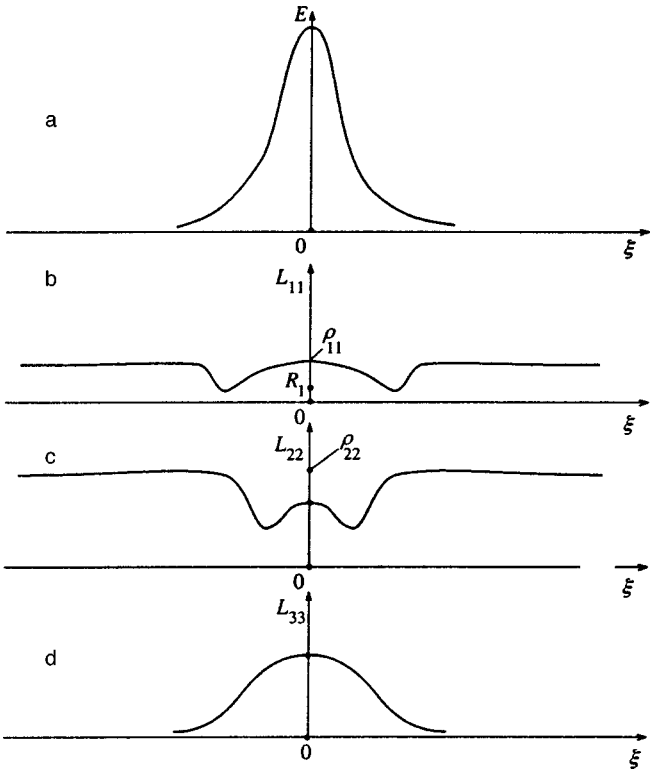


FIG. 3. 4π-videopulse with a steepened crest in a three-level nonequilibrium medium (a) and the corresponding population dynamics (b, c, and d) for $\alpha<0$, $\beta>0$, and $j=1$. Here ρ_{33} is assumed zero and $\rho_{22}>\rho_{11}$.

$$\frac{\sum_{k=1}^{N} d_{kj}^2\omega_{k1}}{\left|\sum_{k=1}^{N} d_{kj}^2\omega_{kj}\right|}>1.$$

At the same time, we have $\alpha<0$ if $\Sigma_{k=1}^{N} d_{kj}^2\omega_{kj}<0$. Since $\omega_{kj}<0$ holds for $k<j$, the condition $\alpha<0$ is sure to be met in an equilibrium medium with $j\neq 1$. Figure 3 depicts the population dynamics in the quantum levels in a three-level nonequilibrium medium with $j=1$ when a steady-state 4π-videopulse with a tapered crest propagates in the medium.

3. $\alpha>0$ and $\beta<0$. The corresponding solution of Eq. (20) for the electric-field profile of an ultrashort pulse is a steady-state traveling 0π-videopulse:

$$E=-\frac{2\hbar}{D_j\tau_p}\sqrt{\kappa_0^2-1}\,\frac{\tanh\xi\,\mathrm{sech}\,\xi}{1+(\kappa_0^2-1)\mathrm{sech}^2\xi}, \tag{34}$$

where

$$\xi=\frac{t-z/v}{\tau_p},\quad \kappa_0^2=\frac{2\alpha}{|\beta|}>1.$$

The velocity v of propagation of the 0π-videopulse is given by the formula

$$\frac{1}{v^2}=\frac{1}{c^2}+\frac{|\beta|}{2}(\kappa_0^2-1)\tau_p^2, \tag{35}$$

and the population dynamics is given by expression of the type

$$L_{jj}=\frac{\rho_{jj}[\cosh^2\xi-(\kappa_0^2-1)]^2+4R_j(\kappa_0^2-1)\cosh^2\xi}{(\kappa_0^2+\sinh^2\xi)^2}, \qquad (36)$$

$$L_{mm}=\rho_{mm}+4\frac{d_{mj}^2}{D_j^2}(\kappa_0^2-1)$$

$$\times\frac{(\kappa_0^2-1)R_j+\rho_{jj}\cosh^2\xi-\rho_{mm}}{(\kappa_0^2+\sinh^2\xi)^2}, \quad m\neq j. \qquad (37)$$

With allowance for (21), the conditions $\alpha>0$, $\beta<0$, and $\kappa_0^2=2\alpha/|\beta|>1$ can be represented in the following manner:

$$R_j<\frac{D_{j\omega\rho}}{D_{j\omega}}<\rho_{jj}. \qquad (38)$$

With $j=1$ in a three-level system, Eq. (38) implies, in particular, $\rho_{33}>\rho_{22}$. If we then assume $\rho_{22}=0$, we get

$$\rho_{11}>\frac{d_{31}^2}{2d_{31}^2+d_{21}^2}, \quad \rho_{33}<\frac{d_{31}^2+d_{21}^2}{2d_{31}^2+d_{21}^2}.$$

In the limit $d_{21}^2\gg d_{31}^2$, these inequalities become

$$\rho_{11}>\left(\frac{d_{31}}{d_{21}}\right)^2, \quad \rho_{33}<1-\left(\frac{d_{31}}{d_{21}}\right)^2.$$

These conditions are also met for $\rho_{33}>\rho_{11}$.

Note that when the pulse propagates in such a medium, the leading edge of the 0π-videopulse lowers the population of only the ground state (Fig. 4). Thus, being nonequilibrium, the medium remains remain absorbing for $\kappa_0^2>1$. On the other hand, the middle section of the pulse ($E=0$) restores the initial population of the ground state and lowers the population of the third level, moving some of the atoms to the middle level. Combining (21) with (27), we find that in the case of thermodynamic equilibrium,

$$\beta=\frac{16\pi n}{\hbar c^2 D_j^2 Z}\sum_{k>m} d_{kj}^2 d_{mj}^2\omega_{km}\left[\exp\left(-\frac{\hbar\omega_{mj}}{k_BT}\right)\right.$$

$$\left.-\exp\left(-\frac{\hbar\omega_{kj}}{k_BT}\right)\right]\geqslant 0.$$

The condition $\beta<0$ is not met, so that 0π-videopulses may form only in a nonequilibrium medium.

4. AMPLIFICATION

Now suppose that the restrictions imposed on the coefficients α and β in Sec. 3 are not met. Then, in the absence of dissipation, a nonequilibrium medium will irreversibly transfer part of its energy to the optical pulse, amplifying the latter. Below we study the specific features of such amplification.

At $\gamma=0$ Eq. (20) has a Hamiltonian representation:

$$\frac{\partial\tilde{P}}{\partial t}=-\frac{\delta H}{\delta\theta}, \quad \frac{\partial\theta}{\partial t}=\frac{\delta H}{\delta\tilde{P}},$$

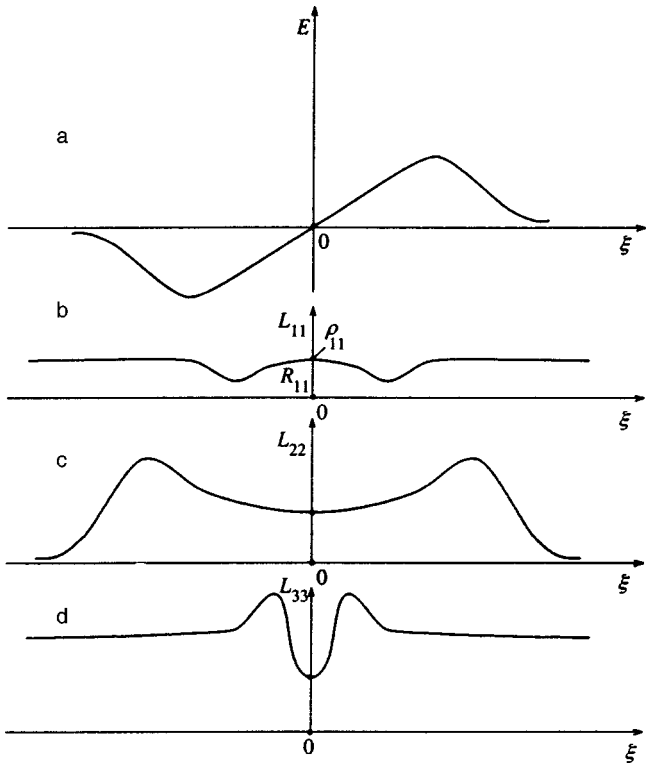


FIG. 4. Dynamics of the electric field of ultrashort pulses in their propagation in a three-level nonequilibrium medium (a) and the population dynamics (b, c, and d) for $\alpha>0$, $\beta<0$, and $j=1$. Here ρ_{22} is assumed zero and $\rho_{33}>\rho_{11}$.

where $H=\int\mathscr{H}d^3r$, and the Hamiltonian density is

$$\mathscr{H}=\frac{c^2}{2}\tilde{P}^2+\frac{1}{2}(\nabla\theta)^2-\alpha\cos\theta-2\beta\cos\frac{\theta}{2}. \qquad (39)$$

The ''vacuum'' values of θ ($\theta=$const), which minimize $\mathscr{H}$, can be found from the condition that the function

$$-\left(\alpha\cos\theta+2\beta\cos\frac{\theta}{2}\right)$$

be at its minimum. The possible areas of the propagating pulses are equal to these ''vacuum'' values θ_v. For instance, in the steady-state propagation modes studied in Sec. 3, $\theta_v=4\pi n$ ($n=0,1,2,\ldots$) for $\beta\neq 0$. Obviously, in the amplification modes the area of the ultrashort pulse is also determined by the values θ_v. For instance, for a two-level medium ($\beta=0$) with $\alpha<0$ we have $\theta_v=\pi$ (see (39)), which coincides with the results of Refs. 14 and 16. In our case, amplification occurs when any one of the following three conditions is met.

1. $\alpha<0$ and $\beta<0$. Here for $\kappa\equiv(2\alpha/\beta)^{1/2}>1$ the Hamiltonian density $\mathscr{H}$ has its minimum at the value θ_m specified by the condition

$$\cos\frac{\theta_v}{2}=-\frac{1}{\kappa}.$$

But for $\kappa<1$ the value of $\mathscr{H}$ is minimal at $\theta_v=2\pi n$ ($n=0,1,2,\ldots$). Suppose than an ultrashort pulse with a small area

$$\Phi \equiv \frac{2D_j}{\hbar}\int_{-\infty}^{\infty} E\,dt'$$

is incident on the medium. In the first case ($\kappa>1$), the pulse is amplified with the area Φ increasing simultaneously to the values $\theta_v = 2\arccos(\kappa^{-1})$. Clearly, the given area may vary from $\Phi=0$ ($\kappa=1$) to $\phi=\pi$ ($\kappa=\infty$) and may equal $q\pi$, where generally q is an irrational number satisfying the condition $0<q<1$. Hence we call this amplified signal an irrational signal. When $\kappa<1$ holds, in the process of being amplified the ultrashort pulse becomes a non-steady-state 2π-videopulse with an increasing amplitude. After the $q\pi$- and 2π-videopulses have traveled in the medium, we can find the population levels by (13) and (14) at $\theta=q\pi$ and $\theta=2\pi$, respectively.

2. $\alpha<0$ and $\beta>0$. Amplification is possible if

$$\kappa_1 \equiv \left(\frac{2|\alpha|}{\beta}\right)^{1/2} > 1,$$

since for $\kappa_1<1$ 4π-videopulses with a steepened crest are formed (see Sec. 3). The Hamiltonian density $\mathscr{H}$ is minimal at $\theta_v = q\pi$, where $q=\pi^{-1}\arccos(\kappa_1^{-1})$. The value $\theta_v=2\pi$ corresponds to the maximum of $\mathscr{H}$, and therefore is not included in our discussion. Setting $\theta=q\pi$ in (13) and (14), we find the level populations after the passage of the amplified $q\pi$-videopulse.

3. $\alpha>0$ and $\beta<0$. Reasoning along the same lines as in the first two cases, we conclude that the initial signal is amplified if

$$\kappa_0 \equiv \left(\frac{2\alpha}{|\beta|}\right)^{1/2} < 1.$$

The largest area of the videopulse is 2π. Accordingly, the final populations can be found from (13) and (14) with $\theta=2\pi$.

Now let us study the dynamical features of the amplification of an ultrashort pulse in a spatially one-dimensional case. Following Refs. 14 and 16, we introduce the self-similar variable $\eta=z^2-c^2t^2$. Then at $\gamma=0$ Eq. (20) becomes

$$\eta\theta'' + \theta' = \frac{\alpha}{4}\sin\theta + \frac{\beta}{2}\sin\frac{\theta}{2}, \tag{40}$$

where the prime on θ indicates a derivative with respect to η.

Suppose that a function $\theta(\eta)$ is a solution of Eq. (40). Then the electric field of the ultrashort pulse is

$$E = -\frac{\hbar c^2}{D_j}\,t\,\theta'(\eta), \tag{41}$$

i.e., the amplitude of the signal increases in proportion to t. Since in the limit the area of the ultrashort pulse reaches its maximum value and ceases to increase, the amplification of this pulse is accompanied by self-squeezing. Here the signal length $\tau_p \sim t^{-1}$. Following Ref. 29, we set $\theta=\theta_v+\varepsilon$ ($\varepsilon \ll 1$) in Eq. (40). As a result the equation reduces to

$$\eta\varepsilon'' + \varepsilon' = \mu\varepsilon, \tag{42}$$

where for a 2π-pulse

$$\mu = \frac{\alpha}{4}\left(1-\frac{\beta}{2\alpha}\right) > 0,$$

and for a $q\pi$-pulse ($0<q<1$)

$$\mu = -\frac{\alpha}{4}\left(1-\left(\frac{\beta}{2\alpha}\right)^2\right) > 0.$$

The solution of Eq. (42) in terms of the variables z and t is

$$\varepsilon \sim J_0(2\sqrt{\mu(c^2t^2-z^2)}), \quad ct>|z|,$$

$$\varepsilon \sim K_0(2\sqrt{\mu(z^2-c^2t^2)}), \quad ct<|z|.$$

For the electric field of an ultrashort pulse we have

$$E \sim \frac{t}{\sqrt{c^2t^2-z^2}}J_1(2\sqrt{\mu(c^2t^2-z^2)}), \quad ct>|z|, \tag{43}$$

$$E \sim \frac{t}{\sqrt{z^2-c^2t^2}}K_1(2\sqrt{\mu(z^2-c^2t^2)}), \quad ct<|z|, \tag{44}$$

where $J_1(x)$ and $K_1(x)$ are, respectively, the Bessel and modified Hankel functions of the order 1. Combining this with (13) and (14), we see that as the trailing edge of the amplified signal passes, the values of the populations of the quantum levels of atoms oscillate about the quasiequilibrium values $L_{jj}(\theta_v)$ and $L_{mm}(\theta_v)$ ($m\neq j$). Here the amplitude of these oscillations decays according to the asymptotic behavior of the Bessel function. On the other hand, the electric field of the pulse varies according to the damped oscillations mode for $ct>|z|$ and decreases very rapidly (exponentially, for all practical purposes) to zero for $ct<|z|$.

Let us study the two limits, $|\alpha|\gg|\beta|$ and $|\alpha|\ll|\beta|$. Then Eq. (20) at $\gamma=0$ becomes an ordinary sine-Gordon equation. Equation (40) corresponding to these case has solutions for E that are nonzero primarily at $\eta=0$ (Refs. 14, 16, and 28). Hence in the vicinity of this principal maximum we can ignore the term $\eta\theta''$ in (40) and write the approximate solutions for the field in the neighborhood of $z=ct$:

$$E = \lambda\frac{\hbar\omega(z)}{D_j}\operatorname{sech}\left[\omega(z)\left(t-\frac{z}{c}\right)\right], \tag{45}$$

where $\omega(z)=|\alpha|cz/2$ and $\lambda=1/2$ for $|\alpha|\gg|\beta|$, and $\omega(z)=|\beta|cz/4$ and $\lambda=1$ for $|\beta|\gg|\alpha|$.

Setting $z^2-c^2t^2 \approx 2z(z-ct)$ in (43), we establish the behavior of the electric field outside the principal maximum:

$$E \sim z\frac{J_1(2\sqrt{\omega(z)(t-z/c)})}{2\sqrt{\omega(z)(t-z/c)}}. \tag{46}$$

Figure 5 depicts the electric-field profiles of an ultrashort pulse two subsequent values of the coordinates $z_m=z_1$ and $z_m=z_2>z_1$ of the principal maximum.

Equations (45) and (46) clearly show that the function $\omega(z)$, which increases linearly with the distance covered, can be interpreted as the frequency of the photons of the given pulse. Equation (45) implies that the frequency and amplitude of the electric field increase in proportion to z. The

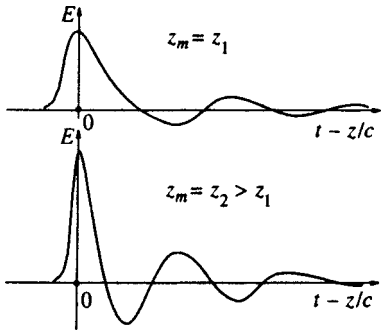


FIG. 5. Electric-field profiles of ultrashort pulses in a nonequilibrium amplifying medium, described by Eqs. (45) and (46), for two subsequent values of the coordinate z_m of the principal maximum (at the extreme left), $z_m = z_1$ and $z_m = z_2 > z_1$. Amplification is accompanied by an increase in the frequency of the pulse photons and the photon number density.

pulse intensity I can be expressed by the formula $I \sim E^2 \sim n_{\rm ph}\hbar\omega$, where $n_{\rm ph}$ is photon number density in the pulse. Since $E \sim z$ and $\omega \sim z$, we conclude that $n_{\rm ph} \sim z$. Hence, the amplification of an ultrashort pulse with $\omega \gg |\omega_{mj}|$ occurs due to the increase in the frequency of each photon and due to the increase in the photon number density. Both parameters increase in proportion to the distance covered by the pulse. This result refines the conclusion drawn by Belenov *et al.*,[14,16] who stated that the amplification of an ultrashort pulse is due solely to the increase in photon frequency.

The last statement is entirely true when an ultrashort pulse interacts with BKP transitions, which are induced by the square of the field rather than by the field proper.[30]

5. AN AUTOSOLITON IN A NONEQUILIBRIUM MEDIUM

Suppose that an ultrashort pulse is propagating in a nonequilibrium medium with dissipation ($\gamma \neq 0$). Then the energy transferred by the excitated atoms to the pulse field due to stimulated transitions can dissipate irreversibly in the medium due to extraneous losses determined by the parameter γ. When these two opposing processes balance each other, a stable structure of in the form of an autosoliton may develop.[18,31] After the autosoliton has traveled through the medium, the medium does not return to its initial state, as would be the case if the medium were conservative. In a dissipative medium the law of energy conservation does not hold for a system consisting of N-level atoms and an electromagnetic field of the pulse. Hence the atoms do not have enough time to go back to the initial quantum states and irreversibly release part of their energy to the other components of the medium.

We seek the solution of Eq. (20) in the form of a wave traveling along the z axis, $\theta = \theta(t - z/v)$. This yields

$$\left(\frac{1}{v^2} - \frac{1}{c^2}\right)\ddot{\theta} - \frac{\gamma}{c^2}\dot{\theta} = \alpha \sin\theta + \beta \sin\frac{\theta}{2}, \tag{47}$$

where the dot on θ stand for the derivative with respect to the variable $t - z/v$.

We choose the ansatz[31]

$$\dot{\theta} = \frac{4}{\tau_p}\sin\frac{\theta}{2}, \tag{48}$$

whence $\ddot{\theta} = (4/\tau_p^2)\sin\theta$. Substituting this into (47) and (48) and equating the corresponding coefficients of $\sin(\theta/2)$ and $\sin\theta$, we find that

$$\tau_p = -\frac{4\gamma}{\beta c^2}, \quad \frac{1}{v} = \frac{1}{c}\sqrt{1 + \alpha\left(\frac{2\gamma}{\beta c}\right)^2}. \tag{49}$$

Integrating (48), we arrive at an expression for the electric field $E = (\hbar/2D_j)\partial\theta/\partial t$ of the autosoliton:

$$E = \frac{\hbar}{D_j \tau_p}\,\mathrm{sech}\,\frac{t - z/v}{\tau_p}. \tag{50}$$

The autosoliton (50) is a unipolar videopulse. It is stable if $E > 0$ holds, since otherwise the vectors E and d_{mj} ($m \neq j$) point in opposite directions, which contradicts the principle of stability of a state with minimal energy. From (50) we conclude that $\tau_p > 0$, which yields $\beta < 0$. Moreover, the solution (50) has physical meaning if $v < c$. Hence (see (49)) $\alpha > 0$. Thus, an autosoliton of the form (50) can develop only in such a nonequilibrium medium for which $\alpha > 0$ and $\beta < 0$. It will be recalled that in this case, in the absence of dissipation, either a 0π-videopulse is formed ($\kappa_0 = (2\alpha/|\beta|)^{1/2} > 1$) or the process of amplification and self-squeezing of a pulse with maximum area equal to 2π ($\kappa_0 < 1$) takes place. The area of the autosoliton (50) is $\Phi = 2\pi$. However, as Eqs. (49) show, no restriction on the parameter κ_0 is needed for such an autosoliton to form. Nevertheless, there is one important restriction on the parameter γ, for which the inequality (5) is responsible. Equations (21) imply $|\beta| \sim 16\pi n d^2 \omega_0/\hbar c^2$, where d and ω_0 are the characteristic values of the dipole moment and the transition frequency, Then from (5) and (49) we find that $\gamma \ll 4\pi d^2 n/\hbar$. Substituting $d \sim 5 \times 10^{-18}$ abs.u. and $n \sim 10^{21}\,\mathrm{cm}^{-3}$ into this inequality yields $\gamma \gg 10^{14}\,\mathrm{s}^{-1}$.

Let us consider one of the possible autosoliton formation processes. Suppose that $\kappa_0 < 1$. Then the amplification of the principal peak in the 2π-videopulse of type (45) and the self-squeezing slow down due to dissipation until stabilization accompanied by a reduction in the propagation velocity sets in.

Now we study the processes that take place at the leading and trailing edges of the pulse. Linearizing Eq. (20) via the substitution $\theta = 2\pi + \varepsilon$ ($\varepsilon \ll 1$) and defining ε as $\psi(\mathbf{r},t)\exp(-\gamma t/2)$, we reduce Eq. (20) to

$$\Delta\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} = 4\tilde{\mu}\psi,$$

where $\tilde{\mu} = \mu - \gamma^2/16c^2$, with $\mu = (\alpha/4)(1 + |\beta|/2\alpha) > 0$.

Introducing the self-similar variable $\eta = z^2 - c^2t^2$ (see Sec. 4), we find that in the spatially one-dimensional case

$$E \propto t\exp\left(-\frac{\gamma t}{2}\right)J_1(2\sqrt{\zeta}),$$

where $\zeta = \tilde{\mu}(c^2t^2 - z^2)$.

When $\zeta>0$ holds, the oscillations in the ''wings'' at first grow in amplitude in proportion to t, but are then quenched by the damping exponential and finally completely disappear in the limit $t\to\infty$. For $\zeta<0$ the Bessel function $J_1(2\sqrt{\zeta})$ becomes the modified Hankel function, which decays exponentially with increasing $|\zeta|$. Here, too, initial amplification ($\propto t$) is replaced by damping. Thus, as a result of smoothing out of the ''tails'' and stabilization of the principal peak, an autosoliton of the form (50) develops.

When an autosoliton travels in the medium, the population dynamics in the quantum levels can easily be established by substituting (50) into (13) and (14). However, we will determine only the final populations, by setting $\theta=2\pi$ in (13) and (14). Then

$$L_{jj}=\rho_{jj},$$

$$L_{mm}=\left(1-4\frac{d_{mj}^2}{D_j^2}\right)\rho_{mm}+4\frac{d_{mj}^2}{D_j^2}R_j, \quad m\neq j.$$

Thus, the population of the jth level common for all transitions returns to its initial value. The populations of all the other levels change irreversibly. As a result we find that when the autosoliton passes through the medium, the state of the latter changes. This is the main difference between our case and that of an autosoliton, apart from the case of a soliton traveling in a conservative medium, since such a medium returns to its initial state after the soliton has passed through it (see Sec. 3). Another important difference between the solution (50) and the videopulses examined in Sec. 3 that the autosoliton has no free parameters: according to (49) and (50), its amplitude, speed of propagation, and length are rigidly determined by the parameters of the medium. The reason is that a pulse in a dissipative medium loses the information about its initial state, and in the limit $t\to\infty$ ''forgets'' everything about the initial conditions. This is not the case for solitons propagating in a conservative medium, where the information about the initial (boundary) conditions is retained, so that the solutions contain at least one free parameter each.

6. CONCLUSION

In this paper we have studied the various modes of propagation of ultrashort pulses in multilevel quantum system in conditions where the spectra of the pulses overlap the quantum transitions under investigation. We have considered a special, and yet fairly broad, class of transitions with one common quantum level. Due to the spectral overlap of all the quantum transitions by the field of an ultrashort pulse, the interaction in this case can formally be considered resonant. Indeed, since the spectrum of the ultrashort pulse substantially overlaps the frequencies of the transition, it always contains spectral components that are resonant to all the transitions. For the same reason, here inhomogeneous broadening does not play a decisive role since the overlap exists even with such broadening.[32] As a result we conclude that for ultrashort pulses satisfying condition (5), an equilibrium absorbing medium becomes nonlinearly transparent in the sense that, after being excited by the leading edge of an ultrashort pulse, the atoms are returned to their initial state by the trailing edge, returning all the absorbed energy to the field. Notwithstanding the similarity to the mechanism of excitation–deexcitation of a medium, there are certain important differences in the case of self-induced transparency for monochromatic resonant pulses. Probably, the most important one is the difference between the velocities of the soliton of an ultrashort pulse and of the soliton of the envelope of a monochromatic signal. Equations (24), (35), and (21) yield an estimate for the velocities of ultrashort pulses:

$$v\sim\frac{c}{\sqrt{1+(8\pi d^2 n/\hbar\omega_0)(\omega_0\tau_p)^2}}.$$

For a broad class of substances not undergoing a superradiant transition, $8\pi d^2 n/\hbar\omega_0<1$ (Refs. 33 and 34). Then in view of the inequality (5) we have $v\leqslant c$. Thus, the velocity of a broadband ultrashort pulse propagating in an absorbing medium is only slightly smaller than the speed of light in vacuum, while the velocity of resonant solitons of the envelope is smaller than c by two to three orders of magnitude.[2] Since in the above sense all the transitions are resonant, probably there can be no area theorem here (such a theorem enables analyzing the variation in the area of an ultrashort pulse as the pulse propagates in an absorbing medium[2,28]). Instead of this we can use the method of minimizing the Hamiltonian, which was proposed in Sec. 4.

In the present paper we have examined π-transitions, for which the matrix $\hat{A}$ (see Eq. (2)) is real. At the same time, magnetic rotation of the polarization plane of the light wave is explained by σ-transitions, for which $\hat{A}$ is complex-valued. Such transitions start at the ground Zeeman sublevels of the s state and end at the higher Zeeman sublevels of the p state ($j=1$). Thus it would interesting to generalize our approach to the case of σ-transitions. In this way we could study the Faraday rotation of the polarization plane of ultrashort pulses for arbitrary initial populations of the excited Zeeman sublevels. In Ref. 35 this effect was studied for zero initial population.

Note that for pulses with a rotating polarization plane condition (7) is not met. Hence the solution (8) becomes invalid, and we must use other methods to solve the material equations (6). Even when there is spectral overlap, finding the general solution of the material equation (1) for a matrix $\hat{A}$ with an arbitrary structure is extremely difficult mathematically. Hence the study of the interaction of an ultrashort pulse and multilevel quantum systems with arbitrary transition is not simple. The approximation (5) has meaning provided that the overlapped quantum levels are far from the other levels of the discrete spectrum, since otherwise we would have to account for transitions into the continuous spectrum, or ionization, and this would complicate matters substantially. Hence there arises the problem of correctly accounting for the effect of these distant levels of the discrete spectrum, which not overlapped by the spectrum of ultrashort pulses. To do this, all quantum transitions are divided into two groups, with one incorporating the transitions that overlapped by the spectrum of ultrashort pulses, and the other incorporating the transitions that are not overlapped by

this spectrum. The latter group form the region of linear transparency. In Ref. 36 these two groups were interpreted as independent subsystems (vibrational and electronic). The solution of the problem when quantum transitions between the two groups of levels are possible would result would mean an important generalization and progress in this area of research.

[1] S. L. McCall and E. L. Hahn, Phys. Rev. Lett. **18**, 908 (1967).
[2] L. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms*, Wiley, New York (1975).
[3] R. K. Bullough and F. Ahmad, Phys. Rev. Lett. **27**, 330 (1971).
[4] C. T. Lee, Opt. Commun. **9**, 1 (1973).
[5] P. J. Caudrey, J. C. Eilbeck, J. D. Gibbon, and R. K. Bullough, J. Phys. A **6**, L53 (1973).
[6] J. C. Eilbeck, J. D. Gibbon, P. J. Caudrey, and R. K. Bullough, J. Phys. A **6**, 1337 (1973).
[7] J. D. Gibbon, P. J. Caudrey, R. K. Bullough, and J. C. Eilbeck, Lett. Nuovo Cimento **8**, 775 (1973).
[8] D. H. Auston, K. P. Cheung, I. A. Valdmanis, and D. A. Kleinmann, Phys. Rev. Lett. **53**, 1555 (1984).
[9] J. T. Darrow, B. B. Hu, X. C. Chang, and D. H. Auston, Opt. Lett. **15**, 323 (1990).
[10] P. C. Becker, H. L. Fragnito, J. Y. Bigot, C. H. Brito-Crus, and C. V. Shank, Phys. Rev. Lett. **63**, 505 (1989).
[11] A. Kujawski, Z. Phys. **66**, 271 (1987).
[12] A. Kujawski, Z. Phys. **85**, 129 (1991).
[13] A. V. Andreev, Zh. Éksp. Teor. Fiz. **108**, 796 (1995) [JETP **81**, 434 (1995)].
[14] É. M. Belenov, P. G. Kryukov, A. V. Nazarkin, A. N. Oraevskiĭ, and A. V. Uskov, JETP Lett. **47**, 523 (1988).
[15] É. M. Belenov and A. V. Nazarkin, JETP Lett. **51**, 288 (1990).
[16] É. M. Belenov, A. V. Nazarkin, and V. A. Ushchapovskiĭ, Zh. Éksp. Teor. Fiz. **100**, 762 (1991) [Sov. Phys. JETP **73**, 422 (1991)].
[17] R. H. Pantell and H. E. Puthoff, *Fundamentals of Quantum Electronics*, Wiley, New York (1969).
[18] S. V. Sazonov, JETP Lett. **53**, 420 (1991).
[19] I. A. Lappo-Danilevskiĭ, *Application of Matrix Functions to the Theory of Linear Systems of Ordinary Differential Equations* [in Russian], Gostekhizdat, Moscow (1957).
[20] R. Braunstain, Phys. Rev. **125**, 475 (1962).
[21] R. Braunstain and N. Ockman, Phys. Rev. A **134**, 499 (1964).
[22] G. B. Al'tshuler, Opt. Spektrosk. **55**, 83 (1983) [Opt. Spectrosc. **55**, 47 (1983)].
[23] F. Bassani and G. Pastori Parravicini, *Electron States and Optical Transitions in Solids* [Russian translation], Nauka, Moscow (1982).
[24] E. B. Aleksandrov and V. S. Zapasskiĭ, *Laser Magnetic Spectroscopy* [in Russian], Nauka, Moscow (1986).
[25] C. Kittel, *Introduction to Solid State Physics*, Wiley, New York (1956).
[26] D. N. Klyshko, *Physical Bases of Quantum Electronics* [in Russian], Nauka, Moscow (1986).
[27] L. S. Yakupova, in: *Nonlinear Evolution Equation and Dynamical Systems* (The 8th Int. Workshop, Dubna, 1992), World Scientific, Singapore (1993), p. 432.
[28] G. L. Lamb, *Elements of Soliton Theory*, Wiley-Interscience, New York (1980).
[29] S. V. Sazonov and E. V. Trifonov, Zh. Éksp. Teor. Fiz. **103**, 1527 (1993) [JETP **76**, 744 (1993)].
[30] É. M. Belenov, P. G. Kryukov, A. V. Nazarkin, and I. P. Prokopovich, Zh. Éksp. Teor. Fiz **105**, 28 (1994) [JETP **78**, 15 (1994)].
[31] S. V. Sazonov, J. Phys.: Condens. Matter **7**, 175 (1995).
[32] S. V. Sazonov and E. V. Trifonov, J. Phys. B **27**, L7 (1994).
[33] 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).
[34] S. V. Sazonov, Izv. Ross. Akad. Nauk, Ser. Fiz. **62**, 430 (1998) [Izv. Ross. Akad. Nauk, Ser. Fiz. **62**, 350 (1998)].
[35] S. V. Sazonov, Zh. Éksp. Teor. Fiz. **107**, 20 (1995) [JETP **80**, 10 (1995)].
[36] S. A Kozlov and S. V. Sazonov, Zh. Éksp. Teor. Fiz. **111**, 404 (1997) [JETP **84**, 221 (1997)].

Translated by Eugene Yankovsky

Numerical simulation of the ionization dynamics of a two-electron quantum system in a femtosecond pulse

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

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

(Submitted 14 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 1618–1635 (November 1998)

The ionization of a simple two-electron model system, viz., the one-dimensional negative hydrogen ion, is investigated using direct numerical integration of the time-dependent Schrödinger equation. The one- and two-electron ionization probabilities as functions of frequency and radiation intensity are obtained. It is shown that two-electron ionization is mediated by both direct and sequential mechanisms. The stabilization of the two-electron system against the ionization process is investigated. The data obtained are compared with calculations performed within the one-dimensional single-particle model of H^-. The photoelectron spectrum is analyzed in the region of parameters corresponding to the single-electron ionization regime.
 [S1063-7761(98)00511-3]

1. INTRODUCTION

The investigation of the dynamics of atomic systems in high-intensity laser fields is presently of great interest.[1] However, the theoretical description of the interaction of electromagnetic radiation with atoms has generally been confined to the one-electron approximation. Within this approximation the quantum system is represented in the form of a single electron moving in a certain effective static potential, which takes into account the Coulomb interaction with the nucleus, as well as the partial screening of the nucleus by the remaining electrons. The one-electron approximation can clearly be used successfully to describe hydrogenic systems characterized by the presence of a single weakly bound valence electron moving in the field of an atomic core. However, even in this case a strong external field can significantly distort the electrostatic potential created by the atomic core, in which the outer electron moves, and can thereby exert an additional influence on it. As for negative ions, they are usually characterized by the presence of several (at least two) electrons in the outer shell, which raises some question as to the possibility of describing their ionization process in the one-electron approximation. The only exception is probably the negative hydrogen ion H^-, which consists of a weakly bound electron revolving around a core, i.e., a hydrogen atom. Therefore, it can be stated that the construction of physically correct models of the photoionization of many-electron atomic system is of definite interest.

Direct numerical integration of the time-dependent Schrödinger equation for a quantum system in an electromagnetic field has permitted the investigation of the ionization of one-dimensional two-electron systems without any simplifying assumptions. For example, the ionization of the model one-dimensional He atom was investigated in Refs. 2 and 3, and the negative hydrogen ion was investigated in Refs. 4–6.

The dependence of the ionization probability of the system on the radiation intensity were obtained, and the photoelectron spectra were analyzed in those studies. However, the key questions which arise in describing the dynamics of two-electron systems in an electromagnetic field, in our opinion, remain unanswered. In particular, the conditions under which the dynamics of the system can be described in the one-electron approximation and the way in which the presence of the second electron influences the stabilization regime predicted within the one-electron model must be ascertained. Finally, in the case of two-electron ionization, which mechanism, the direct or sequential, more faithfully describes the physical essence of the process?

In the present work we obtain an exact numerical solution of the two-particle Schrödinger equation for the one-dimensional negative hydrogen ion in an electromagnetic field. The wave function obtained is then used to calculate the one- and two-electron ionization probabilities as functions of the radiation intensity and frequency. It is shown that two-electron ionization is a result of competition between the direct and sequential processes. The photoelectron energy spectra are obtained in the region for one-electron ionization. The region for the existence of a stabilization regime is determined. The possibility of describing the dynamics of H^- within the one-electron model is investigated.

2. TWO-PARTICLE ONE-DIMENSIONAL MODEL OF THE NEGATIVE HYDROGEN ION

It is assumed within the one-dimensional model that the interaction of the electrons with the nucleus and with one another can be described by a smoothed Coulomb potential.[7]

Therefore, the Hamiltonian of the system can be written in the form

$$H=\sum_{i=1}^{2}\,(T_i+V(x_i))+V_{12}(x_1,x_2). \tag{1}$$

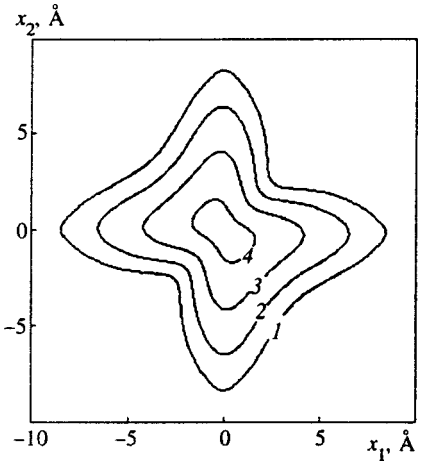


FIG. 1. Distribution of the probability density $\rho(x_1,x_2)$ for the ground state of the one-dimensional negative hydrogen ion. The lines of equal probability density correspond to 10^{-4} (*1*), 10^{-3} (*2*), 0.01 (*3*), 0.1 (*4*).

Here T_i is the kinetic energy of the ith electron, $V(x_i)=-e^2/\sqrt{\alpha^2+x_i^2}$ is the energy of its interaction with the nucleus, $V_{12}=e^2/\sqrt{\alpha^2+(x_1-x_2)^2}$ is the energy of the interaction of the electrons with one another, and α is a smoothing parameter.

It has been reported[7] that a smoothed Coulomb potential provides a qualitatively correct description of the structure of the energy spectrum of the hydrogen atom, particularly the presence of an infinitely large number of Rydberg states, which condense toward the continuum edge. As for the ground state, its energy is very sensitive to the choice of the value of α. Therefore, it can be expected in a two-electron system that the energy of the bound states and their number will depend strongly on the choice of the value of the smoothing parameter.

We obtained stationary states corresponding to the Hamiltonian (1) by solving the two-particle time-independent Schrödinger equation

$$H\varphi(x_1,x_2)=E\varphi(x_1,x_2) \tag{2}$$

on a rectangular grid. In this process we used the symmetry of the Hamiltonian under spatial inversion,

$$H(x_1,x_2)=H(-x_1,-x_2),$$

and interchange of the electrons,

$$H(x_1,x_2)=H(x_2,x_1).$$

The solution method is described in the Appendix.

The ground-state wave function of H^- for $\alpha=0.92$ Å is presented in Fig. 1. The energy of this state equals $E_0=-12.56$ eV. If the binding energy of the electron in the hydrogen atom for this value of α, $E_1=-11.45$ eV, is taken into account, we find that the energy for detaching an electron from H^- (the ionization potential) amounts to $I\sim 1.1$ eV, which is fairly close to the experimental value $I\approx 0.75$ eV. We note that the bound state found for the parameter value chosen is the only such state.

In Refs. 4 and 5, $\alpha=a_0$ (the Bohr radius) was chosen. In this case it turned out that $I\approx 1.7$ eV, and the number of bound states remained unclear: Grobe and Eberly[5] assumed the appearance of a second state in the discrete spectrum near the continuum edge. Our calculations for $\alpha=a_0$ show that such a state probably exists, but the accuracy of the calculations must be improved to determine its energy.

We note that the single bound state in the model (as in the real three-dimensional) negative hydrogen ion is characterized by a wave function which is symmetric under interchange of the electrons. This means that the total spin S in this state equals zero. Since intercombination transitions are forbidden in the electric dipole approximation, the state of the two-electron system always has a singlet character during a laser pulse.

The data in Fig. 1 attest to the relative ''weakening'' of the binding of one of the electrons in the system: states in which one of the electrons is near the attractive center, while the other is at a considerably greater distance from it are more probable. The observed weakening of the binding of one of the electrons, of course, differs significantly from the classical models of a weakly bound electron moving in an orbit with a radius larger than that of the inner electron. From the standpoint of our one-dimensional quantum-mechanical model, such a classical picture should correspond to a distribution of the electron density $\rho(x_1,x_2)$ with local maxima in regions near the points $x_1\approx 0$ and $x_2=\pm x_m$ or $x_2\approx 0$ and $x_1=\pm x_m$ (x_m is the distance from the center to the weakly bound electron). The real distribution of the two-electron density has a ''cross-shaped'' appearance, and it attains its maximum at the origin of coordinates $x_1=x_2=0$. However, if the two-electron function $\Psi(x_1,x_2)$ is approximated by a function constructed in the form of a symmetrized product of two one-electron orbitals $u(x)$ and $v(x)$, i.e.,

$$\Psi(x_1,x_2)\sim(u(x_1)v(x_2)+u(x_2)v(x_1)),$$

then for the two-electron density shown in Fig. 1 be obtained, these orbitals must be characterized by different dispersions of the coordinates, i.e., by regions of spatial localization of the electrons of different sizes. For this reason, we shall henceforth refer to an outer and an inner electron. As an example of a wave function with such a structure we can cite the Chandrasekhar wave function[8]

$$\Psi(r_1,r_2)\propto\exp(-r_1/a_1-r_2/a_2)+\exp(-r_2/a_1-r_1/a_2)$$

($a_1\approx 0.97a_0$, $a_2\approx 3.53a_0$, a_0 is the Bohr radius), which describes the bound state of the three-dimensional negative hydrogen ion.

The wave function obtained for the ground state of the negative hydrogen ion attests to strong electron–electron correlations in the system. In particular, as can be seen from Fig. 1, such states are more probable when the electrons are located on opposite sides of the attractive center. The electron–electron correlations lead to a loss of symmetry for the wave function $\Psi(x_1,x_2)$ under inversion with respect to one of the coordinate axes: $x_1\to -x_1$ or $x_2\to -x_2$. This feature of $\Psi(x_1,x_2)$ cannot be described in terms of one-electron orbitals.

TABLE I.

n	1	2	3	4	5	6	7	8	9
E_n, eV	−11.45	−6.05	−3.85	−2.52	−1.74	−1.26	−0.96	−0.60	−0.41

The degree of correlation of the motion of the electrons can be characterized quantitatively by the correlation coefficient K, which was introduced in Ref. 9. The value of K was calculated in Ref. 10 for the ground state of the one-dimensional H^- and He systems, which were treated within the exact two-particular model. It was found that there are stronger correlations in the H^- system. This is because the electron–electron interaction energy in this system, unlike the helium atom, can no longer be considered small in comparison to the interaction energy of each of the electrons with the nucleus.

Unlike the bound state, all the excited states of H^- belong to the continuous spectrum. They include one-electron continuum states, in which one of the electrons is in the discrete spectrum, while the other is in the continuous spectrum, and two-electron continuum states, which correspond to infinite motion of both electrons.

The one-electron continuum states are ''hydrogen atom + free electron'' systems. If the interactions of the electrons with one another are disregarded for such states and, in addition, if the wave function of the free electron is approximated by a de Broglie plane wave with the wave vector k and the energy $E=\hbar^2k^2/2m$, then in the case of $S=0$ with consideration of the identity principle, the one-electron continuum states will be described by a spatial wave function of the form

$$\varphi_{nk}(x_1,x_2)=\frac{1}{\sqrt{2}}\left\{\Phi_n(x_1)\frac{\exp(ikx_2)}{\sqrt{2\pi}}+\Phi_n(x_2)\frac{\exp(ikx_1)}{\sqrt{2\pi}}\right\}, \tag{5}$$

where $\Phi_n(x)$ is the wave function describing the hydrogen atom in the nth stationary state, which satisfies the one-electron time-independent Schrödinger equation for an electron in a smoothed Coulomb potential (the energies E_n of the nine lowest states in the hydrogen atom are listed in Table I). This approximation is apparently permissible, at least for not very small k, because the potential created in space by the neutral hydrogen atom is short-range.

3. INTERACTION WITH AN ELECTROMAGNETIC FIELD

The dynamics of the system in an external laser field are described by the equation

$$i\hbar\frac{\partial\Psi(x_1,x_2,t)}{\partial t}=[H-e(x_1+x_2)\varepsilon(t)\cos(\omega t)]\Psi(x_1,x_2,t), \tag{6}$$

where ω is the frequency of the electric field of the wave and $\varepsilon(t)$ is the pulse envelope. The envelope $\varepsilon(t)$ used in the calculations had the form

$$\varepsilon(t)=\begin{cases}\varepsilon_0\sin^2\dfrac{\pi t}{2t_f}, & t\leqslant t_f,\\ \varepsilon_0, & t_f\leqslant t\leqslant t_f+t_p,\\ \varepsilon_0\sin^2\dfrac{\pi[t-(2t_f+t_p)]}{2t_f}, & t_f+t_p\leqslant t\leqslant 2t_f+t_p,\end{cases} \tag{7}$$

where t_f and t_p are the rise and decay times and the width of the plateau of the laser pulse. We set $t_f=2T$ and $t_p=5T$, where $T=2\pi/\omega$ is the optical period. The frequency values corresponded to $\hbar\omega=2$ and 5 eV.

For the numerical solution Eq. (6) was rewritten in the coordinates

$$\xi=(x_1+x_2)/\sqrt{2},\quad \eta=(x_1-x_2)/\sqrt{2}.$$

Since an electromagnetic field destroys the symmetry of the Hamiltonian with respect to spatial inversion, the integration region of (6) was chosen in the form

$$\xi\in(-\xi_{\max},\xi_{\max}),\quad \eta\in(0,\eta_{\max}),$$

$$\xi_{\max}=100\ \text{Å},\quad \eta_{\max}=150\ \text{Å}.$$

The number of points on the spatial grid was 810×540, and the integration step with respect to time was a hundredth of the period corresponding to $\hbar\omega=5$ eV. The method used to integrate the time-dependent Schrödinger equation was similar to the one in Ref. 11.

The function $\Psi(x_1,x_2,t)$ obtained as a result of the calculations was used to calculate the probabilities of the populating of various states and to calculate the photoelectron energy spectrum.

For example, the probability of finding the system in the unionized (original) state was calculated from the formula

$$W_0(t)=|C_0(t)|^2=\left|\int\Psi(x_1,x_2,t)\varphi_0(x_1,x_2)\,dx_1\,dx_2\right|^2, \tag{8}$$

where $\varphi_0(x_1,x_2)$ is determined from the solution of (2).

The one-electron ionization probability was calculated using the function

$$\tilde{\Psi}(x_1,x_2,t)=\Psi(x_1,x_2,t)-C_0(t)\varphi_0(x_1,x_2)\exp\left\{-\frac{i}{\hbar}E_0t\right\}, \tag{9}$$

where E_0 is the energy of the stationary state $\varphi_0(x_1,x_2)$.

As can be seen, $\tilde{\Psi}$ is formed as a result of the superposition of one- and two-electron continuum states and does not include the single bound state of the two-electron system. This situation does not diminish the errors in calculating the one-electron ionization probability associated with the completeness of the basis set of plane waves that we used to describe the one-electron continuum states.

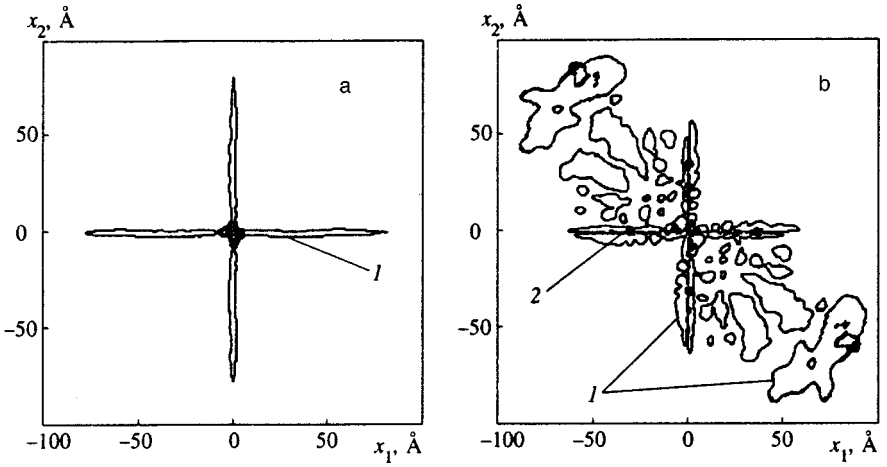


FIG. 2. Distribution of the probability density $\rho(x_1,x_2)$ at the end of a laser pulse for $\hbar\omega=5$ eV and $P=10^{13}$ (a) and 8×10^{14} W/cm^2. The contour levels correspond to the same values as in Fig. 1.

The coefficients in the expansion of the function (9) in the one-electron continuum states (5) were defined as

$$C_{nk}=\int \tilde{\Psi}(x_1,x_2,t)\varphi^*_{nk}(x_1,x_2)dx_1dx_2$$
$$=\sqrt{2}\int \tilde{\Psi}(x_1,x_2,t)\Phi_n(x_1)\frac{\exp\{-ikx_2\}}{\sqrt{2\pi}}$$
$$\times\exp\left\{\frac{i}{\hbar}(E_n+E)t\right\}dx_1\,dx_2 \qquad (1)$$

(here E_n and E are the energies of the excited nth one-electron bound state of hydrogen and of the free electron, $E=\hbar^2k^2/2m$), and the probability of one-electron ionization with passage of the bound electron into the nth state is

$$W_n=\int |C_{nk}|^2\,dk=2\int |\alpha_n(x)|^2\,dx, \qquad (11)$$

where

$$\alpha_n(x,t)=\int \tilde{\Psi}(x_1,x,t)\Phi_n(x_1)\,dx_1. \qquad (12)$$

The total one-electron ionization probability was determined by summing the W_n:

$$W^{(1)}=\sum_{n=1}^{n_{\max}} W_n. \qquad (13)$$

In our calculations we set $n_{\max}=18$. This value was selected on the basis of the condition that the populations of states with $n\geqslant n_{\max}$ are negligibly small in the range of radiation intensities investigated.

The two-electron ionization probability was determined as

$$W^{(2)}=1-W_0-\sum_{n=1}^{n_{\max}} W_n. \qquad (14)$$

If the two-electron ionization probability is small, the expansion coefficients C_{nk} permit determination of the photoelectron spectrum in wave-vector space:

$$w(k)=\sum_n |C_{nk}|^2=2\sum_n \left|\alpha_n(x)\frac{\exp\{-ikx\}}{\sqrt{2\pi}}dx\right|^2. \qquad (15)$$

Here the values of $|C_{nk}|^2$ specify the photoelectron spectra provided the hydrogen atom is in the nth stationary state when H^- undergoes photoionization. In free space the momentum expansion is identical to the energy expansion. Therefore, the expression (15), in effect, specifies the photoelectron energy spectrum $w(k=\sqrt{2mE}/\hbar)$. In this case (15) is normalized according to the condition

$$\int w(k=\sqrt{2mE}/\hbar)(dE/dk)^{-1}\,dE=1.$$

As for two-electron ionization, it is impossible to calculate the energy spectrum without knowledge of the functions in the two-electron continuum. In our opinion, the double momentum expansion performed for this purpose in Ref. 5 is incorrect, since the basis set of plane waves is complete and includes bound states of the electron, precluding separation of the one- and two-electron continuum states. Moreover, it should also be borne in mind that for a long-range potential, such as a one-dimensional smoothed Coulomb potential, momentum expansion is not identical to energy expansion and, therefore, does not permit calculation of the energy spectrum.

4. SIMULATION RESULTS

4.1. Space–time picture of the process

Figure 2 presents typical plots of the spatial distribution of the probability density $\rho=|\Psi(x_1,x_2)|^2$, which were obtained at the time of completion of the laser pulse for the intensity values $P=10^{13}$ W/cm^2 and 8×10^{14} W/cm^2. Figure 2a corresponds to one-electron ionization when one of the electrons (the identity principle does not allow us to say which one) is in a bound state, while the other is distant from the nucleus. Such a situation leads to the formation of the characteristic ''cross'' on the plot of $\rho(x_1,x_2)$. Figure 2b corresponds to an intensity at which both one- and two-electron ionization take place. In this case the electron density $\rho(x_1,x_2)$ localized in the region $\{x_1>0, x_2<0\cup x_1$

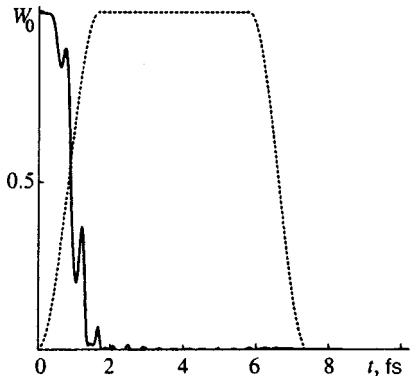


FIG. 3. Time dependence of the population of the ground state of H^- during a laser pulse for $\hbar\omega=5$ eV and $P=8\times10^{14}$ W/cm^2. The broken line is the laser pulse envelope.

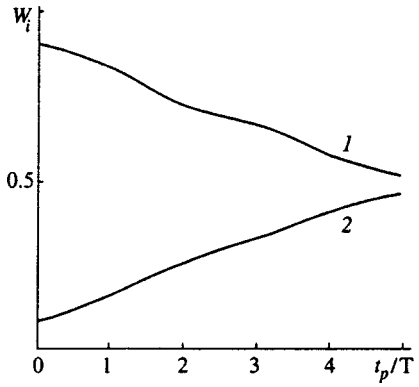


FIG. 4. Dependence of the one- (*1*) and two-electron (*2*) ionization probabilities at the end of a laser pulse on the width of the plateau of the laser pulse for $\hbar\omega=5$ eV and $P=8\times10^{14}$ W/cm^2.

$<0, x_2>0\}$ corresponds to two-electron ionization, i.e., in two-electron ionization the electrons escape in opposite directions for the most part as a result of Coulomb repulsion. We note that such correlated electron motion, which is obtained in exact calculations, is not observed in the Hartree or Hartree–Fock self-consistent-field approximations.

Spatial distributions of $\rho(x_1,x_2)$ obtained at different times enable us to distinguish between two mechanisms of two-electron ionization: a direct process, i.e., passage from the initial state to two-electron continuum states, and a sequential process. Sequential two-electron ionization involves passage from one-electron continuum states to two-electron continuum states under the condition that the saturation of single-electron ionization is observed as a result of essentially total ''depletion'' of the ground state. In fact, when $\hbar\omega=5$ eV and $P=8\times10^{14}$ W/cm^2, the population dynamics of the ground state during a pulse attest to the essentially complete emptying of this state already at the beginning of the plateau of the laser pulse (see Fig. 3).

Since the absorption of at least three photons is needed for a transition to the two-electron continuum when $\hbar\omega=5$ eV, while a transition to the one-electron continuum is possible as a one-photon process, the probability of direct two-photon ionization is low, and the sequential process is dominant. In this case it can be expected that increasing the width of the plateau of the laser pulse will reduce the one-electron ionization probability at the end of the pulse and increase the two-electron ionization probability via the sequential ionization mechanism. The data presented in Fig. 4 completely corroborate the major role of the sequential mechanism: as the width of the plateau is increased (at a fixed value of $t_f=2T$), there is a monotonic rise in the two-electron ionization probability from 7.6 to 47% mainly due a decrease in the percentage of one-electron ionization. Such a description of the process is qualitatively consistent with the data obtained in Ref. 4: a time-resolved analysis of the photoionization spectra for $\hbar\omega=27.2$ eV and $P\sim10^{16}$ W/cm^2 also provides evidence in support of the sequential mechanism of two-electron ionization.

Although the rate of direct two-electron ionization is small under these conditions, the probability of this process is nonzero at the end of the pulse. From the standpoint of the temporal evolution of the two-electron density $\rho(x_1,x_2)$ in the case of the direct mechanism of the process, in which both electrons, the outer and inner, pass into the continuum at essentially the same time and escape in opposite directions, the electron density should be localized in a region roughly equidistant from both coordinate axis $x_1=0$ and $x_2=0$. In the case of the sequential ionization process, the distribution of $\rho(x_1,x_2)$ should exhibit the ''cross'' corresponding to the single-electron ionization regime of the system, which gradually expands as the inner electron moves away from the system. The data presented in Fig. 5 for $P=8\times10^{14}$ W/cm^2, which correspond to the distribution of $\rho(x_1,x_2)$ at different times $t=2T$, $5T$, and $7T$, attest to the simultaneous presence of both mechanisms and, therefore, to competition between the direct and sequential ionization processes.

It would also be interesting to investigate the possibility of realizing the two-electron ionization mechanism[12] based on the rescattering model.[13] According to this model, the outer, weakly bound electron leaves the atom, and then it returns after half of an optical cycle and knocks out the inner electron with a certain probability. The data presented in Fig. 5 attest to delocalization of the electron density over a region with a radius many times greater than the atomic radius even at the beginning of the laser pulse. This applies to both one- and two-electron continuum states. In such a situation there is scarcely any justification to speak about the periodically repeated scattering of one of the electrons on the atomic core. We note in this context that calculations of the ionization dynamics of the one-dimensional He atom in Ref. 14, which were performed in the Hartree–Fock approximation and revealed a contribution of the rescattering effect to the two-electron ionization process, are in need of further interpretation.

4.2. Single- and two-electron ionization probabilities and stabilization

Let us move on to a study of the dependence of the one- and two-electron ionization probabilities on radiation intensity. The data for $\hbar\omega=5$ eV in the intensity range $P=10^{12}-2\times10^{15}$ W/cm^2 are presented in Fig. 6. This figure also shows the nonionization probability, i.e., the probability of

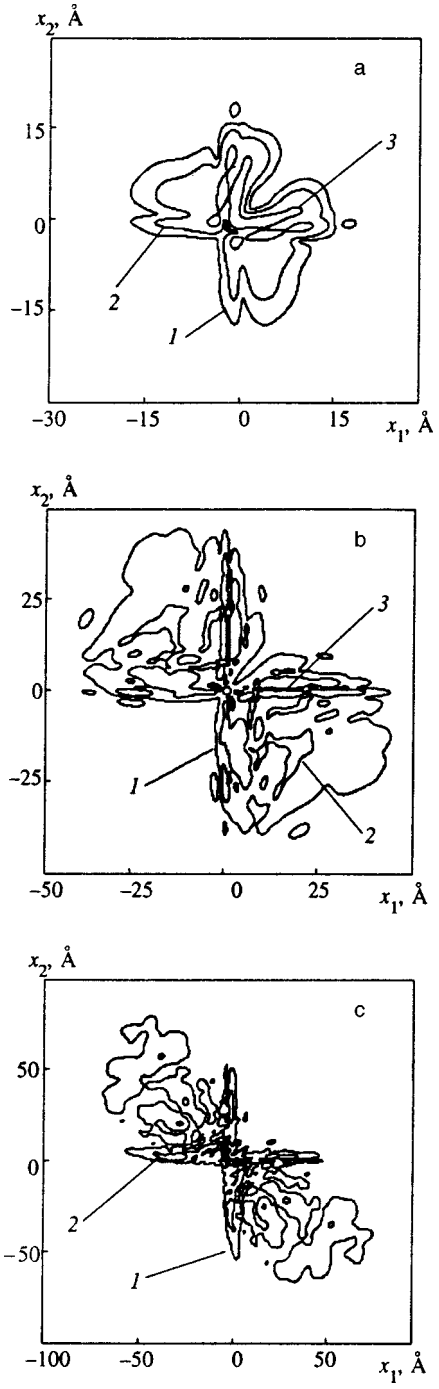


FIG. 5. Distribution of $\rho(x_1,x_2)$ for $\hbar\omega=5$ eV and $P=8\times10^{14}$ W/cm^2 at the times $t=2T$ (a), $5T$ (b), and $7T$ (c). The contour levels correspond to the same values as in Fig. 1.

remaining in a single bound state at the end of the laser pulse. As can be seen, double ionization begins at intensities equal to $\sim10^{14}$ W/cm^2, and the maximum for one-electron ionization is at 2×10^{14} W/cm^2. This value of the intensity is less than the stabilization threshold obtained in the one-dimensional[15] and three-dimensional[11] one-electron models of H^-, which amounts to $(4-5)\times10^{14}$ W/cm^2. This means that the study of the phenomenon of stabilization in H^- requires an investigation of the effect of a laser field on the inner electron.

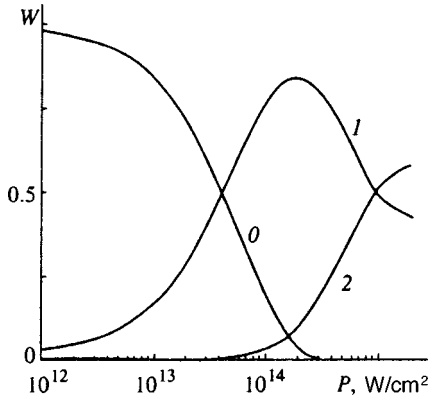


FIG. 6. Probability of remaining in the ground state (*0*) and one- (*1*) and two-electron (*2*) ionization probabilities at the end of a pulse as a function of radiation intensity for $\hbar\omega=5$ eV.

A comparison of the nonionization probabilities for the one-dimensional one- and two-electron models (see Fig. 7) shows that under our conditions the effect of the field on the inner electron is already significant at $\geqslant10^{13}$ W/cm^2 and leads to a significant increase in the probability of detachment of the outer, weakly bound electron. In addition, an increase in the nonionization probability (stabilization) is observed at $(4-8)\times10^{14}$ W/cm^2, but a rapid decrease in the value of $1-W_0$ is observed for $P\geqslant10^{15}$ W/cm^2. It was shown in Ref. 15 in the one-electron approximation that the appearance of stabilization under the conditions considered is due to the formation of a Kramers–Henneberger potential and the populating of stationary states, which are stable toward ionization, in it. Our calculations performed using a two-particle model allow us to state that the effect of the field on the inner electron significantly distorts the Kramers–Henneberger potential and eliminates stabilization.

We also note a tendency for saturation of the two-electron ionization probability in the region $P\geqslant10^{15}$ W/cm^2 (see Fig. 6). This saturation also appears to be a result of stabilization, but in the one-electron system, i.e., the hydrogen atom.

In the case under consideration ($\hbar\omega=5$ eV) three photons are needed to eject the inner electron, and two photons

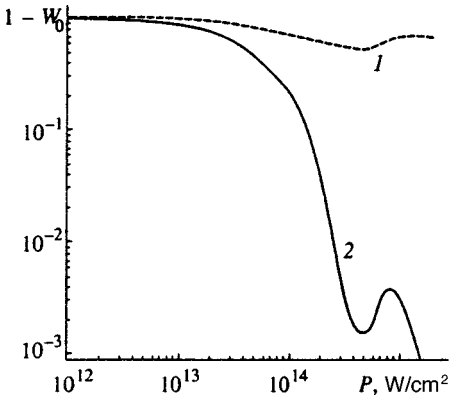


FIG. 7. Probability of remaining in the unionized state at the end of a laser pulse for the one-electron (*1*) and two-electron (*2*) models of H^- as a function of intensity for $\hbar\omega=5$ eV.

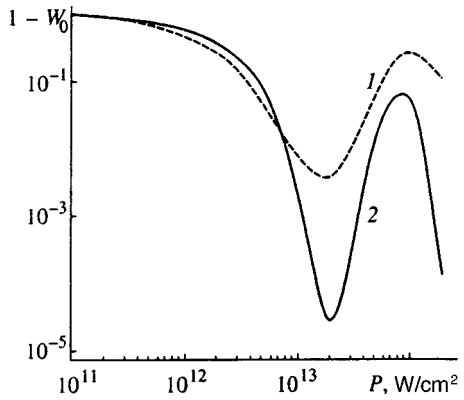


FIG. 8. Same as in Fig. 7, but for $\hbar\omega=2$ eV.

are needed to excite it. The low multiphoton order of the process is also manifested in the effective action of the field at once on the two electrons of the system and its unsatisfactory description within the one-electron model. Therefore, it can be expected that a decrease in the photon energy $\hbar\omega$ and a simultaneous increase in the multiphoton order of the processes for the inner electron (provided they remain one-photon processes for the outer electron) will yield better agreement between the calculations performed using the one- and two-electron models. The calculations of the nonionization probability for $\hbar\omega=2$ eV support this hypothesis (Fig. 8). Up to $P\sim10^{14}$ W/cm^2 the plots of $1-W_0$ qualitatively mimic one another, although stabilization is manifested even more clearly in the two-electron model: the nonionization probability increases by more than three orders of magnitude in the intensity range $2\times10^{13}-8\times10^{13}$ W/cm^2. Raising the intensity above 10^{14} W/cm^2 leads to two-electron ionization and a loss of the stabilization regime of the system.

4.3. Analysis of the photoelectron spectrum and populating of various bound states in the single-electron ionization regime

Figure 9 shows the populations of states of the hydrogen atom with various values of the quantum number n during the one-electron ionization of the negative hydrogen ion, which were calculated using (11) and (12) for $\hbar\omega=5$ eV and

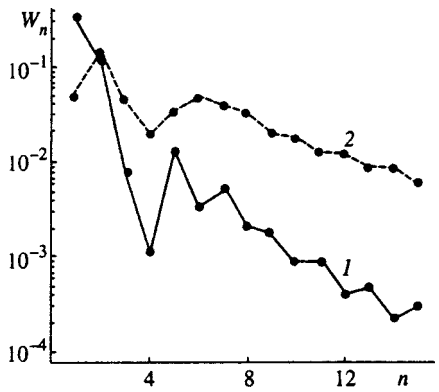


FIG. 9. Populations $\Phi_n(x)$ of various states of a one-dimensional hydrogen atom formed as a result of the one-electron photoionization of H^- for two values of the intensity: *1* — 4×10^{13} W/cm^2, *2* — 10^{15} W/cm^2.

various intensities. It is seen from the figure that the maximum probability is attained for $n=1$ and 2, which correspond to photodetachment of the outer electron. At the same time, there is also a maximum for $n=5$ and 6 for all the intensity values. This maximum may appear as a result of ejection of the inner electron by the field upon the absorption of three photons by the system. Furthermore, as can be seen from Fig. 9, when the radiation intensity is high, the probabilities of both ionization channels are close in value.

These arguments are qualitatively confirmed by the fact that the energies of the fifth and sixth stationary states of the one-dimensional hydrogen atom are fairly close to the binding energy of the outer electron in H^-. The possibility of photodetachment of both the outer and inner electrons in H^- was also considered in Ref. 4 in an analysis of high-energy photoelectrons, but no correlations between the position of the peak in the spectrum and the populations of different atomic states were obtained. Therefore,in addition to the totals we calculated the partial photoelectron energy spectra corresponding to residence of the bound electron in the nth stationary state.

The spectra for $P=2\times10^{14}$ W/cm^2 are shown in Fig. 10. As is seen from the figure, the photoelectron spectrum has a complicated structure because the electron remaining bound following the laser pulse can be in different states. For example, the series of peaks 1, 2, 3, ... is the result of the ionization of H^- with the formation of a hydrogen atom in the ground state, and peaks $1'$, $2'$, $3'$, ... appear as a result of the ionization and simultaneous excitation of the hydrogen atom formed to the $n=2$ state. The atomic system absorbs one additional photon in the latter case. The same number of photons are also absorbed in the case of ionization with excitation of the hydrogen atom to the $n=3$ state. The absorption of one more additional photon occurs for $n=6$, leading to an increase in the photoelectron energy in comparison to the case of $n=3$ (Fig. 10b). Thus, it is apparently convenient to interpret the case of $n=6$ as removal of the inner electron from H^-. We note that the position of the energy peaks in this case is not described exactly by the expression

$$E=l\hbar\omega+(E_0-E_n) \tag{16}$$

(l is the number of photons absorbed, and E_0 and E_n are the energies of the ground state of H^- and the nth excited state of the hydrogen atom) as a consequence of the Stark effect for the lower states of the hydrogen atom and the ponderomotive shift of the continuum edge.

At low intensities ($P\sim10^{13}$ W/cm^2) the shift of the levels due to the Stark effect is small, and the positions of the peaks are consistent with the calculations using Eq. (16). In this case photodetachment of the outer electron takes place, and the inner electron remains in the ground state for the most part. As a result, the photoelectron energy spectrum acquires the structure typical of the ionization of one-electron systems (Fig. 11), and the excitation probabilities of all the states with $n>2$ are negligibly small.

To conclude this section we note that the investigation of the spectra of photoelectrons formed as a result of two-electron ionization is of unquestionable interest. Such calculations were not performed because of the need to isolate the

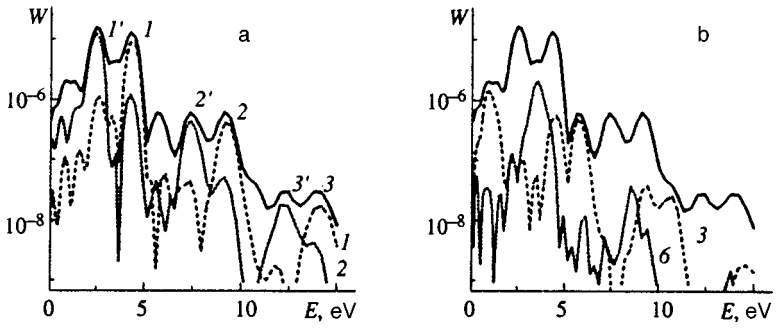


FIG. 10. Energy spectrum of photoelectrons formed in the one-electron photoionization of H^- for $P=2\times10^{14}$ W/cm^2. The numbers of the curves correspond to the states of the electron in the hydrogen atom formed. Solid curve — overall spectrum.

part of the wave function that corresponds to the two-electron continuum, which is a separate complicated problem. We note only that the general form of the wave function at the end of the laser pulse (see Fig. 2b) is evidence that the mean value of the kinetic energy of the photoelectrons produced by two-electron ionization is greater than the value for one-electron ionization.

5. CONCLUSION

Exact two-particle calculations of the ionization dynamics of the one-dimensional negative hydrogen ion in an electromagnetic field have been performed in the present work. The data obtained attest to strong electron–electron correlations in the system. The latter are manifested in the structure of the wave function of the ground state, which is characterized by weak binding of one of the electrons, as well as in the correlated two-electron ionization dynamics, which result in escape of the photoelectrons in opposite directions.

It has been concluded from the data obtained on two-electron ionization dynamics that the rescattering model, which involves the photodetachment of one electron and ejection of the second electron by it when it is scattered by the mother atom, is not feasible under the conditions considered.

It has been shown that the action of a strong field on the inner electron destroys the validity of the one-electron approximation. A comparison of the two-particle calculations carried out and calculations performed in the one-electron approximation has revealed the values of the parameters of the laser radiation (intensity and frequency) at which good correspondence between the two approaches compared is observed: a stabilization regime and agreement between the threshold intensity values for this regime in the two cases are observed.

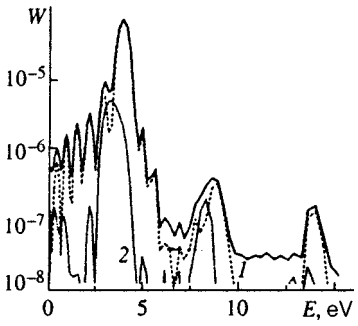


FIG. 11. Same as in Fig. 10, but for $P=10^{13}$ W/cm^2.

The features of the structure of the photoelectron spectrum in the region for one-electron ionization due to the superposition of the contributions of different channels of the process, viz., the photodetachment of one of the electrons and excitation of the remaining atom, have been explained.

In our opinion, it would be of interest to carry out similar investigations for a model neutral atom characterized by a large number of states in the discrete spectrum, including both singlet and triplet states, as well as autoionizing states. It would also be of interest to study the possibilities of using approximate Hartree–Fock methods to describe the features of the ionization dynamics of atomic systems in intense light fields.

This work was performed with financial support from the Russian Fund for Fundamental Research (Grant Nos. 96-02-19286 and 96-15-96447).

APPENDIX A:

We write the time-independent Schrödinger equation (2) in the form

$$\left(\frac{\partial^2}{\partial x_1^2}+\frac{\partial^2}{\partial x_2^2}-\frac{e^2}{\sqrt{\alpha^2+x_1^2}}-\frac{e^2}{\sqrt{\alpha^2+x_2^2}}+\frac{e^2}{\sqrt{\alpha^2+(x_1-x_2)^2}}\right)\varphi(x_1,x_2)=E\varphi(x_1,x_2). \quad \text{(A1)}$$

Taking into account the symmetry of the Hamiltonian under spatial inversion and interchange of the electrons, we introduce the new variables

$$\xi=\frac{x_1+x_2}{\sqrt{2}}, \quad \eta=\frac{x_1-x_2}{\sqrt{2}}, \quad \text{(A2)}$$

which correspond to a coordinate system turned 45° relative to x_1 and x_2. In the ξ,η coordinate system the problem (A1) can be written in the following form:

$$\left(\frac{\partial^2}{\partial \xi^2}+\frac{\partial^2}{\partial \eta^2}-\frac{e^2}{\sqrt{\alpha^2+2(\xi+\eta)^2}}-\frac{e^2}{\sqrt{\alpha^2+2(\xi-\eta)^2}}+\frac{e^2}{\sqrt{\alpha^2+2\eta^2}}\right)\varphi(\xi,\eta)=E\varphi(\xi,\eta). \quad \text{(A3)}$$

The function $\varphi(\xi,\eta)$ is symmetric or antisymmetric in the coordinate axes and can be defined in the first quadrant:

$$0\leqslant\xi,\quad \eta\leqslant\infty. \tag{A4}$$

To obtain the entire set of stationary states for (A1), we must consider the problem (A3) in the region (A4) for four cases with different boundary conditions on the coordinate axes:

$$\begin{aligned}
&1.\quad \left.\frac{\partial\varphi}{\partial\xi}\right|_{\xi=0,\,0\leqslant\eta\leqslant\infty}=0,\quad \left.\frac{\partial\varphi}{\partial\eta}\right|_{\eta=0,\,0\leqslant\xi\leqslant\infty}=0,\\
&2.\quad \left.\frac{\partial\varphi}{\partial\xi}\right|_{\xi=0,\,0\leqslant\eta\leqslant\infty}=0,\quad \varphi_{\eta=0,\,0\leqslant\xi\leqslant\infty}=0,\\
&3.\quad \varphi_{\xi=0,\,0\leqslant\eta\leqslant\infty}=0,\quad \left.\frac{\partial\varphi}{\partial\eta}\right|_{\eta=0,\,0\leqslant\xi\leqslant\infty}=0,\\
&4.\quad \varphi_{\xi=0,\,0\leqslant\eta\leqslant\infty}=0,\quad \varphi_{\eta=0,\,0\leqslant\xi\leqslant\infty}=0.
\end{aligned} \tag{A5}$$

The discrete analog of the problem (A3)–(A5) was obtained using the finite element method with a cubic approximation of the function in an element.[11] The two-dimensional region $0\leqslant\xi,\eta\leqslant 30$ Å was considered. The condition of equality of the derivative to zero was imposed on the distant boundary (the eigenfunctions for the states of the discrete spectrum of interest to us vanish on the distant boundary, and the actual assignment of the boundary condition is of no importance). The use of rectangular Lagrangian elements in a two-dimensional region enables us to write the following finite-element form of the problem:[16]

$$\mathbf{L}\varphi=E\mathbf{M}\varphi, \tag{A6}$$

where $\varphi=\{\varphi_1,\ldots,\varphi_n\}$ is the vector of the nodal values of the eigenfunctions, $N=KL$ is the number of nodes in the two-dimensional region, which is equal in the case of Lagrangian elements to the product of the number of nodes along each of the coordinates,

$$\mathbf{L}=\mathbf{M}_\xi\otimes\mathbf{D}_\eta+\mathbf{D}_\xi\otimes\mathbf{M}_\eta+\mathbf{P}$$

is the matrix of the Hamiltonian of the two-particle system, and

$$\mathbf{M}=\mathbf{M}_\xi\otimes\mathbf{M}_\eta$$

is the weighting matrix formed by the direct matrix product of the weighting matrices of the one-particle systems. The one-particle weighting matrix (or one-dimensional weighting matrix corresponding to one of the coordinates) is specified by a sum of element matrices:

$$\mathbf{M}_{\xi,\eta}=\sum_{k=1}^{K_e,L_e}\mathbf{m}^k_{\xi,\eta},$$

where K_e and L_e are the number of elements along each of the coordinates. For cubic elements we have $K=3K_e+1$, and $L=3L_e+1$.

The single-particle matrices of the derivatives $\mathbf{D}_{\xi,\eta}$ are defined in a similar manner.

The matrix elements are specified by the form functions in an element:[17]

$$\mathbf{m}^k_\xi=\int_{\xi_k}^{\xi_{k+1}}\mathbf{N}^{\mathbf{T}}\mathbf{N}d\xi,$$

$$\mathbf{d}^k_\xi=\int_{\xi_k}^{\xi_{k+1}}\frac{d\mathbf{N}^{\mathbf{T}}}{d\xi}\frac{d\mathbf{N}}{d\xi}d\xi,$$

where $\mathbf{N}=\{N_1,N_2,N_3,N_4\}$ is the vector of form functions for a cubic element.

The matrix **P** is defined by the interaction potential of the particles and was calculated using the following approximate formula

$$\mathbf{P}=\sum_{l=1}^{L_e}\sum_{k=1}^{K_e}\mathbf{m}^k_\xi\otimes\mathbf{m}^l\,\eta V_{k,l},$$

where

$$V_{k,l}=-\frac{e^2}{\sqrt{\alpha^2+2(\tilde\xi_k+\tilde\eta_l)^2}}-\frac{e^2}{\sqrt{\alpha^2+2(\tilde\xi_k-\tilde\eta_l)^2}}+\frac{e^2}{\sqrt{\alpha^2+2\tilde\eta_l^2}}$$

is the value of the interaction potential in the element with the indices k and l, and $\tilde\xi_k$ and $\tilde\eta_l$ are the mean values of the coordinates at the element.

Thus, the problem of finding stationary states has been reduced to the eigenvalue problem (A6). The boundary conditions 2–4 in (A5) are taken into account in the following manner: zero values of the function are assigned on the boundary. The off-diagonal elements of **L** in rows with the indices of nodes having zero assigned values are set equal to zero, and the diagonal elements in those rows are set equal to unity. The zero values of the derivatives are natural boundary conditions and do not require alteration of the matrix.

The eigenvalue problem was solved by iterating within a subspace.[18] The main goal of the iterative subspace approach is to find the p smallest eigenvalues and eigenvectors of the original problem. These eigenvectors form a p-dimensional orthogonal subspace. The effectiveness of the method is due to the fact that the entire subspace is iterated as a whole, rather than each eigenvector individually.

The matrix **L** must be positive definite to apply this method. In order to satisfy this condition, a displacement of the matrix must be introduced. We introduce the new matrix

$$\tilde{\mathbf{L}}=\mathbf{L}+s\mathbf{M},$$

choosing s so as to ensure the positive definiteness of $\tilde{\mathbf{L}}$. Then the new eigenvalue problem is written in the form

$$\tilde{\mathbf{L}}\tilde\varphi=\tilde E\mathbf{M}\tilde\varphi. \tag{A7}$$

It is easy to show that the solutions of (A6) and (A7) are related by the expressions

$$\tilde\varphi=\varphi,\quad \tilde E=E+s.$$

Subtracting the shift s from the modified values, we obtain the eigenvalues of the original matrix equation. In the problem under consideration a shift equal to $e^2/\sqrt{\alpha^2}$ was chosen.

The eigenvalue problem (A6) has a solution consisting of N eigenvalues and eigenvectors. However, we are interested only in a few low eigenvalues that correspond to the discrete spectrum of the Hamiltonian under consideration. In the formulation of the problem (A6) presented above these eigenvalues $E=\tilde{E}-s$ should be negative. Positive eigenvalues correspond to free motion of the electrons, and their discrete structure and finite number are specified by the finite size of the region considered and the discrete nature of the representation of the problem.

Therefore, the use of the proposed algorithm, which permits the determination of only the necessary number of eigenvalues, rather than all of them, seems very efficient. In addition, the algorithm under consideration permits working with a banded matrix **L** and makes it possible to use external memory.

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

[1] N. B. Delone and V. P. Krainov, *Multiphoton Processes in Atoms*, Springer-Verlag (1994).
[2] M. S. Pindzola, D. C. Griffin, and C. Bottcher, Phys. Rev. Lett. **66**, 2305 (1991); M. S. Pindzola, P. Gavras, and T. W. Gorczyca, Phys. Rev. A **51**, 3999 (1995).
[3] D. G. Lappas, A. Sanpera, J. B. Watson *et al.*, J. Phys. B **29**, L619 (1996).
[4] R. Grobe and J. H. Eberly, Phys. Rev. Lett. **68**, 2905 (1992).
[5] R. Grobe and J. H. Eberly, Phys. Rev. A **48**, 4664 (1993).
[6] A. I. Artemiev, R. Grobe, and J. H. Eberly, in *Super-Intense Laser-Atom Physics IV*, H. G. Muller and M. V. Fedorov (Eds.), Kluwer Academic, Dordrecht–Boston (1996), p. 285.
[7] J. Javanainen, J. H. Eberly, and Q. Su, Phys. Rev. A **38**, 3430 (1988).
[8] H. A. Bethe and E. Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, Springer, Berlin; Academic Press, New York (1957) [Russ. transl., Fizmatgiz, Moscow (1960)]; B. M. Smirnov, *Negative Ions*, McGraw-Hill, New York–London (1982) [Russ. original, Atomizdat, Moscow (1978)].
[9] K. Rzazewski, in *Super-Intense Laser-Atom Physics IV*, H. G. Muller and M. V. Fedorov (eds.), Kluwer Academic, Dordrecht–Boston (1996), p. 213.
[10] S. L. Haan, R. Grobe, and J. H. Eberly, Phys. Rev. A **50**, 378 (1994).
[11] E. A. Volkova, A. M. Popov, and O. V. Tikhonova, Zh. Eksp. Teor. Fiz. **108**, 436 (1995) [JETP **81**, 235 (1995)].
[12] K. C. Kulander, J. Cooper, and K. J. Schafer, Phys. Rev. A **51**, 561 (1995).
[13] P. B. Corkum, Phys. Rev. Lett. **71**, 1994 (1993).
[14] J. B. Watson, A. Sanpera, D. G. Lappas *et al.*, in *7th International Conference on Multiphoton Processes (ICOMP), Book of Abstracts*, Garmisch-Partenkirchen, Germany (1996).
[15] A. M. Popov, O. V. Tikhonova, and E. A. Volkova, Laser Phys. **5**, 1029, 1184 (1995).
[16] C. A. J. Fletcher, *Computational Galerkin Methods*, Springer-Verlag, New York (1984) [Russ. transl., Mir, Moscow (1988)].
[17] L. J. Segerlind, *Applied Finite Element Analysis*, Wiley, New York (1976) [Russ. transl., Mir, Moscow (1979)].
[18] H. Rutishauser, ''Computational aspects of F. L. Baner's simultaneous iteration method,'' Numer. Math. **13**, 4 (1969).

Translated by P. Shelnitz

Using effective operators in calculating the hyperfine structure of atoms

V. A. Dzuba and V. V. Flambaum

University of New South Wales, 2052 Sydney, Australia

M. G. Kozlov*) and S. G. Porsev†)

St. Petersburg Nuclear Physics Institute, 188350 Gatchina, Leningrad Region, Russia
(Submitted 16 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 1636–1645 (November 1998)

We propose a method for calculating the hyperfine structure (hfs) of multielectron atoms based on a combination of configuration superposition and many-body perturbation theory. The latter is used to construct an effective Hamiltonian and an effective hfs operator in configurational space. The method can be applied in calculations of the matrix elements of any one-electron operators. By way of an example we calculate the magnetic hfs constant A for several lowest levels of neutral thallium. We show that the method achieves a calculation accuracy of about 1%, which earlier was possible only for atoms with a single valence electron. © *1998 American Institute of Physics.* [S1063-7761(98)00611-8]

1. INTRODUCTION

Recently we proposed a method for calculating the lowest energy levels of multielectron atoms.[1] The calculations done for Tl (Ref. 1) and Ca, Sr, Ba, and Yb (Refs. 2 and 3) demonstrated its effectiveness. In this paper we wish to show that the method can be used to calculate not only energies but also other observables, such as the hyperfine structure (hfs) constants and transition amplitudes. For the sake of definiteness we focus on calculations of the magnetic dipole hfs constant. Generalization to other one-electron operators is obvious. Not that hfs calculations are one of the main tests in calculating amplitudes that do not conserve spatial parity.

At present several methods for calculating multielectron atoms are available. For atoms with one electron in addition to the electrons in filled shells there is the many-body perturbation theory in the residual Coulomb interaction (see, e.g., Ref. 4). For atoms with several valence electrons there is the configuration-superposition method and the multiconfiguration Hartree–Fock method.[5] Lately the coupled-cluster method has gained wide acceptance.[6–8] The hfs constants have been repeatedly calculated by all these methods (see, e.g., Refs. 9–14).

The most complicated problem encountered in atomic calculations is the need to correctly account for the correlations between valence electrons and the correlations incorporating core electrons. Correlations of the first type are too strong to be accounted for by ordinary many-body perturbation theory. However, if the number of the valence electrons is not too large, these correlations are taken into account fairly well by the configuration-superposition method or the multiconfiguration Hartree–Fock method. Correlations of the second type are accounted for more simply by many-body perturbation theory, since the number of configurations needed to describe these correlations by the configuration-superposition method or the multiconfiguration Hartree–Fock method is too large.

All this suggests a combination of the configuration-superposition method and many-body perturbation theory. The latter is used to set up an effective Hamiltonian for the valence electrons. After that the Schrödinger equation with the effective Hamiltonian can be solved by the configuration-superposition method. At this stage only valence electrons are accounted for explicitly. In Ref. 1 we showed that the Brillouin–Wigner perturbation theory and the ordinary diagrammatic technique are sufficient for constructing the effective Hamiltonian. In the present paper we discuss the setting-up of effective Hamiltonians for other observables.

In Sec. 2 we define the valence subspace and give the main formulas for the effective Hamiltonian. Section 3 is devoted to a discussion of other effective operators. In Sec. 4 we calculate the hyperfine structure in thallium.

2. EFFECTIVE HAMILTONIAN FOR VALENCE ELECTRONS

Here we are interested in low-energy atomic states with energies $E_i - E_0 < \varepsilon$, where E_0 is the ground-state energy of the atom. Then to a first approximation we can assume that the inner electrons, whose Hartree–Fock energies ε_n are much higher (in absolute value) than ε, form a core, which is described by the wave function

$$\Psi_{\rm core} = (N_c!)^{-1/2} \det(\phi_1, \phi_2, \ldots, \phi_{N_c}), \tag{1}$$

$$h_{\rm HFD} = \varepsilon_n \phi_n, \tag{2}$$

where $h_{\rm HFD}$ is the Hartree–Fock–Dirac operator. Although this operator is used to define the atomic core, it may incorporate the field of all the valence electrons or of several such electrons. For instance, below we examine thallium as an atom with three valence electrons and the core $[1s^2 \ldots 5d^{10}]$, while the $h_{\rm HFD}$ operator is set up for the $1s^2 \ldots 5d^{10}6s^2$ configuration (the V^{N-1} approximation, where N is the number of electrons in the atom).

We call the subspace of the multielectron states Ψ for which the core electrons are in state (1) the valence subspace and introduce the projector P on this subspace. The complementary subspace, with a corresponding projector $Q=1-P$, is characterized by the fact that at least one of the core electrons is excited to states lying higher than ϕ_{N_c}.

The Schrödinger equation

$$H\Psi = E\Psi \tag{3}$$

can be shown[1] to be equivalent to the following equation in the P space for the function $\Phi = P\Psi$:

$$[PHP + \Sigma(E)]\Phi = E\Phi, \tag{4}$$

$$\Sigma(E) = PV'R_Q(E)V'P, \tag{5}$$

where V' is the operator of the residual Coulomb interaction, and $R_Q(E)$ is the Green's function in the Q space, i.e.,

$$V' \equiv H - H_0, \tag{6}$$

$$R_Q(E) = Q\frac{1}{E-QHQ}Q. \tag{7}$$

The operator H_0 can be expressed in terms of the operator in Eq. (2) as follows:

$$H_0 = \sum_{i=1}^{N} h_{\rm HFD}(\mathbf{r}_i) - W. \tag{8}$$

The constant W appears on the right-hand side of this equation because the sum of single-particle energies ε_n allows for the electrostatic interaction of the electrons between each other twice and cannot serve as a good approximation for the total energy E of the atom. This constant can be found from the condition[1]

$$\langle \Psi_{\rm core}|H_0|\Psi_{\rm core}\rangle = \langle \Psi_{\rm core}|H|\Psi_{\rm core}\rangle \equiv E_{\rm core}, \tag{9}$$

or $\langle\Phi_0|H_0|\Phi_0\rangle = \langle\Phi_0|H|\Phi_0\rangle$, where Φ_0 is the ground-state wave function of the atom. Finally, this constant can be used as an adjustable parameter, selected by the best match between theory and experiment. Note that in the lowest perturbation order for the operator $\Sigma(E)$, the redefinition of this constant as $W \to W+\delta$ is equivalent to a shift in energy: $\Sigma(E) \to \Sigma(E+\delta)$.

The solutions of Eqs. (3) and (4) are related by

$$\Psi = [P + R_Q(E)V'P]\Phi, \tag{10}$$

which implies $\Phi = P\Psi$.

The orthonormalization condition $\langle\Psi_i|\Psi_k\rangle = \delta_{i,k}$ can be approximately reduced to a condition imposed on the functions Φ,

$$\langle\Phi_i|1 - \partial_E\Sigma(\overline{E})|\Phi_k\rangle \approx \delta_{i,k}, \tag{11}$$

where $\overline{E} \approx (E_i+E_k)/2$. Note that of the equations in (4)–(11) only the last is an approximation.

It is natural to call the operator in the square brackets in Eq. (4) the effective Hamiltonian $H_{\rm eff}$ for the valence electrons. Equations (4)–(7) make it possible to use this Hamiltonian for implementing the usual methods of many-body perturbation theory and, in particular, the diagrammatic technique.[1,2] To this end we use the representation of the exact Green's function in terms of the Green's function in the Hartree–Fock–Dirac representation:

$$R_Q(E) = R_Q^0(E) + R_Q^0(R)V'R_Q(E), \tag{12}$$

$$R_Q^0(E) = Q\frac{1}{E-QH_0Q}Q. \tag{13}$$

3. EFFECTIVE OPERATORS FOR VALENCE ELECTRONS

We assume that we know the solutions of Eq. (4), which we will use to find an observable a corresponding to the one-electron operator A:

$$a = \langle\Psi|A|\Psi\rangle. \tag{14}$$

We define an effective operator $A_{\rm eff}$ such that

$$a = \langle\Phi|A_{\rm eff}|\Phi\rangle. \tag{15}$$

Combining (10), (14), and (15), we arrive at an expression for $A_{\rm eff}$:

$$A_{\rm eff} = PAP + PV'R_Q(E)AP + PAR_Q(E)V'P + PV'R_Q(E)AR_Q(E)V'P. \tag{16}$$

Equations (12) and (16) allow the operator $A_{\rm eff}$ to be expanded in a power series in V'. Unfortunately, in most cases this series converges very slowly, so that usually instead of the consistent perturbation-theory approach one uses approximations that partially allow for all orders in V'. The random-phase approximation (RPA) is the one most often used in this case (see, e.g., Ref. 15). Let us see how this approximation agrees with Eq. (16).

On the right-hand side of Eq. (16), the operator A is always combined with the exact Green's function. We introduce a new operator $\tilde{A}$ such that

$$\begin{pmatrix} A_{PP} & A_{PQ}R_Q \\ R_QA_{QP} & R_QA_{QQ}R_Q \end{pmatrix} = \begin{pmatrix} \tilde{A}_{PP} & \tilde{A}_{PQ}R_Q^0 \\ R_Q^0\tilde{A}_{QP} & R_Q^0\tilde{A}_{QQ}R_Q^0 \end{pmatrix}, \tag{17}$$

where $A_{PP} \equiv PAP$, etc. If we were able to construct $\tilde{A}$, we would easily derive the operator $A_{\rm eff}$, since by substituting $\tilde{A}$ into Eq. (16) we arrive at a situation in which all exact Green's functions are replaced by Hartree–Fock Green's functions. Equation (17) is equivalent to the following system of operator equations:

$$\tilde{A}_{PP} = A_{PP}, \tag{18}$$

$$\tilde{A}_{PQ} = A_{PQ} + \tilde{A}_{PQ}R_Q^0V'Q, \quad \tilde{A}_{QP} = \tilde{A}_{PQ}^{\dagger}, \tag{19}$$

$$\tilde{A}_{QQ} = A_{QQ} + \tilde{A}_{QQ}R_Q^0V'Q + QV'R_Q^0\tilde{A}_{QQ} - QV'R_Q^0\tilde{A}_{QQ}R_Q^0V'Q. \tag{20}$$

The RPA equations for the core electrons (Fig. 1) resemble Eqs. (19) and (20). The main difference between the operator $\tilde{A}$ and the RPA operator $A_{\rm RPA}$ is that the former is not single-particle. Moreover, the random-phase approximation does not incorporate a number of single-particle corrections allowed by Eqs. (19) and (20). However, the most important terms of Eqs. (19) and (20) are taken into account by $A_{\rm RPA}$.

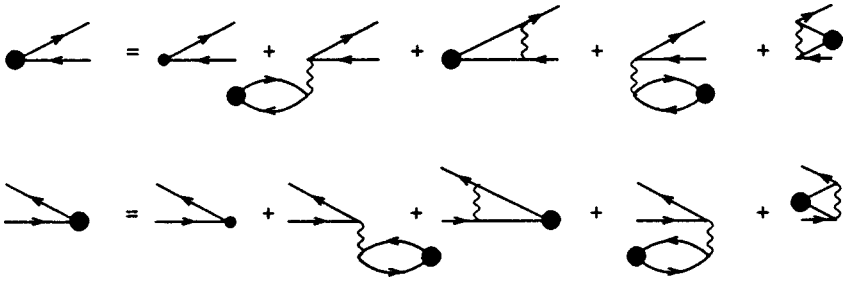


FIG. 1. Diagrammatic representation of the random-phase approximation method for the core. A small filled circle stands for the bare matrix element of the one-electron operator A, and a large filled circle corresponds to a similar matrix element in the random-phase approximation. The wavy line stands for the Coulomb interaction.

All this implies that we can write an approximate equation for the effective operator (16),

$$A_{\rm eff}\approx PAP+PV'R_Q^0(E)A_{\rm RPA}P + PA_{\rm RPA}R_Q^0(E)V'P. \tag{21}$$

The equation does not contain a term similar to the fourth term on the right-hand side of Eq. (16). This term is partially included in the second and third terms. The remaining part corresponds to what is known as structural radiation, which in most cases is very weak.[14]

Equation (21) also resembles the RPA equation for valence electrons. However, in addition to the RPA diagrams (Fig. 2a), this equation contains two additional types of diagram (Figs. 2b and 2c). The diagrams of Fig. 2b can be called subtractional (SBT) by analogy with similar diagrams for the operator Σ (Ref. 1). They appear only when the Hartree–Fock operator, which incorporates the field of (several) valence electrons, is used to solve the RPA equations. It should be recalled that such an operator is used in setting up the core wave function (1). The diagrams in Fig. 2c correspond to the two-particle corrections $A_{\rm TP}$ to the effective operator. There is also an important type of diagram (Fig. 2d) not included in the approximation (21). Such diagrams refer to a higher order in the many-body perturbation theory, but their contribution is extremely large, while similar diagrams in which the $A_{\rm RPA}$ vertex is connected to the other part of the diagram by a particle line rather than by a hole line are taken into account when (21) is substituted in (15).

Combining all these corrections, we arrive at an approximate expression for the effective operator:

$$A_{\rm eff}\approx P(A_{\rm RPA}+A_{\rm SBT}+A_{\rm TP}+A_\sigma)P, \tag{22}$$

where the four terms on the right-hand side correspond to the four types of diagram in Figs. 2a–2d. Below, in calculating the hfs constants, we use Eqs. (15) and (22). We note once more that this approximation allows only for the first order of many-body perturbation theory. Even in second order there are corrections of the structural-radiation type, which are not included in (22). On the other hand, Eq. (22) takes into account some of the most important higher-order corrections.

To conclude this section we note that the operator $A_{\rm eff}$ is used to solve Eq. (4). This means that excitations of valence electrons are taken into account in all orders, which guarantees an accuracy higher than that achieved by many-body perturbation theory.

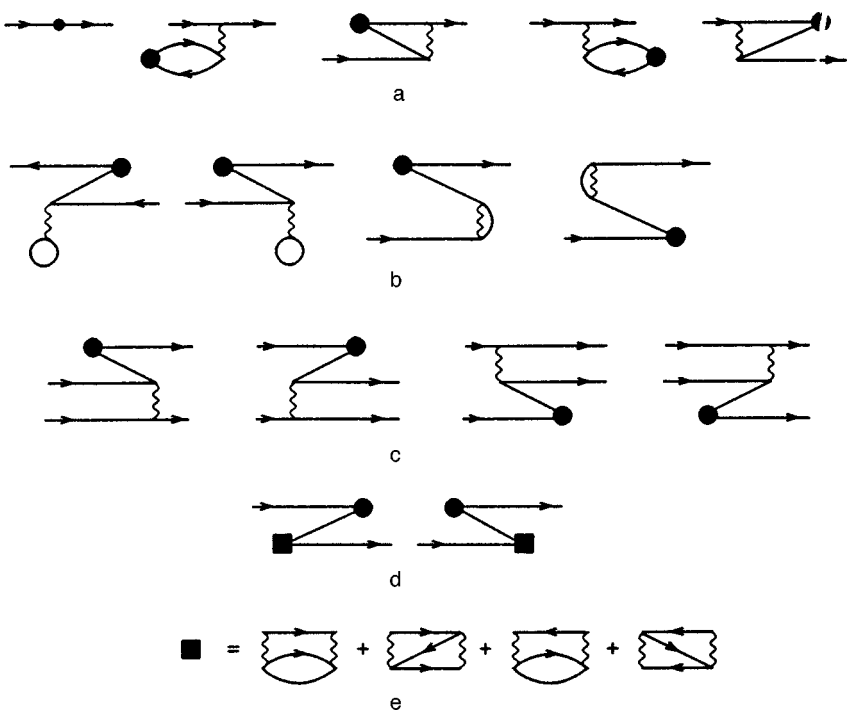


FIG. 2. Diagrammatic representation of the effective operator $A_{\rm eff}$ for the valence electrons: (a) the random-phase approximation $A_{\rm RPA}$, (b) the subtractional correction $A_{\rm SBT}$, (c) the two-particle correction $A_{\rm TP}$, and (d) the self-energy correction A_σ. A filled square stands for the self-energy block (e).

TABLE I. Valence energies of several lowest levels of Tl.

Level	$E_{\rm val}$, a.u.	ΔE, $\rm cm^{-1}$	
		Calculation	Experiment[18]
$6p_{1/2}$	−2.072 084	0	0
$6p_{3/2}$	−2.036 471	7808	7793
$7s_{1/2}$	−1.951 435	26 472	26 478
$7p_{1/2}$	−1.916 363	34 169	34 160
$7p_{3/2}$	−1.911 804	35 170	35 161
$6d_{3/2}$	−1.907 348	36 148	36 118
$6d_{5/2}$	−1.907 011	36 222	36 200

Note: The effective Hamiltonian was set up for the energy $E_{\rm val}=-1.64$ a.u., which was chosen on the grounds of best agreement between that the calculated energy intervals between the levels and the experimental values of these intervals.

4. CALCULATIONS OF THE HYPERFINE STRUCTURE

In the point-nucleus approximation, the operator of the magnetic hyperfine structure has the form

$$H_{\rm hfs}=\frac{\mu_n\mu_0}{I}\,\frac{\mathbf{I}\cdot(\boldsymbol{\alpha}\times\mathbf{r})}{r^3}, \tag{23}$$

where μ_n is the magnetic moment of the nucleus, μ_0 is the Bohr magneton, and $\boldsymbol{\alpha}$ is the Dirac matrix. For a finite nucleus, this expression can be used only if $r>r_n$ holds, where r_n is the radius of the nucleus. Inside the nucleus the hfs operator depends on the nuclear structure rather than on the magnetic moment of the nucleus. The exact formulas can be found in Ref. 13. There it is also shown that the corresponding corrections are extremely small, so that here we use an approximation in which inside the nucleus the dependence of the operator (23) on the radius, $1/r^3$, is replaced by r/r_n^4. Such an approximation is quite sufficient if we need an accuracy of about 1%.

We begin our calculations of the hyperfine structure of the thallium atom by constructing an effective Hamiltonian and solving Eq. (4). This part of our calculations differs somewhat from the procedure adopted in Ref. 1. We enlarged the basis set of the radial functions, which now includes orbitals up to $21s$, $21p$, $21d$, $18f$, $18g$, and $14h$ (the procedure of constructing orbitals is similar to the one described in Refs. 1, 16, and 17). We also significantly increased the number of configurations. Moreover, in calculating the diagrams for the operator $\Sigma(E)$ we also calculated their first energy derivatives. This made it possible to allow in the first approximation for the energy dependence of the matrix elements of the effective Hamiltonian. The thallium spectrum calculations employed an effective Hamiltonian for different valence electrons, which, as noted earlier, corresponds to different constants W in (8). We found that the agreement with the experimental spectrum is best when $\Sigma(E)$ is taken at an energy $E_{\rm val}=E-E_{\rm core}=-1.64$ a.u. (see Table I).[1)] What is important is that the low-energy part of the spectrum of neutral thallium was well reproduced by these calculations and the valence energy of the ground state was found to coincide perfectly with the experimental value of 2.0722 a.u.[19]

The next step involved constructing the effective operator (22). All corrections except $A_{\rm TP}$ are determined by single-particle operators and can be reduced to calculating effective one-electron integrals. Calculating the expectation value (15) for such operators is easy. However, allowing for the two-electron correction $A_{\rm TP}$ required more involved calculations. Hence in calculating these corrections we discarded the contributions of configurations whose weight in the wave function was no larger than 10^{-4}.

To determine the error associated with the incompleteness of the configuration-superposition method, we estimated the contribution of virtual orbitals not included in the superposition of configurations in the lowest perturbation order. We call the corresponding corrections the valence corrections. These corrections are important only for the $6p_{3/2}$ and $7p_{3/2}$ levels, where the other corrections are small or cancel each other almost perfectly.

Finally, the last correction arises if we allow for the conditions (11). For the lowest levels with fixed angular momentum J and parity P the conditions simply determine the corrections to normalization. For excited states, strictly speaking, we must allow for violation of orthogonality with the lowest states, but here we took into account only the normalization condition. The results of calculating the hyperfine structure of thallium are listed in Table II.

5. CONCLUSION

A comparison of the results listed in Table I and those of Ref. 1 shows that even approximate allowance for the energy dependence of the operator Σ combined with an optimum choice of the constant W, which determines the staring approximation (8) of the Hamiltonian, make it possible to enhance the accuracy of calculations of the atomic spectrum. The largest deviation from the experimental data in such calculations amounts only to 0.2%, while in the previous calculations it amounted to about 1%. This refinement of the spectrum was found to have an effect on the results of hfs calculations.

The results listed in Table II show that the Hartree–Fock–Dirac approximation for the magnetic hfs constant agrees very poorly with the experiment. In some cases it even yields an incorrect sign or an incorrect order of magnitude of the constant. Allowing only for configuration superposition does not improve the results significantly.

Only when many-body perturbation theory and the configuration-superposition method are combined does the accuracy improve substantially. Many-body perturbation theory yields three types of correction, resulting from (i) replacing the valence Hamiltonian by the effective Hamiltonian (allowance for $\Sigma(E)$), (ii) using the effective hfs operator ($A_{\rm SBT}$,$A_{\rm TP}$, and A_σ), and (iii) the normalization condition (11). The last correction is less than 1%, while the other two are much larger.

Note that the various many-body perturbation-theory corrections often cancel one another because summation over intermediate states leads to contributions that violate the Pauli exclusion principle. Such contributions, obviously, cancel each other, so that their net effect is zero. For in-

TABLE II. Magnetic hfs constants for several lowest levels of ^{205}Tl (MHz).

	$6p_{1/2}$	$6p_{3/2}$	$7s_{1/2}$	$7p_{1/2}$	$7p_{3/2}$	$6d_{3/2}$	$6d_{5/2}$
HFD	17 554	1302	7612	1957	188	21	9
CS	+195	−1369	+3655	−88	+114	−149	+307
$\Sigma(E)$	+3197	+48	+421	+290	−41	+90	−135
RPA	+1359	+327	+1043	+134	+71	+6	+14
$A_{\rm SBT}$	−1225	−120	−72	−116	−17	−16	−11
$A_{\rm TP}$	+1130	+44	−10	+105	+6	+12	+8
A_σ	−1090	−19	−303	−105	−10	+2	−6
Valence	−13	+53	+43	−6	+6	−1	−1
Normalization	−214	−2	−82	−14	−2	0	−1
Theory of Ref. 14	21 760	−1919	12 470	2070	195		
Theory of Ref. 13	21 300	339	12 760				
Present work	21 623	264	12 307	2157	315	−35	184
Experiment	21 311	265	12 297	2155	309	−43	229

Note: We list the values of the hfs constants in the Hartree–Fock–Dirac (HFD) approximation and the correction yielded by configuration superposition (CS) and the correction Σ to the Hamiltonian. We allow for the corrections to the hfs operator arising from the use of the random-phase approximation (RPA), the subtractional correction $A_{\rm SBT}$, the two-particle correction $A_{\rm TP}$, the self-energy correction A_σ, the valence correction reflecting the incompleteness of configuration substitution, and the normalization correction.

stance, the sum $A_{\rm RPA}+A_{\rm SBT}+A_{\rm TP}$ obeys the Pauli exclusion principle, while each term in it does not. For this reason the separate contributions to this sum have no physical meaning. Moreover, in most cases there is partial balance between the correction due to the use of an effective Hamiltonian and the correction A_σ. In our case this balance results from the fact that the contributions of the intermediate electron and hole state have different signs but comparable values. As a result, the net many-body perturbation-theory correction is much smaller than the separate contributions.

The data in Table II suggest that for all the levels except the d levels the agreement with experiment is extremely good. The largest many-body perturbation-theory corrections arise for the $6p_{1/2}$ level, where they amount to more than 3000 MHz. This explains the lower accuracy achieved in calculating this constant (1.4%). Two other large constants (for the $7s$ and $7p_{1/2}$ levels) were found to be calculated with an accuracy ten times higher, which corresponds to appreciably smaller values of the many-body perturbation-theory corrections. For the $6p_{3/2}$,$6d_{3/2}$, and $6d_{5/2}$ levels we have an extremely poor starting approximation. Hence the perfect agreement between our results for the $6p_{3/2}$ level and the experimental data can to a certain extent be considered a coincidence. Note that to obtain a correct result we needed to allow for all the corrections here.

On the whole we can say that the adopted method proved to be very effective in calculating the spectrum and the hfs constants. Note that in a somewhat simplified form this method has recently been used to calculate the hfs constants and P,T-odd matrix elements for the BaF molecule,[20] with the accuracy of this method also increasing significantly. Further refinements require increasing the number of higher-order corrections taken into account. This can be done by resorting to the methods used earlier in calculations for the cesium and thallium atoms by a pure perturbation-theory approach.[14,21] First one must allow for higher-order corrections to the operator Σ. The leading corrections of this type correspond to the inclusion of a polarization operator in phonon lines. To first order these corrections can be taken into account by introducing screening coefficients (see, e.g., Ref. 3). One must also allow for the correction to the effective hfs operator related to structural radiation.

This work was made possible by a grant from the Russian Fund for Fundamental Research (Grant 98-02-17663).

*)E-mail: mgk@mfl309.spb.edu
†)E-mail: porsev@thd.pnpi.spb.ru

1)The problem of selecting the energy is discussed in greater detail in a paper submitted to the journal Optika i Spektroskopiya (translated as Optics and Spectroscopy).

[1] V. A. Dzuba, V. V. Flambaum, and M. G. Kozlov, JETP Lett. **63**, 882 (1996); Phys. Rev. A **54**, 3948 (1996).
[2] M. G. Kozlov and S. G. Porsev, Zh. Éksp. Teor. Fiz. **111**, 838 (1997) [JETP **84**, 461 (1997)].
[3] V. A. Dzuba and W. R. Johnson, Phys. Rev. A **57**, 2459 (1998).
[4] I. Lindgren and J. Morrison, *Atomic Many-Body Theory*, Springer, Berlin (1985).
[5] I. P. Grant and H. M. Quiney, Adv. At. Mol. Phys. **23**, 37 (1988).
[6] S. A. Blundell, W. R. Johnson, and J. Sapirstein, Phys. Rev. A **43**, 3407 (1991).
[7] E. Ilyabaev and U. Kaldor, Phys. Rev. A **47**, 137 (1993).
[8] E. Eliav, U. Kaldor, and Y. Ishikawa, Phys. Rev. A **49**, 1724 (1994); **50**, 1121 (1994); **51**, 225 (1995).
[9] I. Lindgren, Rep. Prog. Phys. **47**, 345 (1984).
[10] T. Olsson, A. Rosen, B. B. Fricke, and G. Torbohm, Phys. Scr. **37**, 730 (1988).
[11] D. R. Beck and D. Datta, Phys. Rev. A **48**, 182 (1993).
[12] P. Jönsson and C. Froese Fisher, Phys. Rev. A **48**, 4113 (1993).
[13] A.-M. Mårtensson-Pendrill, Phys. Rev. Lett. **74**, 2184 (1995).
[14] V. A. Dzuba, V. V. Flambaum, O. P. Silvestrov, and O. P. Sushkov, J. Phys. B **20**, 1399 (1987).
[15] A.-M. Mårtensson-Pendrill, in: *Methods of Computational Chemistry*, Vol. 5: *Atomic and Molecular Properties*, S. Wilson (ed.), Plenum Press, New York (1992).
[16] P. Bogdanovich and G. Žukauskas, Sov. Phys. Collect. **23**, 13 (1983).

[17] P. Bogdanovich, G. Žukauskas, and S. Šandžiuvienė, Sov. Phys. Collect. **24**, 20 (1984).
[18] C. E. Moore, *Atomic Energy Levels*, Circ. No. 467 Nat. Bureau Standards USA, Washington, D.C. (1958).
[19] A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions*, Springer, Berlin (1985).
[20] M. G. Kozlov, A. V. Titov, N. S. Mosyagin, and P. V. Souchko, Phys. Rev. A **56**, R3326 (1997).
[21] V. A. Dzuba, V. V. Flambaum, and O. P. Sushkov, Phys. Lett. A **141**, 147 (1989).

Translated by Eugene Yankovsky

Inelastic processes in collisions of multicharged fast ions with atoms

V. I. Matveev*) and Kh. Yu. Rakhimov

Department of Thermal Physics, Uzbek Academy of Sciences, 700135 Tashkent, Uzbekistan
(Submitted 16 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 1646–1660 (November 1998)

We use the eikonal approximation to develop a general formula for the cross sections of inelastic collisions of multicharged fast ions (including relativistic ions) with atoms that is applicable within a broad range of collision energies, has the standard nonrelativistic limit, and becomes, in the ultrarelativistic case, the well-known result that follows from the exact solution of the Dirac equation. As an example we study the excitation and ionization of a hydrogenlike atom, the single and double excitation and ionization of a heliumlike atom, and multiple (up to the eighth order) ionization of the neon atom and (up to eighteenth order) ionization of the argon atom. We derive simple analytical expressions for the inelastic cross sections and establish recurrence relations linking the cross sections of ionization of different orders. Finally, we compare our results with the experimental data. © *1998 American Institute of Physics.* [S1063-7761(98)00711-2]

1. INTRODUCTION

The primary reason for the considerable interest in inelastic processes accompanying the collisions of atoms with multicharged ions is that the effective strengths of the fields generated by the ions over atomic distances can exceed those of the characteristic internal fields by several orders of magnitude. Such fields are very difficult to generate by other means. Thus, collision experiments involving multicharged ions are actually the only way to study the behavior of atoms and molecules in ultrahigh electromagnetic fields. From the fundamental viewpoint, research into the behavior of matter in ultrahigh electromagnetic fields constitutes one of the most important problems of modern physics.

Moreover, a number of problems of applied science, such as measurements of the energy spectra of nuclear fission fragments, synthesis of superheavy elements, interpretation of data on ultrahard cosmic rays, pumping active media of high-power lasers, ion diagnostics and spectroscopy of plasma, implantation and sputtering of solids by ion bombardment, and design and application of high-energy heavy-ion accelerators, have stimulated research in the field of collisions involving multicharged fast ions. Studies in inelastic collisions of fast (including relativistic) multicharged ions with atoms have lately attracted a lot of attention (see, e.g., Refs. 1 and 2, recent reviews in Refs. 3–5, and the literature cited therein). The cross sections of inelastic processes in collisions of this type are very large, and for this reason these studies are of interest from the practical viewpoint. The strong field of an ion with a large charge makes it impossible (even when the velocities of the colliding particles are relativistic) to use perturbation-theory techniques, which complicates calculations substantially, since nonperturbative approaches are required.

Examples abound. We would like to mention Refs. 6–9, which use the sudden-perturbation approximation, Refs. 10–12, which use the eikonal approximation and its modifications, calculations of Becker *et al.*[13] based on the numerical solution of the time-dependent Dirac equation, and the recently found exact solution[14] of the Dirac equation in the ultrarelativistic limit.

The most general basis for studying the cross sections of inelastic processes involving collisions of multicharged fast ions with atoms is the Glauber approximation,[15] which is valid for $Z/v \ll 1$, where Z is the ion's charge and v is the collision velocity (atomic units). This approximation is based on the eikonal approximation,[16,17] which is close to the semiclassical domain, and is closely related to the sudden-perturbation approximation.[4,6,12] The use of perturbation theory in describing collisions of multicharged ions with atoms when $Z/v \sim 1$ violates the range of applicability of the Born approximation and, due to its nonunitary nature, to values of probability greater than unity.[5] We also note that for ions with very large charges the applicability range ($Z/v \ll 1$) of the Born approximation is never reached, no matter how high the collision energies are.

According to Ref. 15, the eikonal approximation[18] for relativistic potential scattering can be generalized (see, e.g., Ref. 12) to the case of an inelastic collision of a low-Z atom (nonrelativistic before and after the collision) and an ion moving with a relativistic velocity v. Then, in the Glauber approximation, the general expression for the amplitude of the inelastic collision accompanied by the transition of the atom from state $|\Psi_i\rangle$ to state $|\Psi_f\rangle$ has the form[12] (cf. Ref. 15)

$$f_{if}(\mathbf{q}) = \frac{ik_i}{2\pi} \int e^{-i\mathbf{q}\cdot\mathbf{b}} \langle \Psi_f | \left[1 - \exp\left\{ -\frac{i}{v} \int U\, dx \right\} \right] | \Psi_i \rangle\, d^2b, \tag{1}$$

where the momentum transfer $\mathbf{q} = \mathbf{k}_f - \mathbf{k}_i$. The scattering potential $U = U(x, \mathbf{b}; \{\mathbf{r}_a\})$ depends not only on the ion position $\mathbf{R} = (x, \mathbf{b})$ but also on the instantaneous positions of the atomic electrons, whose sets of coordinates we denote by $\{\mathbf{r}_a\}$, $a = 1,2, \ldots, N$, with N the number of electrons.

In this paper we use the eikonal approximation to derive a general formula for the cross sections of inelastic collisions of multicharged fast ions (including relativistic ions) with atoms that is applicable within a broad range of collision energies, has the standard nonrelativistic limit, and becomes, in the ultrarelativistic case, the well-known result that follows from the exact solution[14] of the Dirac equation. As an example we study the excitation and ionization of a hydrogenlike atom, the single and double excitation and ionization of a heliumlike atom, multiple (up to the eighth order) ionization of the neon atom and (up to eighteenth order) ionization of the argon atom. We derive simple analytical expressions for the inelastic cross sections and establish recurrence relations linking the cross sections of ionization of different orders. We also compare our results with the experimental data.

2. GENERAL CONSIDERATIONS

To generalize the eikonal approximation to the case of a collision of a relativistic ion and a high-Zrelativistic atom one must allow for the following: (a) the behavior of the atomic electrons are described by the Dirac equation, and (b) by definition the Glauber approximation to the scattering potential $U=U(x,\mathbf{b};\{\mathbf{r}'_a\})$ is the static Coulomb potential generated by the atomic nucleus and atomic electrons in fixed (and simultaneous, from the viewpoint of the incident ion) positions $\mathbf{r}'_a=x'_a,y'_a,z'_a$:

$$U(x,\mathbf{b};\{\mathbf{r}'_a\})=\frac{Z_aZ}{\sqrt{x^2+b^2}}-\sum_{a=1}^{N}\frac{Z}{\sqrt{(x-x'_a)^2+(\mathbf{b}-\mathbf{s}'_a)^2}},$$

where Z_a is the charge of the atomic nucleus, with $Z_a=N$, and

$$\frac{1}{v}\int_{-\infty}^{\infty}U\,dx=\sum_{a=1}^{N}\chi_a(\mathbf{b},\mathbf{s}'_a),\qquad \chi_a(\mathbf{b},\mathbf{s}'_a)=\frac{2Z}{v}\ln\frac{|\mathbf{b}-\mathbf{s}'_a|}{b},$$

with the x axis directed along the initial momentum $\mathbf{k}_i$ of the ion, and the two-dimensional vector $\mathbf{s}'_a$ is (y'_a,z'_a). Let us assume, for the sake of definiteness, that we are dealing with simultaneous positions of electrons at time $t'=0$ in the reference frame co-moving with the ion; the corresponding instantaneous positions of the atomic electrons are $\mathbf{r}'_a$ and the electron wave function is $\Psi'(\mathbf{r}'_a,t')$. Then instead of (1) we have

$$f_{if}(\mathbf{q})=\frac{ik_i}{2\pi}\int e^{-i\mathbf{q}\cdot\mathbf{b}}\langle\Psi'_f|\left[1-\exp\left\{-\frac{i}{v}\int U\,dx\right\}\right]|\Psi'_i\rangle\,d^2b.$$

In the reference frame in which the atom is at rest (at $t'=0$),

$$x_a=\gamma x'_a,\quad \mathbf{s}_a=\mathbf{s}'_a,\quad t=-x_a\frac{v}{c^2},$$

$$\Psi(\mathbf{r}_a,t)=\psi(\mathbf{r}_a)\exp(-iEt)=\psi(\mathbf{r}_a)\exp\left(iEx_a\frac{v}{c^2}\right)$$

$$=S_a^{-1}\Psi'(\mathbf{r}'_a,t'=0),$$

$$d^3r_a=dx_a\,dy_a\,dz_a=\gamma\,dx'_a\,dy'_a\,dz'_a=\gamma\,d^3r'_a,$$

where $\gamma=1/\sqrt{1-v^2/c^2}$, and S_a is the Lorentz transformation matrix for the wave function, acting on the bispinor indices that refer to the atomic electron with the number a (the respective Dirac matrices are α_a), with $S_a^{-2}=\gamma(1-\mathbf{v}\alpha_a/c)$ (see Ref. 2). Hence in the Glauber approximation the general expression for the probability amplitude of an inelastic collision of a relativistic ion and a high-Z relativistic atom accompanied by the transition of the atom from state $|\psi_i\rangle$ with energy E_i to state $|\psi_f\rangle$ with energy E_f has the form

$$f_{if}(\mathbf{q})=\frac{ik_i}{2\pi}\int\langle\psi_f|\left[1-\exp\left\{-i\sum_{a=1}^{N}\chi_a(\mathbf{b},\mathbf{s}_a)\right\}\right]\gamma^{-N}S^{-2}$$

$$\times\exp\left[i\sum_a\frac{vx_a}{c^2}(E_f-E_i)\right]|\psi_i\rangle\exp(-i\mathbf{q}\cdot\mathbf{b})\,d^2b,\qquad(2)$$

where $S^2=\Pi_{a=1}^{N}S_a^{-2}$. This is the final expression corresponding to the generalization of Eq. (1) to the case of a collision of a relativistic ion and a high-Z relativistic atom obeying the same applicability conditions: the collision time must be much shorter than the characteristic atomic time. If we are not interested in the scattering angles of the ion, then, in accordance with the expression for the cross section in the small-angle approximation (as in Ref. 15), we can integrate over them: for small scattering angles we have

$$d\Omega\approx\frac{d^2q}{k_ik_f}\approx\frac{d^2q}{k^2}.$$

Then, representing $|f_{if}|^2$ in (2) by a double integral in d^2b and d^2b', we integrate in d^2q using the integral representation of the delta function, which is then removed by integrating in d^2b'. As a result we arrive at a formula for the cross section of the atomic transition from state $|\psi_i\rangle$ to state $|\psi_f\rangle$ in the collision of the atom and the relativistic ion:

$$\sigma=\int d^2b\left|\langle\psi_f|\left[1-\exp\left\{-i\sum_{a=1}^{N}\chi_a(\mathbf{b},\mathbf{s}_a)\right\}\right]\right.$$

$$\left.\times\gamma^{-N}S^{-2}\exp\left[i\sum_a\frac{vx_a}{c^2}(E_f-E_i)\right]|\psi_i\rangle\right|^2.\qquad(3)$$

Accordingly, the integrand can be interpreted as the probability of the atom going from state $|\psi_i\rangle$ to state $|\psi_f\rangle$ in a collision with the impact parameter $\mathbf{b}$. In this form, for a one-electron atom this probability, obviously, coincides with the exact probability[14] for the amplitude of the transition in the ultrarelativistic case and has a standard nonrelativistic limit.[15] Generally speaking, for long-range potentials, the integral with respect to the impact parameter in (3) diverges at large impact parameters. However, this divergence is unimportant:[4,8] at large impact parameters the ion field is weak and the Born approximation can be employed, with the applicability ranges of the Born and eikonal approximations overlapping, which makes a meaningful match with respect to the impact parameter possible.

The above formulas are of a general nature and can be applied to collisions of atoms with ions of arbitrary charge. What makes collisions of atoms with highly charged ions so special is that the cross sections of inelastic processes are usually very large and exceed the atomic dimensions significantly. Bearing all this in mind, we examine the matching procedure using the example of one-electron transitions in collisions of relativistic ions with nonrelativistic (before and after the collision[1)] atoms, when the ψ_i and ψ_f in (3) are two-component spinors and we can assume that $S^{-2}=1$; in addition, $\exp[i\Sigma_a x_a(E_f-E_i)/c]=1$. For this case simple analytical expressions for cross sections can be derived in the matching process. Suppose that an atomic electron goes from a ground state $|i\rangle$ into a state $|\mathbf{k}\rangle$ of the one-electron continuum with electron momentum $\mathbf{k}$. We denote the upper limit of integration with respect to the parameter b in Eq. (3) by b_0. For large values of b (i.e., $b \gg s$) and orthogonal states $|\mathbf{k}\rangle$ and $|i\rangle$, the generalized inelastic form factor is

$$\langle f|1-\exp\left\{-i\frac{2Z}{v}\ln\frac{|\mathbf{b}-\mathbf{s}|}{b}\right\}|i\rangle \approx \langle f|\exp\{i\mathbf{q}\cdot\mathbf{s}\}|i\rangle, \qquad (4)$$

with $\mathbf{q}=2Z\mathbf{b}/vb^2$; this form factor tends to $i\mathbf{q}\cdot\langle f|\mathbf{s}|i\rangle$ for small $\mathbf{q}$. Hence the integral in (3) depends on b_0 logarithmically, so that the contribution of the region $b<b_0$ to the cross section can be written as

$$\sigma(b<b_0)=8\pi\frac{Z^2}{v^2}\lambda_i\ln\frac{2\alpha_i}{q_0}, \quad q_0=\frac{2Z}{vb_0}, \qquad (5)$$

where λ and a depend solely on the atomic characteristics; they do not depend on the collision parameters, i.e., the charge of the incident ion and the ion velocity:

$$\lambda_i=\int\frac{d^3k}{3}|\langle\mathbf{k}|\mathbf{r}|i\rangle|^2,$$

$$\alpha_i=\lim_{q_0\to 0}\frac{q_0}{2}\exp\left\{\frac{1}{\lambda_i}\int_{q_0}^{\infty}\frac{dq}{q^3}\int d^3k\,|\langle\mathbf{k}|\exp(-i\mathbf{q}\cdot\mathbf{r})|i\rangle|^2\right\}.$$

In the region of large values of $b>b_0$, the ion field is a weak perturbation, and we can use what is known as the Bethe asymptotic formula:

$$\sigma_i(b>b_0)=8\pi\frac{Z^2}{v^2}\lambda_i\left(\ln\frac{2v}{\eta b_0\omega_i\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right); \qquad (6)$$

here $\eta=e^B=1.781$($B=0.5772$ is the Euler constant), and the ''average'' ionization frequency ω_i is defined as

$$\ln\omega_i=\frac{\int d^3k\,|\langle\mathbf{k}|\mathbf{r}|i\rangle|^2\ln\Omega_{\mathbf{k}i}}{\int d^3k\,|\langle\mathbf{k}|\mathbf{r}|i\rangle|^2}, \qquad (7)$$

where $\Omega_{\mathbf{k}i}=\epsilon_{\mathbf{k}}-\epsilon_i$ is the transition frequency. Summing (5) and (6), we arrive at expression for the total cross section:

$$\sigma_i=8\pi\frac{Z^2}{v^2}\lambda_i\left(\ln\frac{2\alpha_i v^2}{\eta Z\omega_i\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right). \qquad (8)$$

Note that the dependence on the cutoff parameter b_0 disappears after matching.

But if the collision changes the states of more than one electron or if the transitions are dipole-forbidden, the integration with respect to the impact parameter in (3) can be extended to the entire impact-parameter plane (since the integrand ensures convergence) and there is no need for a matching with perturbation theory.

3. COLLISIONS WITH HYDROGENLIKE ATOMS

Below we give the formulas for the cross section of the transition of a hydrogenlike atom (with a nuclear charge Z_α) from its ground state to states with the principal quantum number n as a result of a collision with a multicharged ion. These were obtained in the approximation of large impact parameters (Eq. (4)) by the matching method:

$$\sigma_n=\frac{2^{11}\pi}{3}\frac{Z^2}{v^2}\frac{n^7}{(n^2-1)^5}\left(\frac{n-1}{n+1}\right)^{2n}\frac{1}{Z_a^2}\times\left\{\ln\frac{\gamma_n v^2 Z_a}{Z\Omega_n\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}, \qquad (9)$$

where $\Omega_n=\epsilon_n-\epsilon_1$, and the first few numbers γ_n are

$$\gamma_2=0.30,\quad \gamma_3=0.44,\quad \gamma_4=0.49,\quad \gamma_5=0.53,$$
$$\gamma_6=0.54,\quad \gamma_7=0.55,\quad \gamma_8=0.56,\quad \gamma_9=0.57,$$
$$\gamma_{10}=0.57,\quad \gamma_{11}=0.57.$$

Similarly, the total ionization cross section is

$$\sigma_i=8\pi\frac{Z^2}{v^2}\frac{0.283}{Z_a^2}\left\{\ln\frac{5.08v^2}{ZZ_a\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}. \qquad (10)$$

Next, summing (9) over all values of n, we arrive at an expression for the total cross section of excitation of the discrete states of the hydrogen atom,

$$\sigma_{\rm exc}=\sum_{n=2}^{\infty}\sigma_n=8\pi\frac{Z^2}{v^2}\frac{0.717}{Z_a^2}\left\{\ln\frac{0.84v^2}{ZZ_a\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}, \qquad (11)$$

and the total inelastic cross section,

$$\sigma_r=\sigma_{\rm exc}+\sigma_i=8\pi\frac{Z^2}{v^2}\frac{1}{Z_a^2}\left\{\ln\frac{1.4v^2}{ZZ_a\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}. \qquad (12)$$

Equations (9) and (10) can be used to estimate the cross sections of excitation and ionization of the K-shells of complex nonrelativistic atoms as a result of a collision of the atom and a relativistic ion when the K-shell electrons are described by hydrogenlike functions with an effective nuclear charge Z_a. To obtain estimates of the cross sections of ionization or excitation of L-shells one can use the cross sections of ionization or excitations of hydrogenlike atoms from $2s$- and $2p$-states.

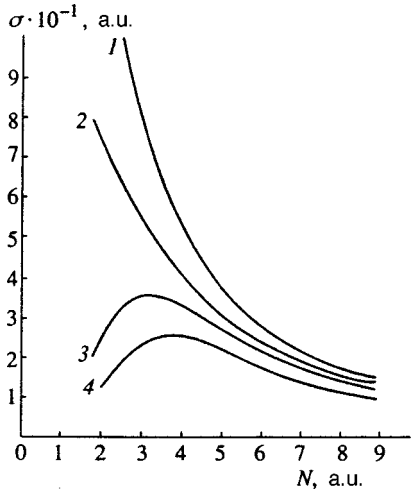


FIG. 1. Cross sections for ionization of the hydrogen atom by ions C^{6+} obtained in the Born approximation (curve *1*), by the matching method of Eq. (10) (curve *2*), in the Glauber approximation[10] (curve *3*), and in the sudden-perturbation approximation[8] (curve *4*).

For hydrogenlike atoms that initially were in $2s$-states we have

$$\sigma_i = 8\pi \frac{Z^2}{v^2} 0.82 \left\{ \ln \frac{17.1 v^2}{Z Z_a \sqrt{1-\beta^2}} - \frac{\beta^2}{2} \right\}, \tag{13}$$

$$\sigma_n = 8\pi \frac{Z^2}{v^2} \frac{2^{17}}{3} n^7 \frac{(n-2)^{2n-6}}{(n+2)^{2n+6}} (n^2-1) \times \left\{ \ln \frac{\beta_n v^2 Z_a}{Z \Omega_n \sqrt{1-\beta^2}} - \frac{\beta^2}{2} \right\}, \tag{14}$$

where $n \geqslant 3$, $\Omega_n = \epsilon_n - \epsilon_2$, and the numbers β_n are

$$\beta_3 = 0.18, \quad \beta_4 = 0.28, \quad \beta_5 = 0.34,$$

$$\beta_6 = 0.39, \quad \beta_7 = 0.41, \quad \beta_8 = 0.42,$$

$$\beta_9 = 0.44, \quad \beta_{10} = 0.45, \quad \beta_{11} = 0.46.$$

For hydrogen atoms that initially were in the $2p$-state we have (after averaging over the projections of angular momentum in this state)

$$\sigma_i = 8\pi \frac{Z^2}{v^2} 0.53 \left\{ \ln \frac{271 v^2}{Z Z_a \sqrt{1-\beta^2}} - \frac{\beta^2}{2} \right\}, \tag{15}$$

$$\sigma_n = 8\pi \frac{Z^2}{v^2} \frac{2^{15}}{3} n^{11} \frac{(n-2)^{2n-7}}{(n+2)^{2n+7}} \left(\frac{11}{3} - \frac{4}{n^2} \right) \times \left\{ \ln \frac{\beta_n v^2 Z_a}{Z \Omega_n \sqrt{1-\beta^2}} - \frac{\beta_2}{2} \right\}, \tag{16}$$

where $n \neq 2$, and the numbers β_n are

$$\beta_1 = 0.27, \quad \beta_3 = 0.13, \quad \beta_4 = 0.30, \quad \beta_5 = 0.46$$

$$\beta_6 = 0.58, \quad \beta_7 = 0.67, \quad \beta_8 = 0.73, \quad \beta_9 = 0.79$$

$$\beta_{10} = 0.82, \quad \beta_{11} = 0.85.$$

Figure 1 illustrates the nature of the results obtained by the matching method with the ionization cross sections for the hydrogen atom obtained in the Born representation (curve *1*), by the matching method of Eq. (10) (curve *2*), in the Glauber approximation[10] (curve *3*), and in the sudden-perturbation approximation[8] (curve *4*). We see that the cross section values obtained by Eq. (10) within its applicability range are close to those obtained in the Glauber approximation, and with increasing velocity tend to the values of the Born approximation. We also note that although for $Z \sim v \ll 1$ the expressions derived in this section do not transform directly into the Born approximation σ_B, which is a general property of the approximation (3). Instead, their relative difference tends to zero as the velocity increases, i.e., $(\sigma_B - \sigma)/\sigma \to 0$ as $v \to c$. But when $Z \sim v \sim 1$, our results, which are based on the unitary approximation (3), as well as the Glauber approximation, agree much better with the experimental results as compared to the Born approximation, which is known to be nonunitary and in this region overestimates inelastic cross sections substantially (by a factor of 1.5). Here, in contrast to direct application of the Glauber approximation in the form (2), which requires significant computer times even in the nonrelativistic case,[10,11] our formulas are analytic.

4. EXCITATION AND IONIZATION OF A HELIUMLIKE ATOM

If one uses perturbation theory in calculating the cross sections of inelastic collisions of fast charged particles with complex atoms, one-electron excitation or ionization is a first-order effect in the interaction of the incident particle and the atomic electrons. A two-electron transition corresponds to the second order of perturbation theory, whence the interaction of the incident particle and an atomic electron is taken into account once and the electron–electron interaction, also once. Multielectron transitions are calculated in a similar way,[3,19] i.e., the interaction of the incident particle and atomic electrons is always taken into account once, while all the rest amounts to taking into account the electron–electron interaction the necessary number of times.

The situation changes, however, when the interaction of the atomic electrons and the incident particle is much stronger than the electron–electron interaction. In this case a multielectron transition should be considered the result of the direct action[3,20] of the strong field of the incident particle, and it is to this mechanism of direct excitation that formulas (2) and (3) correspond (see Ref. 4). Below we give the cross sections of one- and two-electron transitions from the ground state of a nonrelativistic heliumlike atom in collisions with a multicharged relativistic ion. The formulas were derived in the approximation of large impact parameters (Eq. (4)). For all cases, in deriving these formulas we described the two-electron states of the heliumlike atom by symmetrized products of hydrogenlike one-electron wave functions.

To avoid the normalization procedure (which is usually ambiguous because, strictly speaking, one must orthogonalize all the states belonging to the continuous and discrete spectra), we selected one-electron hydrogenlike wave functions in the field of a nucleus with the same effective nuclear charge, which was equal to Z_1 for one-electron transitions

and Z_2 for two-electron. In this connection we denote arbitrary two-electron states $|n_1,n_2\rangle$ of the heliumlike atom by two sets of one-electron hydrogenlike quantum numbers n_1 and n_2. Then, according to (3), the cross section of the transition from the ground state $|0,0\rangle$ to a state $|n_1,n_2\rangle$ in the approximation (4) (large impact parameters) is given by the formula

$$\sigma=\int d^2b\,|\langle n_1,n_2|\exp\{-i\mathbf{q}\cdot(\mathbf{r}_1+\mathbf{r}_2)\}|0,0\rangle|^2. \tag{17}$$

We see that the cross section is expressed in terms of an impact-parameter integral of the product of well-known[15] hydrogenlike form factors. The cross section of two-electron transitions (in which the two electrons were sure to change state) were obtained directly from Eq. (17) by integrating over the entire impact-parameter plane. On the other hand, the cross sections of inelastic processes that incorporate one-electron transitions (e.g., single excitations of ionizations and total inelastic cross sections) are obtained via matching with perturbation theory, so that the corresponding formulas contain the characteristic logarithmic dependence on the velocity of the multicharged ion and the relativistic parameter $\gamma=1/\sqrt{1-v^2/c^2}$.

The cross section of single ionization, corresponding to one atomic electron arriving in any state of the continuum and the other, in any state of the discrete spectrum (or in any state belonging to the complete set of the discrete and continuous spectrum, but in this case we must subtract the contribution corresponding to two electrons being in states that belong to the two-electron continuum, i.e., in double-ionization states), is

$$\sigma^{1+}=16\pi\frac{Z^2}{v^2}0.283 \times\left\{\frac{1}{Z_1^2}\left[\ln\frac{5.08v^2}{ZZ_1\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right]-\frac{1}{Z_2^2}\ln 3.72\right\}. \tag{18}$$

The total cross section for one-electron excitations of the discrete spectrum, when only one atomic electrons is excited to a one-electron state belonging to the discrete spectrum and the other remains in the ground state (in an unexcited state, to be precise), has the form

$$\sigma^{1*}=16\pi\frac{Z^2}{v^2}\frac{0.375}{Z_1^2}\left\{\ln\frac{0.256v^2}{ZZ_1\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}. \tag{19}$$

The total double-ionization cross section can be obtained by summing (17) over all n_1 and n_2 belonging to the two-electron continuum. The result is

$$\sigma^{2+}=16\pi\frac{Z^2}{v^2}\frac{0.283}{Z_2^2}\ln 3.72=9.36\frac{Z^2}{Z_2^2v^2}. \tag{20}$$

The expression for the total cross section of the transition of a heliumlike atom to all doubly excited states of the discrete spectrum (including all possible autoionization states) can be found after the respective summation over n_1 and n_2 is done:

$$\sigma^{2*}=16\pi\frac{Z^2}{v^2}\frac{0.283}{Z_2^2}\ln 1.15=2.03\frac{Z^2}{v^2Z_2^2}. \tag{21}$$

The total cross sections are linked by the following obvious general relationship:

$$\sigma_r=\sigma^{1+}+\sigma^{1*}+\sigma^{2+}+\sigma^{2*}, \tag{22}$$

where the total inelastic cross section σ_r, corresponding to the possibility of any excitation of the heliumlike atom, is

$$\sigma_r=16\pi\frac{Z^2}{v^2}\frac{0.717}{Z_a^2}\left\{\ln\frac{1.03v^2}{ZZ_a\sqrt{1-\beta^2}}-\frac{\beta^2}{2}\right\}, \tag{23}$$

where Z_a is the effective charge of the nucleus of the heliumlike atom in the ground state $(1s^2)$, equal to the charge of the bare atom minus 5/16.

As another example of two-electron transitions into discrete states we examine the cross section of excitation of autoionization states with the principal quantum number $n=2$ (the L-shell) of a heliumlike atom. Since in the collisions considered here the electron spins do not change, excitation of the following autoionization states is possible: $2s^2\,{}^1S$, $2s2p\,{}^1P$, $2p^2\,{}^1S$, and $2p^2\,{}^1D$, with the corresponding cross sections being

$$\sigma(2p^2\,{}^1D)=2\sigma(2p^2\,{}^1S)=\frac{10}{3}\sigma(2s2p^1P) =30\sigma(2s^2\,{}^1S)=\pi\frac{Z^2}{v^2Z_2^2}\frac{2^{31}}{3^{19}}\frac{1}{11}.$$

In Table I we compare the experimental data of Refs. 21 and 22 with the results of our calculations and with the results of numerical calculations done by Fritsch and Lin[22] for the total cross sections of excitation to the autoionization states $2s2p+2p^2$ of the helium atom. The first column gives the incident-ion energies reduced to the atomic unit of mass, the second column gives the ion charge, the third gives the experimental results taken from Refs. 21 and 22, the fourth gives our results ($Z_2=1.97$), and the fifth the results of numerical calculation done in Ref. 22.

TABLE I. Sum of cross sections ($\times 10^{19}$ cm^2) of excitation of the autoionization $2s2p^1P$- and $2p^2\,{}^1D$-states of the helium atom.

Energy, MeV/nucleon	Ion charge	Experiment, 10^{-19} cm^2	Theory, 10^{-19}cm^2	Calculation,[22] 10^{-19} cm^2
1.84	6	8.305 ± 1.744	18.45	25.6
1.5	6	20.1 ± 7.20	22.61	31.8
1.5	9	48.99 ± 17.66	50.79	111.6

Figure 2 shows the results of experiments conducted by Berg *et al.*[23] and the results of calculations (by formulas (20) and (18)) of the cross sections of double ($Z_2=1.97$) and single ($Z_1=1.37$) ionizations of the helium atoms in collisions with relativistic uranium ions U^{90+} at 60, 120, and 420 MeV/nucleon and their ratio σ^{2+}/σ^{1+}. The proper choice of the values of the effective charges $Z_1=1.37$ and $Z_2=1.97$ is confirmed by the adequate agreement of our results listed in Table II with the experimental data of Refs. 24

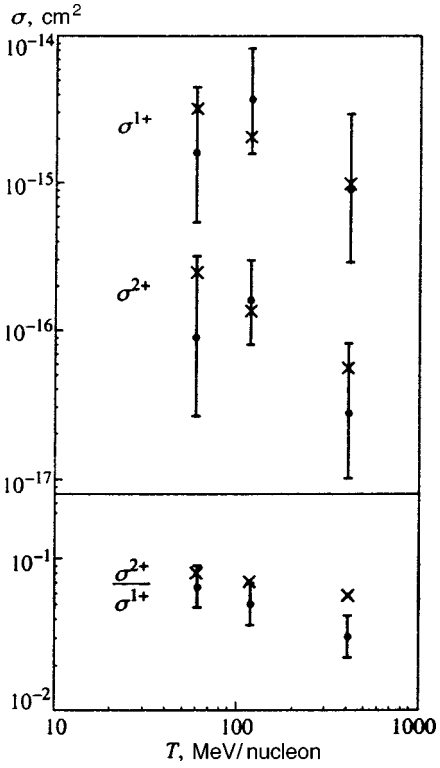


FIG. 2. The results of experiments conducted by Berg *et al.*[23] and the results of calculations (by formulas (20) and (18)) of the cross sections of double and single ionizations of the helium atoms in collisions with relativistic uranium ions U^{90+} at 60, 120, and 420 MeV/nucleon, and their ratio σ^{2+}/σ^{1+}: ●—experiment, and ×—calculation.

and 20 on double and single ionization of a helium atom as a result of a collision with a multicharged fast nonrelativistic ion. Note that here we list the results of calculations of cross sections of direct single ionization (σ^{1+}) and of cross sections ($\sigma^{1+}+\sigma^{2*}$), which may serve as estimates for the total cross sections of formation of singly charged helium ions as a result of direct ionization and Auger decay of all types of doubly excited states of the helium atom. The reason is that in low-Z atoms the Auger decay is the dominating channel of decay of the doubly excited states of the discrete spectrum, with the exception of a relatively small number of states (compared to the total number of all types of such states) for which Auger decay is forbidden by selection rules.[25–27]

5. EXCITATION AND IONIZATION OF A MULTIELECTRON ATOM

Although the strong field of a multicharged ion results in very high probabilities of electronic transitions, for calculations of cross sections of excitation or ionization of a high order the large-impact-parameter approximation (4) may become invalid, with the result that the corresponding cross sections may become comparable to the characteristic atomic dimensions. Hence we must consider the more general formula (3). We assume that the atomic electrons of a multielectron (before the collision) and nonrelativistic (after the collision) atom are not identical and assign to each electron a one-electron hydrogenlike wave function. Then the initial wave function is

$$\Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_{N_0}) = \prod_{i=1}^{N_0} \phi_i(\mathbf{r}_i),$$

and the final wave function is

$$\Psi_f(\mathbf{r}_1, \ldots, \mathbf{r}_{N_0}) = \prod_{i=1}^{N_0} \psi_i(\mathbf{r}_i).$$

Hence the total probability of (N_0-N)th-orderionization of a nonrelativistic N_0-electron atom corresponding to N_0-N electrons reaching the continuum and the remaining N electrons in any of the states of the discrete spectrum is, according to (3) and with allowance for unitarity,

$$W^{(N_0-N)+}(\mathbf{b}) = \frac{N_0!}{(N_0-N)!\,N!} \times \prod_{i=1}^{N_0-N} p_i(\mathbf{b}) \prod_{j=N_0-N+1}^{N_0} (1-p_j(\mathbf{b})), \qquad (24)$$

where $\Pi_{j=N_0-N+1}^{N_0}(\cdots)=1$ for $N=0$, and the generalized one-electron inelastic form factor is,

$$p_i(\mathbf{b}) = \int d^3k_i \left| d^3r_i\, \psi^*_{\mathbf{k}_i}(\mathbf{r}_i) \exp\{-i\chi_i(\mathbf{b},\mathbf{r}_i)\}\phi_i(\mathbf{r}_i)\right|^2, \qquad (25)$$

with $\mathbf{k}_i$ the momentum of the ith electron in the continuum. We see that this probability depends on the vector $\mathbf{b}$. However, after being averaged over the projection of the total orbital angular momentum of the initial state, the probability becomes a function of only $|\mathbf{b}|$. Let us introduce the average

TABLE II. Cross sections ($\times 10^{16}$ cm^2) of double and single ionization of the helium atom.

Energy, MeV/nucleon	Ion charge	σ^{2+} expt.	σ^{2+} theory	σ^{1+} expt.	σ^{1+} theory	$\sigma^{1+}+\sigma^{2*}$ theory
0.64	8	1.32	1.687	7.9	10.231	10.597
1.00	8	1.06	1.08	6.7	8.11	8.344
1.44	8	0.45	0.75	5.9	6.518	6.68
1.4	15	2.91	2.712	17.9	17.798	18.385
1.4	18	4.50	3.905	22.4	23.322	24.168
1.4	20	5.41	4.821	26.0	27.146	28.191
1.4	36	16.0	15.621	57.2	58.206	61.59
1.4	37	16.8	16.501	59.5	60.02	63.594
1.4	44	23.0	23.335	72.1	71.779	76.833

over the orbital angular momentum l and the projection of this momentum m of the value of the one-electron inelastic form factor for each electron of the shell (which is then averaged over all the shells of the atom):

$$p(b)=\frac{1}{n_0}\sum_{n=1}^{n_0}\frac{1}{M_n}\sum_{l,m}\int d^3k\left|\int d^3r\,\psi_{\mathbf{k}}^*(\mathbf{r})\times\exp\{-i\chi(\mathbf{b},\mathbf{r})\}\phi_{nlm}(\mathbf{r})\right|^2, \tag{26}$$

where the inner sum is over all possible values of l and m for a given nth shell, M_n is the number of such values, n is the principal quantum number, and n_0 is the number of shells. Obviously, $p(b)=p(|\mathbf{b}|)$ is independent of the angles determining the direction of $\mathbf{b}$, with the result that $p(b)$ has the meaning of the average one-electron ionization probability. Then, replacing each one-electron form factor in (24) by the average (26), we arrive at an expression for the probability of (N_0-N)th-order ionization common for the independent-electron approximation.[3,9] However, the effective charge Z^* of the atomic nucleus depends on the order of ionization. To allow for this, we make the following substitutions in (26):

$$\mathbf{k}=\frac{\mathbf{k}}{Z^*},\quad \mathbf{b}=\mathbf{b}Z^*,\quad \mathbf{r}=\mathbf{r}Z^*,$$

which means that we are introducing Coulomb units.[15] Then the right-hand side of Eq. (26) can be calculated by using the wave functions of the hydrogen atom with a unit charge, and the dependence on Z^* amounts to the substitution $b=bZ^*$. Hence, if from now on we interpret $p(b)$ as the form factor of the hydrogen atom averaged in accordance with (26), such a substitution makes it possible to calculate the ionization cross section under assumptions that are more general than those used in the independent-electron model. Below we consider high-order ionization, i.e., $N_0\gg 1$ and $N_0-N\gg 1$. We start with the cross section of total ionization of the atom (when the atom is stripped of all of its N_0 electrons). Then in (24) we have N_0, and W is simply the product of N_0 one-electron form factors. We introduce an effective nuclear charge corresponding to total ionization of the atom, $Z_{N_0}^*$. Replacing each one-electron form factor by the average (26), we arrive at an expression for the probability of total ionization:

$$W^{N_0+}=[p(b)]^{N_0},$$

where $b=bZ_{N_0}^*$. The integral (3) of this probability can be evaluated asymptotically $(N_0\gg 1)$ by the Laplace method under the assumption that $p(b)$ has its maximum at the left limit $b=b_0=0$ of the integration interval. The fact that such a maximum exists can easily be verified by studying the results of calculations in Refs. 8 and 14. All this leads to a formula for the cross section of N_0-order ionization of the shell:

$$\sigma^{N_0+}=\pi\frac{1}{(Z_{N_0}^*)^2}\left[\frac{-2\pi}{p''(b_0)N_0}\right]^{1/2}[p(b_0)]^{N_0+1/2}. \tag{27}$$

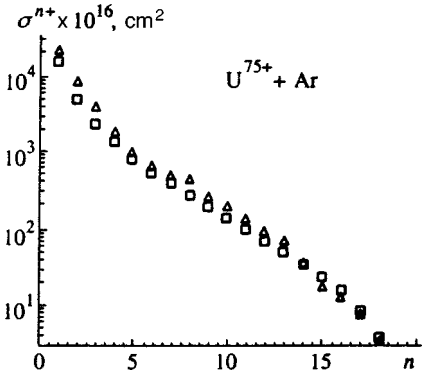


FIG. 3. Cross section of multiple ionization of the argon atom by multicharged uranium ions U^{75+} with an energy of 15 MeV/nucleon as a function of the order of ionization n: □—the results of the experiment of Ullrich *et al.*,[28] and △—our results.

Here and in what follows, b_0 is the point of maximum of $p(b)$, and $p''(b_0)$ is the value of the second p-derivative of $p(b)$ at point b_0. In the case of (N_0-1)st-order ionization, the ionization probability is simply the difference of two terms. The first term contains the product of N_0-1 one-electron form factors and corresponds to a situation in which N_0-1 electrons are in the continuum (the effective charge is $Z_{N_0-1}^*$), and the second term contains the product of N_0 one-electron form factors and corresponds to a situation in which N_0 electrons are in the continuum (the effective charge is $Z_{N_0}^*$). Integrating each term separately by the Laplace method, we arrive at a formula for the cross section of (N_0-1) st-orderionization:

$$\sigma^{(N_0-1)+}=N_0\sigma^{N_0+}\left[\left(\frac{Z_{N_0}^*}{Z_{N_0-1}^*}\right)^2\left(\frac{N_0}{N_0-1}\right)^{1/2}\frac{1}{p(b_0)}-1\right]. \tag{28}$$

Reasoning along similar lines in the general case of (N_0-N) th-order ionization, we obtain

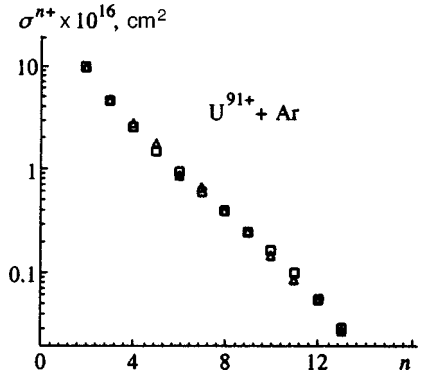


FIG. 4. Results of the experiments described in Refs. 28 and 29 and of the calculations (by formulas (27)–(29)) of the cross sections of multiple ionization of neon atoms in collisions with relativistic uranium ions U^{91+} with an energy of 120 MeV/nucleon as functions of the order of ionization n: □—experiment, and △—our results.

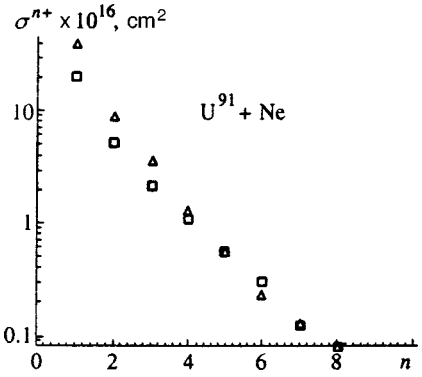


FIG. 5. Results of the experiments described in Refs. 28 and 29 and of the calculations (using formulas (27)–(29)) of the cross sections of multiple ionization of argon atoms in collisions with relativistic uranium ions U^{91+} with an energy of 120 MeV/nucleon as functions of the order of ionization n: □—experiment, and △—our results.

$$\sigma^{(N_0-N)+} = \frac{N_0!\,\sigma^{N_0+}}{(N_0-N)!\,N!}\sum_{m=0}^{N}(-1)^m \times\left(\frac{Z^*_{N_0}}{Z^*_{N_0-N+m}}\right)^2 \frac{N!\,\sqrt{N_0/(N_0-N+m)}}{(N-m)!\,m!} \times[p(b_0)]^{-N+m}, \tag{29}$$

where $Z^*_{N_0-N+m}$ is the effective charge for the case of (N_0-N+m)th-order ionization.

Equations (27)–(29) make it possible, at least in principle, to calculate the cross sections for ionization of any order (provided that $N_0 \gg 1$ and $N_0-N \gg 1$) or, knowing any two cross sections from experiments, to reproduce the remaining cross sections. The simplest approach here is to assume that the known cross sections are σ^{N_0+} and $\sigma^{(N_0-1)+}$. Then we can easily find $p(b_0)$ and insert it in (29). As a result, the cross section of arbitrary (N_0-N)th-order ionization will be expressed in terms of σ^{N_0+} and $\sigma^{(N_0-1)+}$. The results of such calculations for multiple (up to the eighth order) ionization of the neon atom and (up to eighteenth order) ionization of the argon atom are displayed in Figs. 3–5. Here the effective charge was assumed equal to the order of ionization, i.e., $Z^*_N = N$. The figures clearly show that the agreement of the results of calculations with those of the experiments described in Refs. 28 and 29 is good even for low-order ionization, which formally lies outside the applicability range $(N_0-N \gg 1)$ of Eqs. (27)–(29).

*)E-mail: matveev@otf.gimli.com

1)Strictly speaking, as a result of ionization by an ion moving with a relativistic velocity, the atomic electrons arriving in the continuum may acquire relativistic velocities. However, as shown in Ref. 9, such processes occur at small impact parameters, so that the corresponding contribution to the total cross sections of ionization by the impact of a multicharged ion can be ignored.

[1] R. K. Janev, L. P. Presnyakov, and V. P. Shevel'ko, *Physics of Highly Charged Ions*, Springer, New York (1985).
[2] J. Eichler and W. E. Meyrhof, *Relativistic Atomic Collisions*, Academic Press, New York (1995).
[3] J. H. McGuire, in: *Advances in Atomic, Molecular and Optical Physics*, Academic Press, New York (1992), p. 217.
[4] V. I. Matveev, Fiz. Élem. Chastits At. Yadra **26**, 780 (1995) [Phys. Part. Nuclei **26**, 329 (1995)].
[5] J. Eichler, Phys. Rep. **193**, 167 (1990).
[6] J. H. Eichler, Phys. Rev. A **15**, 1856 (1977).
[7] G. L. Yudin, Zh. Éksp. Teor. Fiz. **80**, 1026 (1981) [Sov. Phys. JETP **53**, 523 (1981)].
[8] A. Salop and J. H. Eichler, J. Phys. B **12**, 257 (1979).
[9] V. I. Matveev and M. M. Musakhanov, Zh. Éksp. Teor. Fiz. **105**, 280 (1994) [JETP **78**, 149 (1994)].
[10] J. H. McGuire, Phys. Rev. A **26**, 143 (1982).
[11] D. S. F. Crothers and S. H. McCann, J. Mol. Struct. **16**, 3229 (1983).
[12] V. I. Matveev and S. G. Tolmanov, Zh. Éksp. Teor. Fiz. **107**, 1780 (1995) [JETP **80**, 989 (1995)].
[13] U. Becker, N. Grun, W. Scheid, and G. Soff, Phys. Rev. Lett. **56**, 2016 (1986).
[14] A. J. Baltz, Phys. Rev. Lett. **78**, 1231 (1997).
[15] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-relativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).
[16] M. Goldberger and K. Watson, *Collision Theory*, Wiley, New York (1964).
[17] N. F. Mott and H. S. W. Massey, *The Theory of Atomic Collisions*, 3rd ed., Clarendon Press, Oxford (1965).
[18] A. I. Akhiezer and V. B. Berestetskiĭ, *Quantum Electrodynamics*, Wiley, New York (1974).
[19] V. I. Matveev and É. S. Parilis, Usp. Fiz. Nauk **138**, 583 (1982) [Sov. Phys. Usp. **25**, 881 (1982)].
[20] J. H. McGuire, A. Müller, B. Shuch, W. Groh, and E. Salzborn, Phys. Rev. A **35**, 2479 (1987).
[21] J. O. R. Pedersen and P. Hvelplund, Phys. Rev. Lett. **62**, 2373 (1989).
[22] W. Fritsch and C. D. Lin, Phys. Rev. A **41**, 4776 (1990).
[23] H. Berg, O. Jagutzki, R. Dörner, R. D. DuBois, C. Kelbch, H. Schmidt-Böcking, J. Ullrich, J. A. Tanis, A. S. Schlachter, L. Blumenfeld, B. d'Etat, S. Hagmann, A. Gonzales, and T. Quinteros, Phys. Rev. A **46**, 5539 (1992).
[24] H. Knudsen, L. H. Andersen, P. Hvelplund *et al.*, J. Phys. B **17**, 3445 (1984).
[25] G. Wentzel, Z. Phys. **43**, 524 (1927).
[26] P. Kh. Propin, Opt. Spektrosk. **8**, 300 (1960) [Opt. Spectrosc. (USSR) **8**, 158 (1960)].
[27] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Springer, Berlin (1958).
[28] J. Ullrich, R. E. Olson, H. Berg *et al.*, Nucl. Instrum. Methods Phys. Res. B **40/41**, 149 (1989).
[29] H. E. Berg, GSI-93-12, Report, May, ISSN 0171-4546 (Doktorarbeit, Universität Frankfurt) (1993), p. 1.

Translated by Eugene Yankovsky

Radiative damping of a relativistic electron in classical electrodynamics

G. F. Efremov

Nizhniĭ Novgorod State Teachers' Training University, 603600 Nizhniĭ Novgorod, Russia
(Submitted 28 May 1998)
Zh. Éksp. Teor. Fiz. **114**, 1661–1671 (November 1998)

An exact solution of the problem of the reaction of the field generated by a relativistic classical electron is derived. It is found that the solution differs dramatically from the known formulas by the presence of a component that is even under time reversal. It is also shown that the component of the generalized radiative damping force that is odd under time reversal coincides with the well-known relativistic damping force obtained from the approximate nonrelativistic formula via a Lorentz transformation. © *1998 American Institute of Physics.* [S1063-7761(98)00811-7]

1. INTRODUCTION

To use V. L. Ginzburg's clever comment, physics has several literally eternal problems, which have been debated for decades in the scientific community. One such problem is radiative damping.[1]

The classical expression for the radiative damping force (see formula (75.8) of Ref. 2) leads, as is well known, to instability in the electron's motion, or what is known as the self-accelerating paradox. Actually this means that not only the law of energy conservation but also the causality principle breaks down.[3] These paradoxes, which arose at the beginning of the 20th century, continue to draw attention.[4–10] Nevertheless, the unflagging interest yielded no solution in terms of classical electrodynamics. Physically speaking, the problem of radiative damping must be solved with allowance for quantum effects.[3]

The above paradoxes can be resolved if one allows for laws that are essentially quantum mechanical, which describe the interaction of the electron and the radiation field.[3] Such an approach also leads to a rigorous justification for the conditions under which the common classical formulas of radiative damping can be used.[1]

In this connection it becomes essential to close the gaps that exist in classical radiative damping theory.

1. The classical theory has no consistent relativistic approach to the derivation of the formula for the radiative damping force. The common transition to the relativistic domain is accomplished via Lorentz transformations from the approximate nonrelativistic formula (75.8) of Ref. 2, which is insufficient for a physical theory.

2. There is no procedure for removing divergences inherent in classical electrodynamics.

3. In the nonrelativistic model, the electromagnetic mass enters into the formulas with an incorrect factor of 4/3, which was first obtained by Thomson and then by Abragam and Lorentz.[11]

The present paper provides an exact expression for the reaction of the radiation field emitted by the electron. The essentially relativistic nature of this formula is determined by the Green's function of the field. Hence the damping force has a component that is even under time reversal and one that is odd. A procedure based on the integral representation of the Green's function is proposed that makes it possible to remove the divergences and renormalize the mass. It will be shown that in the relativistic theory (in contrast to the nonrelativistic Abragam–Lorentz model[11]) there is no factor of 4/3 in the expression for the electromagnetic mass. An exact analytic formula is derived for the generalized radiative damping force, and the similarities and discrepancies between this formula and the well-known relativistic expression[1,2] are established. It is found that the odd component (under time reversal) coincides with the well-known relativistic formula. What is new here is that the radiative force has an even component (under time reversal), absent from earlier formulas. In the reference frame co-moving with the electron, this force can be interpreted as a new effect of the dependence of mass on acceleration.

Thus, the exact solution of the physical problem, which is important in itself, makes it possible in the present case to establish the essentially new laws governing the radiative damping of a classical electron both in the relativistic domain and in the nonrelativistic domain.

2. REACTION OF THE RADIATION FIELD GENERATED BY A RELATIVISTIC ELECTRON

We wish to derive, in a rigorous manner, an expression for the reaction of the field of a relativistically moving classical point electron without resorting to the common nonrelativistic approximation. From the standpoint of the correspondence principle of consistent quantum theory,[3] it is natural to use the Hamiltonian method of solving the classical problem.[1]

The relativistic Hamiltonian of an electron placed in a potential field with potential energy $V(\mathbf{r})$ and interacting with its own field has the form

$$H=c\sqrt{\left[\mathbf{p}-\frac{e}{c}\mathbf{A}(\mathbf{r},t)\right]^2+(mc)^2}+eA_0(\mathbf{r},t)+V(\mathbf{r})+F, \tag{1}$$

where $A_0(\mathbf{r},t)$ and $\mathbf{A}(\mathbf{r},t)$ are the potentials of the field, and F is the Hamiltonian of the radiation field. When analyzing the problems of radiative damping it is convenient to use the gauge symmetry, selecting the transverse (Coulomb) gauge for the potentials of the field:

$$\operatorname{div}\mathbf{A}(\mathbf{r},t)=0. \tag{2}$$

The scalar potential $A_0(\mathbf{r},t)$, which is responsible in this case for the Coulomb interaction,[12] leads to no observable effects in the single-particle problem and can be dropped from the system Hamiltonian (1).

To derive an expression for the radiative damping force we must write the exact equations of motion for the dynamical variables of the electron. The system Hamiltonian (1) gives rise to the well-known Lorentz equation:

$$\frac{d}{dt}\left[\frac{m\dot{r}_j(t)}{\sqrt{1-v^2/c^2}}\right]+\nabla_j V(\mathbf{r}(t),t)$$
$$=-\frac{e}{c}\frac{\partial}{\partial t}A_j(\mathbf{r}(t),t)+\frac{e}{c}\dot{r}_\alpha(t)(\nabla_j A_\alpha(\mathbf{r}(t),t)$$
$$-\nabla_\alpha A_j(\mathbf{r}(t),t)), \tag{3}$$

where $\dot{r}_j(t)=\partial H/\partial p_j\equiv v_j(t)$ are the components of the electron velocity.

We write the field potentials in the Lorentz equation (3), which are functions of the electron's radius vector $\mathbf{r}(t)$ at the same time t, in the form of Fourier expansions:

$$A_j(\mathbf{r}(t),t)=\int\frac{d\mathbf{k}}{(2\pi)^3}\exp[i\mathbf{k}\cdot\mathbf{r}(t)]\,A_j(\mathbf{k},t), \tag{4}$$

where the Fourier transforms $A_j(\mathbf{k},t)$ do not contain the electron variables explicitly. The components of the current density, which are the canonical conjugates of the $A_j(\mathbf{k},t)$, can be found from the system Hamiltonian (1) via the relationship

$$-\frac{\delta H}{\delta A_j(\mathbf{k},t)}=\frac{e}{c}\dot{r}_j(t)\exp[i\mathbf{k}\cdot\mathbf{r}(t)]. \tag{5}$$

The Hamiltonian leads to equations for the Fourier transforms $A_j(\mathbf{k},t)$ of the field potentials:

$$\left(k^2+\frac{1}{c^2}\frac{d^2}{dt^2}\right)A_j(\mathbf{k},t)=\frac{4\pi}{c}e\dot{r}_j(t)\exp[i\mathbf{k}\cdot\mathbf{r}(t)], \tag{6}$$

which are the Maxwell equations for the radiation field generated by an electron with a current density

$$j_\alpha(\mathbf{r},t)=e\,\delta^3(\mathbf{r}-\mathbf{r}(t))\dot{r}_\alpha(t).$$

Equation (6) yields an exact solution for the radiation field generated by the electron:

$$A_j(\mathbf{k},t)=\frac{e}{c}\int dt_1\,D_{jl}(\mathbf{k},t-t_1)\dot{r}_l(t_1)\exp[-i\mathbf{k}\cdot\mathbf{r}(t_1)], \tag{7}$$

where $D_{jl}(\mathbf{k},t-t_1)$ is known as the retarded Green's function of the field. In the adopted gauge, the Fourier spectrum of the Green's function has the form[13]

$$D_{jl}(\mathbf{k},\omega)=\int_{-\infty}^{\infty}d\tau\exp(i\omega\tau)\,D_{jl}(\mathbf{k},\tau)$$
$$=4\pi(k^2+\varkappa^2)^{-1}\left(\delta_{jl}-\frac{k_jk_l}{k^2}\right), \tag{8}$$
$$k=|\mathbf{k}|,\quad \varkappa^2=\left(\frac{i\omega}{c}\right)^2(1+i\varepsilon\,\operatorname{sgn}\omega).$$

The presence of $i\varepsilon\,\operatorname{sgn}\omega$ $(\varepsilon>0)$ ensures the correct traversal of the pole in the retarded Green's function. In classical electrodynamics there is no such concept as electromagnetic vacuum, so that (7) contains no solution for an Eq. (6) with a zero right-hand side.

As a result, the reaction of the radiation field generated by the electron is found to depend exclusively on the electron variables. Substituting the exact solution (7) in (4) yields

$$A_j(\mathbf{r}(t),t)=\frac{e}{c}\int dt_1\int\frac{d\mathbf{k}}{(2\pi)^3}D_{jl}(\mathbf{k},t-t_1)\dot{r}_l(t_1)$$
$$\times\exp(i\mathbf{k}\cdot\Delta\mathbf{r}), \tag{9}$$

where $\Delta\mathbf{r}=\mathbf{r}(t)-\mathbf{r}(t_1)$. Finally, by inserting (9) into the Lorentz equation (3) we arrive at a rigorous relativistic expression for the radiative damping force acting on a classical electron:

$$\frac{d}{dt}\left[\frac{mr_j(t)}{\sqrt{1-v^2(t)/c^2}}\right]+V(\mathbf{r}(t))$$
$$=F_j(t)\equiv-\left(\frac{e}{c}\right)^2\frac{\partial}{\partial t}\int dt_1\int\frac{d\omega}{2\pi}$$
$$\times\exp[-i\omega(t-t_1)]\,D_{jl}(\Delta\mathbf{r},\omega)\dot{r}_l(t_1)+\left(\frac{e}{c}\right)^2\dot{r}_\alpha(t)$$
$$\times\int dt_1\int\frac{d\omega}{2\pi}\exp[-i\omega(t-t_1)]$$
$$\times[\nabla_j D_{\alpha l}(\Delta\mathbf{r},\omega)-\nabla_\alpha D_{jl}(\Delta\mathbf{r},\omega)]\dot{r}_l(t_1), \tag{10}$$

where

$$D_{jl}(\Delta\mathbf{r},\omega)=\int\frac{d\mathbf{k}}{(2\pi)^3}\exp(i\mathbf{k}\cdot\Delta\mathbf{r})\,D_{jl}(\mathbf{k},\omega) \tag{11}$$

determines the spatial dependence of the Green's function of the field. By passing from the partial time derivative to the total time derivative we can transform the expression for the radiative damping force $F_j(t)$ in (10) into

$$F_j(t)=-\left(\frac{e}{c}\right)^2\frac{d}{dt}\int dt_1\int\frac{d\omega}{2\pi}$$
$$\times\exp[-i\omega(t-t_1)]\,D_{jl}(\Delta\mathbf{r},\omega)\dot{r}_l(t_1)$$
$$+\left(\frac{e}{c}\right)^2\dot{r}_\alpha(t)\nabla_j\int dt_1\int\frac{d\omega}{2\pi}$$
$$\times\exp[-i\omega(t-t_1)]\,D_{\alpha l}(\Delta\mathbf{r},\omega)\dot{r}_l(t_1). \tag{12}$$

The same result can be obtained by passing to the limit $\hbar \to 0$ in the quantum mechanical formula (24) obtained in Ref. 3 at absolute zero

The above rigorous definition of the radiative damping of a classical electron, which does not involve the nonrelativistic limit, is of special interest since it shows directly that the paradox of a self−accelerating electron cannot be resolved in classical electrodynamics.

3. THE RETARDED GREEN′S FUNCTION; RENORMALIZATION OF MASS IN CLASSICAL ELECTRODYNAMICS

To obtain an explicit analytical formula for the radiative force $F_j(t)$, we must first find the Fourier transform (11) of the Green′s function (8). The features of the Green′s function (8) as a tensor are determined by the fact that in the adopted Coulomb gauge (2) the field′s vector potential is transverse. For convenience we separate the principal scalar part in tensor (8), which we denote by

$$G(k,\omega)=\frac{4\pi}{k^2+\varkappa^2}, \quad \varkappa=\frac{i\omega}{c}(1+i\varepsilon\ \mathrm{sgn}\,\omega), \tag{13}$$

and write (8) as

$$D_{jl}(\mathbf{k},\omega)=\delta_{jl}G(k,\omega)+\frac{1}{\varkappa^2}k_jk_l[G(k,\omega)-G(k,0)]. \tag{14}$$

Using the analytic properties of the function (13) in the upper half ω-plane, we can write the Fourier transform of (13),

$$G(\Delta r,\omega)=\int\frac{d\mathbf{k}}{(2\pi)^3}\exp(i\mathbf{k}\cdot\Delta\mathbf{r})\frac{4\pi}{k^2+\varkappa^2}$$

$$=\frac{1}{\Delta r}\exp(\varkappa\Delta r), \tag{15}$$

$$\Delta r=|\mathbf{r}(t)-\mathbf{r}(t_1)|, \quad \varkappa=i\omega/c,$$

and the Fourier transform of the Green's function (11),

$$D_{jl}(\Delta\mathbf{r},\omega)=\delta_{jl}G(\Delta r,\omega)-\frac{1}{\varkappa^2}\nabla_j\nabla_l\tilde{G}(\Delta r,\omega). \tag{16}$$

For the function

$$\tilde{G}(\Delta r,\omega)=G(\Delta r,\omega)-G(\Delta r,0). \tag{17}$$

we introduce an integral representation, which will be extremely useful in further calculations:

$$\tilde{G}(\Delta r,\omega)=\varkappa\int_0^1 d\beta\exp(\varkappa\Delta r\beta), \tag{18}$$

and correspondingly,

$$G(\Delta r,\omega)=\varkappa\int_0^1 d\beta\exp(\varkappa\Delta r\beta)+\frac{1}{\Delta r}. \tag{19}$$

Substituting (18) in (16) and calculating the derivatives ∇_j and ∇_l, we arrive at an expression for the Green's function,

$$D_{jl}(\Delta\mathbf{r},\omega)=\delta_{jl}G(\Delta r,\omega)-\frac{\Delta r_j\Delta r_l}{\Delta r^2}\int_0^1 d\beta\,\beta^2\exp(\varkappa\Delta r\beta)$$

$$-\left(\delta_{jl}-\frac{\Delta r_j\Delta r_l}{\Delta r^2}\right)\int_0^1 d\beta\,\beta\exp(\varkappa\Delta r\beta)\frac{1}{\Delta r}. \tag{20}$$

Integrating by parts,

$$\int_0^1 d\left(\frac{1}{2}\beta^2\right)\exp(\varkappa\Delta r\beta)\frac{1}{\Delta r}$$

$$=\frac{1}{2}\frac{1}{\Delta r}\exp(\varkappa\Delta r)-\frac{\varkappa}{2}\int_0^1 d\beta\exp(\varkappa\Delta r\beta), \tag{21}$$

and using the integral representation (19), we can reduce the expression (16) for Green's function (16) to

$$D_{jl}(\Delta r,\omega)=D^0_{jl}(\Delta r)+\tilde{D}_{jl}(\Delta r,\omega). \tag{22}$$

In (22) we have separated the terms that are frequency-independent and have a singularity in the limit $\Delta r\to 0$:

$$D^0_{jl}(\mathbf{r})=\frac{1}{\Delta r}\delta_{jl}-\frac{1}{2\Delta r}\left(\delta_{jl}-\frac{\Delta r_j\Delta r_l}{\Delta r^2}\right). \tag{23}$$

Accordingly, the regular part $\tilde{\Delta}_{jl}(\Delta\mathbf{r},\omega)$ of the Green's function has no singularities for $\Delta r\to 0$:

$$\tilde{D}_{jl}(\Delta\mathbf{r},\omega)=\delta_{jl}\varkappa\int_0^1 d\beta\exp(\varkappa\Delta r\beta)-\delta_{jl}\frac{\varkappa}{2}\int_0^1 d\beta\,(1-\beta^2)$$

$$\times\exp(\varkappa\Delta r\beta)+\frac{\Delta r_j\Delta r_l}{\Delta r^2}\frac{\varkappa}{2}\int_0^1 d\beta\,(1-3\beta^2)$$

$$\times\exp(\varkappa\Delta r\beta). \tag{24}$$

Bearing in mind that

$$\int_0^1 d\beta\,(1-3\beta^2)\exp(\varkappa\Delta r\beta)$$

$$=-\Delta r\int_0^1 d\beta\,(1-\beta^2)\varkappa\beta\exp(\varkappa\Delta r\beta), \tag{25}$$

we can write the regular part (24) of the Green's function as

$$\tilde{D}_{jl}(\Delta\mathbf{r},\omega)=\delta_{jl}\varkappa\int_0^1 d\beta\exp(\varkappa\Delta r\beta)$$

$$-\delta_{jl}\frac{\varkappa}{2}\int_0^1 d\beta\,(1-\beta^2)\exp(\varkappa\Delta r\beta)$$

$$-\Delta r_j\frac{\varkappa}{2}\int_0^1 d\beta\,(1-\beta^2)\varkappa\beta\frac{\Delta r_l}{\Delta r}\exp(\varkappa\Delta r\beta). \tag{26}$$

By writing the Green's function as the sum of the singular and regular parts we can analyze the divergences inherent in classical electrodynamics and discuss the problem of electromagnetic mass. Since the singular part (23) of the Green's

function is frequency-independent, if we substitute it into (10) and allow for the radiative damping force in the form (12), we find

$$\frac{d}{dt}\left[\frac{m\dot{r}_j(t)}{\sqrt{1-v^2/c^2}}\right]+\nabla_j V(\mathbf{r}(t))=-\delta m\,\ddot{r}_j(t), \tag{27}$$

where the additional electromagnetic mass

$$\delta m=\frac{e^2}{c^2}(\Delta r)^{-1}_{t_1=t},\quad (\Delta r)^{-1}_{t_1=t}=\int\frac{d\mathbf{k}}{(2\pi)^2}4\pi k^{-2}, \tag{28}$$

is formally divergent. The common procedure of mass renormalization consists in including the additional electromagnetic mass δm in the total renormalized mass

$$m^*=m+\delta m, \tag{29}$$

a value assumed by the experimentally measured mass. After the renormalization procedure has been completed, the regular part (26) of the Green's function remaining in the equations of motion leads to a finite value for the radiative damping force determined by (10). Note, however, that this fact does not resolve the self-accelerating paradox.

Concluding this section, we note that the physical reason for divergences in classical electrodynamics is that by integrating in (28) over the entire **k**-space we step outside the limits of of classical electrodynamics. The restriction on integration in **k**-space may be associated only with one characteristic length in classical electrodynamics, the classical electron radius $a_{\rm cl}$, i.e.,

$$\frac{e^2}{a_{\rm cl}}=mc^2. \tag{30}$$

Because of this we can assume that

$$(\Delta r)^{-1}_{t_1=t}=\frac{1}{a_{\rm cl}}=\frac{mc^2}{e^2}. \tag{31}$$

If we assume that Δr in (28) corresponds to the classical electron radius $a_{\rm cl}$ (Eq. (31)), the additional mass δm coincides with the electron mass. From this viewpoint the electron mass is of a purely electromagnetic nature and is related to the radiative damping force acting on the charge. The allowance for the fact that the field is transverse in calculations of the singular part (23) of the Green's function resolves the problem of the factor of 4/3, which initiallyarose in Thomson's work and then in the nonrelativistic Abragam–Lorentz model.[11]

4. A RELATIVISTIC FORMULA FOR THE RADIATIVE DAMPING FORCE OF A CLASSICAL ELECTRON

As is well known, the retarded Green's function of the field (Eqs. (8) and (11)) is a complex-valued function of the frequency ω. Due to this, the radiative damping force $F_j(t)$ in (10) contains a component that is even under time reversal and one that is odd. Hence the generalized radiative damping force $F_j(t)$ differs dramatically from the well-known radiative damping force, which is odd under time reversal (see formula (3.12) in Ref. 1). Thus, it is important to rigorously determine the time dependence of radiative damping.

The integral representation of the Green's function makes it possible to obtain an explicit analytical formula for the generalized radiative damping force without using approximation schemes. What is especially important here is the procedure of mass renormalization proposed in Sec. 3. First we use the representation (16) of the Green's function and substitute it in the expression for the radiative damping force $F_j(t)$ defined in Eq. (10). Bearing in mind the obvious fact that the derivatives ∇_j,∇_α, and ∇_l commute and allowing for mass renormalization, we find that

$$\begin{aligned}F_j(t)=&-\frac{e^2}{c^2}\frac{\partial}{\partial t}\int dt_1\int\frac{d\omega}{2\pi}\exp[-i\omega(t-t_1)]\\&\times\tilde{D}_{jl}(\Delta r,\omega)v_l(t_1)+\frac{e^2}{c^2}v_\alpha(t)\\&\times\int dt_1\int\frac{d\omega}{2\pi}\exp[-i\omega(t-t_1)]\\&\times[\delta_{\alpha l}\nabla_j-\delta_{jl}\nabla_\alpha]\tilde{G}(\Delta r,\omega)v_l(t_1),\end{aligned} \tag{32}$$

where $v_l(t_1)=\dot{r}_l(t_1)$, and $v_\alpha(t)=\dot{r}_\alpha(t)$. In accordance with the mass−renormalization procedure, in (32) we keep only the regular parts of the Green′s function, $\tilde{D}_{jl}(\Delta r,\omega)$ and $\tilde{G}(\Delta r,\omega)$, defined by (26) and (18), respectively. Below we discuss the main stages in calculating the radiative damping force (32).

After inserting the integral representations (26) and (18) of the Green's functions in (32), we use the expression for the differential,

$$-d\{\exp(\varkappa\Delta r\beta)\}=\varkappa\beta\frac{\Delta r_l}{\Delta r}\exp(\varkappa\Delta r\beta)\,\dot{r}_l(t_1)\,dt_1,$$

and integrate by parts, taking into account the causality principle and the resulting properties of the Green's function.

Next, using obvious transformations we can write the radiative damping force as

$$\begin{aligned}F_j(t)=&-\frac{e^2}{c^3}\int_0^1 d\beta\int dt_1\int\frac{d\omega}{2\pi}\exp(i\omega z)\left[(i\omega)^2v_j(t_1)\right.\\&\left.-\frac{1}{2}(1-\beta^2)(i\omega)^3\Delta r_j\right]\\&+\frac{e^2}{c^4}v_\alpha(t)\int_0^1 d\beta\,\beta\int dt_1\int\frac{d\omega}{2\pi}\\&\times\exp(i\omega z)\,(i\omega)^2[\Delta r_j v_\alpha(t_1)-\Delta r_\alpha v_j(t_1)]\frac{1}{\Delta r},\end{aligned} \tag{33}$$

where we have introduced the notation

$$z=t_1-t+\frac{\Delta r}{c}\beta,\quad \Delta r_j=r_j(t)-r_j(t_1).$$

The calculations that follow are quite obvious. As result of integrating by parts with respect to ω we arrive at derivatives of the delta functions with respect to the argument z. This is followed by integration with respect to t_1. Finally, integration with respect to the auxiliary parameter β introduced into

the integral representation (18) yields the final expression for the generalized radiative damping force $F_j(t)$ in the relativistic equation (10):

$$F_j(t)=\frac{2}{3}\frac{e^2}{c^3}\frac{1}{1-u}\left(\ddot{v}_j+\frac{1}{2}\frac{u}{1-u}\ddot{v}_j^{(l)}\right)+\frac{e^2}{c^3}\frac{\dot{u}}{(1-u)^2}\left(a_j^{(c)}+\frac{1}{4}a_j^{(l)}+\frac{1}{2}\frac{a_j^{(l)}}{1-u}\right), \quad (34)$$

where we have introduced the notation $v=dr(t)/dt$, $u=v/c$,

$$a_j^{(c)}=\left(\delta_{jl}-\frac{v_jv_l}{v^2}\right)\dot{v}_l=\dot{v}_j-v_j\frac{\dot{v}}{v} \quad (35)$$

is the centripetal acceleration, and

$$a_j^{(l)}=\frac{v_jv_l}{v^2}\dot{v}_l \quad (36)$$

is the longitudinal acceleration. By analogy with (35) and (36) we define the normal and longitudinal components of the second derivative of velocity:

$$\ddot{v}_j^{(c)}=\left(\delta_{jl}-\frac{v_jv_l}{v^2}\right)\ddot{v}_l, \quad \ddot{v}_j^{(l)}=\frac{v_jv_l}{v^2}\ddot{v}_l. \quad (37)$$

Interestingly, the ultrarelativistic ($u\to 1$) expression for the radiative damping force is

$$F_j(t)\simeq\frac{1}{3}\frac{e^2}{c^3}\frac{\ddot{v}_j^{(l)}}{(1-u)^2}+\frac{1}{2}\frac{e^2}{c^3}\frac{\dot{u}a_j^{(l)}}{(1-u)^3}. \quad (38)$$

This force is directed parallel to the velocity of the ultrarelativistic particle.

Let us see how formula (34) for the generalized radiative damping force is related to the well-known relativistic formula (3.12) of Ref. 1, which is obtained, via Lorentz transformations, from the approximate nonrelativistic formula (75.8) of Ref. 2,

$$F_j(t)=\frac{2}{3}\frac{e^2}{c^3}\ddot{v}_j. \quad (39)$$

Since this formula (39) and its relativistic generalization are odd under time reversal, we must isolate in the generalized radiative damping force (34) its odd component. Then

$$F_j^{(\mathrm{odd})}(t)=\frac{2}{3}\frac{e^2}{c^3}\frac{1}{1-u^2}\left\{\ddot{v}_j^{(c)}+\frac{\ddot{v}_j^{(l)}}{(1-u)^2}+\frac{3u\dot{u}}{1-u^2}\left[a_j^{(c)}+\frac{a_j^{(l)}}{1-u^2}\right]\right\}. \quad (40)$$

If we use the appropriate notation, Eq. (40) coincides with the relativistic formula (3.12) of Ref. 1.

What sets the expression (34) for the generalized radiative damping force apart from formula (3.12) of Ref. 1 is the presence of a component that is even under time reversal. This is seen most clearly in the ultrarelativistic limit. The study of new effects associated with this difference is of special interest and requires a separate investigation.

To directly verify the approximate nature of the nonrelativistic formula (75.8) of Ref. 2, we write the exact formula (34) in a reference frame co-moving with the electron. What is remarkable in this transition is that the centripetal acceleration tends to zero as $v\to 0$, while the longitudinal acceleration coincides with the ordinary acceleration. Then from (34) we obtain the nonrelativistic formula for the radiative damping force:

$$\frac{1}{mc}F_j(t)=\tau_0\left(\ddot{u}_j+\frac{9}{8}\dot{u}_j\dot{u}\right), \quad \tau_0=\frac{2}{3}\frac{e^2}{mc^3}, \quad (41)$$

and the corresponding Lorentz equation for the electron in the co-moving reference frame is

$$\dot{u}_j=\frac{e}{mc}E_j(t)+\tau_0\left(\ddot{u}_j+\frac{9}{8}\dot{u}_j\dot{u}\right), \quad (42)$$

where $E_j(t)$ is the strength of the external electric field. What sets (41) apart from the well-known formula (75.8) of Ref. 2 is the presence of an additional nonlinear dependence on acceleration. Equation (31) can be written as follows:

$$\left(1-\frac{9}{8}\tau_0\dot{u}\right)\dot{u}_j=\frac{e}{mc}E_j+\tau_0\ddot{u}_j,$$

or

$$m\left(1-\frac{15}{8}\frac{\tau_0}{c}\dot{v}\right)\dot{v}_j=eE_j(t)+m\tau_0\ddot{v}_j. \quad (43)$$

Then the given effect can be interpreted as the dependence of the effective mass m^* on acceleration:

$$m^*=m\left(1-\frac{9}{8}\frac{\tau_0}{c}\dot{v}\right). \quad (44)$$

The common approach to solving problems that involve radiative damping is to formally use the fact that the radiative damping force is weak.[1] Such a procedure yields a physically meaningful result. However, the weakness of the radiative damping force can be rigorously proved only if quantum effects are taken into account.[3] The physical reason for this lies is that classical electrodynamics has no small parameter. The exact solution of the equations of motion for an electron with allowance for the relativistic force (34) also leads to the self-acceleration paradox. Here the velocity reaches the speed of light in a time τ_0 for.

In conclusion we note one more important feature of the classical theory of radiative damping. As is known, both in classical and quantum systems the processes of relaxation (dissipation) and fluctuation phenomena are of the same physical nature. This fact is reflected, in particular, in the fundamental theorems of statistical physics: the Nyquist–Källén–Welton fluctuation–dissipation theorem and the nonlinear fluctuation–dissipation theorems.[14,15] At the same time, the classical radiative damping problem does not demonstrate such a relation. The physical explanation for this is that in classical electrodynamics there is no such concept as electromagnetic vacuum. For this reason an electron in free space is not subjected to fluctuations. At finite temperatures

there is thermal radiation, which gives rise to an additional contribution to the radiative damping of the electron and at the same time is the reason why the electron fluctuates.

This work was supported by a grant from the Russian Fund for Fundamental Research (Grant No. 98-02-17069).

[1] V. L. Ginzburg, *Theoretical Physics and Astrophysics*, Pergamon Press, Oxford (1979).
[2] L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th ed., Pergamon Press, Oxford (1975).
[3] G. F. Efremov, Zh. Éksp. Teor. Fiz. **110**, 1629 (1996) [JETP **83**, 896 (1996)].
[4] G. W. Ford and R. F. O'Connell, Phys. Lett. A **157**, 217 (1991); **158**, 31 (1991); Phys. Rev. A **44**, 6386 (1991).
[5] J. Seke, J. Phys. A **25**, 5415 (1992).
[6] J. Seke, Phys. Rev. A **45**, 542 (1992).
[7] G. W. Ford and R. F. O'Connell, Phys. Lett. A **174**, 182 (1993).
[8] K. Pachucki, Phys. Rev. A **52**, 1079 (1995).
[9] F. V. Hartemann, Phys. Rev. Lett. **74**, 1107 (1995).
[10] A. I. Nikishov Zh. Éksp. Teor. Fiz. **109**, 1889 (1996) [JETP **82**, 1015 (1996)].
[11] J. D. Jackson, *Classical Electrodynamics*, 2nd ed., Wiley, New York (1975).
[12] G. F. Efremov, *Stochastic Equations for Open Quantum Systems* [in Russian], Gorkiĭ State Univ. Press, Gorkiĭ (1982).
[13] A. A. Abrikosov, L. P. Gor'kov, and I. Ye. Dzyaloshinskiĭ, *Methods of the Quantum Theory of Fields in Statistical Physics*, Prentice-Hall, Englewood Cliffs, N.J. (1963).
[14] G. F. Efremov, Zh. Éksp. Teor. Fiz **55**, 2322 (1968) [Sov. Phys. JETP **28**, 1232 (1969)].
[15] G. F. Efremov, Izv. Vyssh. Uchebn. Zaved. Radiofiz. **15**, 1207 (1972).

Translated by Eugene Yankovsky

Instability and melting of a crystal of microscopic particles in a radio-frequency discharge plasma

I. V. Schweigert*)

Institute of Semiconductor Physics Siberian Branch, Russian Academy of Sciences, 630090 Novosibirsk, Russia

V. A. Schweigert and V. M. Bedanov

Institute of Theoretical and Applied Mechanics Siberian Branch, Russian Academy of Sciences, 630090 Novosibirsk, Russia

A. Melzer, A. Homann, and A. Piel

Institut für Experimentelle und Angewandte Physik, Christian-Albrechts-Universität, 24098 Kiel, Germany
(Submitted 11 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 1672–1690 (November 1998)

The development of instability, heating, and melting of a two-layer crystal of dust particles in the sheath of a radio-frequency discharge has been studied by the method of Langevin molecular dynamics. The interaction forces between particles are determined in terms of a model developed earlier, in which the ion clouds under the upper particles are replaced by effective point charges. Both a pure Coulomb interaction and a screened interaction are considered. Various regimes of particle motion in the crystal are discussed. The experimental and calculated results for the mean energy of particles and the number of defects in the crystals are compared. © *1998 American Institute of Physics.* [S1063-7761(98)00911-1]

1. INTRODUCTION

Interest in the study of the behavior of microscopic particles in a plasma has traditionally been due to their important role in plasma technologies, particularly in plasma-enhanced etching and deposition processes in microelectronics. The essentially simultaneous discovery by several experimental groups[1–3] in 1994 of the formation of a crystal of dust particles in a gas discharge showed that microparticles in a plasma are also a new, unique object that makes it possible to study phenomena lying at the junction of the physics of a nonideal plasma, solid-state theory, and the physics of phase transitions. In a low-temperature plasma, micron-size particles acquire a significant charge of $eZ=(10^3-10^4)e$. Therefore, by altering the density of the microparticles, it is possible to experimentally vary the degree of nonidealness $G=e^2Z^2/a_cT_p$ of the plasma over a wide range, where T_p is the temperature of the particles ($a_c=\sqrt{1/\pi n}$ in the two-dimensional (2D) case and $a_c=(3/4\pi n)^{1/3}$ in the three-dimensional (3D) case) and n is the particle density. According to the theory of a one-component plasma, a system of microparticles must undergo a transition into the crystalline state when $G_*=130-140$ (Ref. 4) and $G_*\approx 170$ (Ref. 5) for the 2D and 3D cases, respectively. The possibility that a crystal of dust particles would appear in a plasma was predicted by Ikezi[6] in 1986, but it was the experimental discovery of such crystals in magnetron[1,7,8] and high-frequency[2,3,9,10] discharges that attracted attention to the problem of highly nonideal systems of charged microparticles. Structural ordering of an ensemble of microparticles has also been observed in a thermal plasma[11] and in the strata of the positive column of a dc glow discharge in gas.[12] A dust crystal in a radio-frequency discharge, which has been most thoroughly studied experimentally, is considered below.

In a radio-frequency discharge, the crystal is located in the sheath of the lower electrode, where the force of gravity is balanced in the vertical direction by the electric field acting on the negatively charged particles. Silicon,[1] carbon,[13] or polymer[10] particles form an extended crystal lattice in the plane of the electrode, constrained by specially designed projections on the electrodes and consisting of over 100×100 unit cells. The crystal consists of several layers in the longitudinal direction.

Even the first observations of dust crystals showed two features of its behavior that did not fit into the classical theory of Coulomb systems. First, along with various types of close packing (face-centered[1] and body-centered[1,13,14]) of multilayer crystals, alignment of the crystals into columns was also observed, in which particles of the lower layer are located directly under the upper particles.[1,2,9,10,13] In this case, the particles form a hexagonal lattice in the transverse plane, which is typical of 2D systems. Recent experiments show that just such packing usually occurs in multilayer crystals for the most thoroughly studied range of particle radius, $R=3-5$ μm. According to experimental observations of a crystal of growing carbon particles in a methane plasma,[13] the transition from a cubic body-centered lattice to one aligned in columns occurs when $R\sim 1$ μm. Note that the theory of Coulomb systems[15–17] or Yukawa systems[18] gives only close packings as stable states of multilayer crystals, the type of which varies, depending on the distance between

layers. As shown in Refs. 19–22, the reason that particles align into columns is that the particle fields focus the trajectories of the ions, whose flux is directed from the quasineutral plasma toward the electrode. Regions of enhanced ion concentration—ion clouds—are formed downstream of the particles. These ion clouds attract the particles of the lower layer and thereby ensure vertical alignment of the particles.

Another unexpected feature of the behavior of the crystal is observed as the pressure is decreased or the discharge power is increased: the multilayer crystal begins to melt, even though $G=10\,000$–$20\,000$ according to estimates of the particles at room temperature, and the crystalline state must set in.[2] According to the calculations of Ref. 23, screening of the charge of the particles in the sheath is too small and cannot cause the melting, as was assumed in Ref. 24. As was explained as a result of experimental observations of the motion of particles,[25,26] a multilayer crystal becomes unstable below a certain critical gas pressure, and the particles begin to oscillate about their equilibrium positions. As the pressure decreases, the mean kinetic energy of the particles reaches some tens of electron volts, and this causes the crystal to melt.

Note that a single-layer crystal remains stable for the same discharge parameters. As shown by Refs. 21 and 22, instability is induced in multilayer crystals by the formation of ion clouds. A dust crystal in the sheath is under the influence of the ion flux and is an open system, in which the energy of directed motion of the ions is transformed into vibrational energy of the particles. On the basis of a linear analysis, Refs. 21 and 22 found that the main characteristics of the instability of a two-layer crystal—the critical gas pressure, the vibrational frequency, the ratio of the vibrational amplitudes, and the phase shift between the particles of the upper and lower layers—qualitatively agree with the experimental data.

Below, a nonlinear analysis of the instability of a dust crystal is carried out by means of Langevin molecular dynamics. The behavior of the mean particle energy, the distribution functions of the particles over velocities, and the autocorrelation functions of the velocities at various pressures are considered for purely Coulomb and screened interactions. The crystal–liquid phase transition is studied, and a comparison is made with the experimental data.

2. MODEL OF A DUST CRYSTAL

In this experiment we study a two-layer crystal of particles with hexagonal symmetry in each layer. The crystal levitates in the sheath of a radio-frequency discharge in helium. The discharge power is 12 W. The dust crystal consists of spherical monodisperse particles with a radius of $R=4.7$ μm and a mass of $M=6.73\times10^{-13}$ kg. The mean distance between particles in the layer, $a\approx450\,\mu$m, and the distance between layers of the two-layer crystal, $d\approx380\,\mu$m, show little variation as the gas pressure varies in the range $P=30$–100 Pa. A more detailed description of the experimental apparatus and the measurement procedure can be found in Refs. 21, 22, and 25.

Conforming to the experimental conditions, we also consider a two-layer crystal of particles with a hexagonal lattice in the plane of the electrodes (xy) and a columnar structure extending toward the electrode (z). According to experimental data from Refs. 21, 22, and 25, the vibrational amplitude of the particles in the plane of the electrodes significantly exceeds the amplitude of the longitudinal oscillations. We therefore consider particle motion only in the transverse plane $\boldsymbol{\rho}_{ik}=(x_{ik},y_{ik})$ for fixed longitudinal coordinates z_k, where subscript i labels the particle in the layer, and $k=1,2$ corresponds to the upper $(k=1)$ or lower $(k=2)$ layer. The motion of a particle in the dust crystal is governed by collisions with the atoms of the gas, as well as by electrostatic interaction with the other particles of the crystal and the spatially inhomogeneous ionic charge distribution in the sheath.

The effect of the electrons is described by introducing a screened interaction potential. Monte Carlo calculations of the motion of ions in the sheath through a two-layer dust crystal[21,22] show that the ion flux screens the charged particles nonsymmetrically, and ion clouds form downstream of the particles. On the basis of calculations of the restoring force on the particles of the lower layer when it is displaced from the equilibrium position, Ref. 22 proposes the analytic model of the crystal used below, in which the actual spatially inhomogeneous ionic charge distribution is replaced by an effective point charge rigidly bound to the particle that generates it. The magnitude and position of the effective charge are determined from a calculation of the restoring force. The model thus includes the interaction of particles with gas atoms, the electrostatic repulsion of particles, and the attraction between particles in the lower layer and point ion clouds located under the upper particles.

The Langevin equations for our system can be written in the form[27]

$$\frac{d^2\boldsymbol{\rho}_{ik}}{dt^2}=\frac{1}{M}\mathbf{F}_{ik}-\nu\frac{d\boldsymbol{\rho}_{ik}}{dt}+\frac{1}{M}\mathbf{F}, \tag{1}$$

where ν is the coefficient of friction of the particles in the gas, and $\mathbf{F}$ is the random Langevin force that describes particle heating (bearing in mind the translational degrees of freedom) as a result of collisions with neutral atoms. The electrostatic forces

$$\mathbf{F}_{ik}=\mathbf{F}_{ik,pp}+\mathbf{F}_{ik,pi}$$

include the repulsion $\mathbf{F}_{ik,pp}$ between the particles and the attraction of the particles of the lower layer toward the ion clouds, $\mathbf{F}_{ik,pi}$. The interparticle forces can be written as

$$\mathbf{F}_{ik,pp}=-\partial U_{pp}/\partial\boldsymbol{\rho}_{ik},$$

where

$$U_{pp}=\sum_{i>j}\sum_{k=1}^{2}U(|\boldsymbol{\rho}_{ik}-\boldsymbol{\rho}_{jk}|)+\sum_{ij}U(|\boldsymbol{\rho}_{i1}-\boldsymbol{\rho}_{j2}+\mathbf{e}_z d|) \tag{2}$$

is the potential energy of the interparticle interaction, $\mathbf{e}_z$ is the unit vector in the longitudinal direction, and d is the distance between layers. The first term in Eq. (2) describes the interaction of the particles in the same layer, while the

second term describes the interaction of particles in different layers. The interaction forces between particles in the lower layer and the effective positive point charges eZ_i located a distance $d-d_i$ under the particles of the upper layer are given by

$$\mathbf{F}_{ik,pi} = -\delta_{k2}\partial U_{pi}/\partial \boldsymbol{\rho}_{ik},$$

where

$$U_{pi} = -\epsilon\sum_{ij} U(|\boldsymbol{\rho}_{i1}-\boldsymbol{\rho}_{j2}+\mathbf{e}_z d_i|), \quad \epsilon = Z_i/Z,$$

and the Kronecker delta δ_{k2} reflects the fact that the attractive force acts only on the lower particles. The Debye–Hückel potential

$$U(\boldsymbol{\rho}_i-\boldsymbol{\rho}_j) = \frac{e^2Z^2}{|\boldsymbol{\rho}_i-\boldsymbol{\rho}_j|}\exp(-\lambda|\boldsymbol{\rho}_i-\boldsymbol{\rho}_j|)$$

is used for the interaction of two point charges in the plasma, where λ is the screening length by the electrons. Note that the particle-interaction model described above is based on non-self-consistent calculations of the spatial distribution of the ion concentration[21,22] and, for that reason, can only give a qualitative description of the heating and melting of an actual dust crystal. To explain the role of the screening, we considered both a strongly screened interaction, $\lambda = 2/a$, and a purely Coulomb interaction, $\lambda = 0$.

The behavior of a dust crystal is governed by the pressure P and temperature T of the gas, the lattice parameters a, d, d_i, and the charges Z and Z_i. By measuring distances in terms of lattice constants a of the crystal and the time in periods $1/\omega_p$ of the plasma oscillations of the crystal, where $\omega_p = \sqrt{4\pi e^2Z^2/Ma^3}$ is the plasma frequency, the number of variables can be reduced. We are interested in the behavior of the crystal at various gas pressures or, in other words, various dimensionless coefficients of friction ν/ω_p. The gas temperature is $T = 300$ K, which determines the amplitude of the Langevin random force; the charges of the particles and the lattice parameters are assumed to be invariant with gas pressure. Under the conditions of the experiment of Refs. 21, 22, and 25, $Z \approx 13\,000e$, the distance between particles is $a \approx 450\ \mu$m, the distance between layers is $d \approx 360\ \mu$m, and the plasma frequency of the crystal is $\omega_p \approx 90\ \text{sec}^{-1}$. The coefficient of friction linearly increases with gas pressure[28] and varies in the range 12–$40\ \text{sec}^{-1}$ for gas pressures $P = 30$–100 Pa. The other parameters $Z_i = 0.5Z$ and $d_i = 0.4a$ are taken from Monte Carlo calculations of the motion of the ions.[22]

Using periodic boundary conditions, we model only a fragment of the crystal, assumed to be infinite in the horizontal plane, including $N = 448$ particles—224 particles in each layer. The calculation region is a rectangle $X \times Y = 14 \times 8\sqrt{3}a^2$, in which a fragment of the hexagonal lattice is inscribed with no distortion. Since $X, Y \gg a$, to determine the forces on particles in a screened potential we took into account only the interaction between particles and their nearest images. For a Coulomb potential ($\lambda = 0$), the interpolation method[29] was used. The interaction between the nearest particles was taken into account directly, while the interaction forces with the images were found by linear interpolation of the forces calculated previously by Ewald's method.[30] The approach of Ref. 31 was used to solve the Langevin equations. By gradually reducing the gas pressure, we considered the equilibrium states of the system for different values of the coefficient of friction. For each gas pressure, we modeled the motion of particles until the characteristics of the system reached their steady-state values, and then averaged over time. The resulting state was used as the initial data for the subsequent gas pressure.

To analyze particle motion apart from trajectories, we considered the behavior of the autocorrelation function of velocity $\mathbf{v}$,

$$Z_{vk}(\tau) = \frac{\Sigma_i\langle \mathbf{v}_{ik}(t)\mathbf{v}_{ik}(t-\tau)\rangle}{\Sigma_i\langle \mathbf{v}_{ik}(t)\mathbf{v}_{ik}(t)\rangle},$$

where $\langle \ldots \rangle$ denotes an average over time, the spectrum of the autocorrelation function,

$$Z_{\omega k}(\omega) = 2\int_0^\infty \cos(\omega\tau)Z_{vk}(\tau)\,d\tau, \tag{3}$$

and the velocity distribution functions of particles in each layer. To determine the percentage of defects in the hexagonal lattice—particles with other than six nearest neighbors—we used Voronoĭ diagrams. In analyzing the crystal–liquid phase transition, we considered the behavior of the pairwise correlation function $g(r)$ and the relative deviation $\mathbf{u}_{ik}$ of particles from their equilibrium positions in the crystal, which was characterized by the modified Lindemann parameter[32]

$$\gamma_k = \frac{2}{Na_c^2}\left\langle \sum_{i=1}^{N/2}\frac{1}{N_b}\sum_{j=1}^{N_b}|\mathbf{u}_{ik}-\mathbf{u}_{jk}|^2\right\rangle,$$

where the inner summation is carried out over nearest neighbors. The displacement of particles in the lower layer relative to the nearest particles in the upper layer was characterized by the parameter

$$\gamma_i = \frac{2}{Na_c^2}\left\langle \sum_{i=1}^{N/2}|\mathbf{u}_{i1}-\mathbf{u}_{i2}|^2\right\rangle.$$

The statistical error of the calculations, which is related to the finite observation time of the system, is indicated on the dependences on the gas pressure of the average energy of the particles in different layers,

$$E_k = M\left\langle \sum_i |\mathbf{v}_{ik}|^2\right\rangle/N,$$

and of the Lindemann parameters. To compare the behavior of a multilayer dust crystal and ordinary Wigner 2D classical crystals, and to check the validity of the algorithm, we modeled a single-layer crystal with the same number of particles in the layer—224—for Coulomb and screened interactions. The single-layer crystal is a Hamiltonian system with a total potential energy

$$U_p = \sum_{i>j} U(|\boldsymbol{\rho}_i-\boldsymbol{\rho}_j|).$$

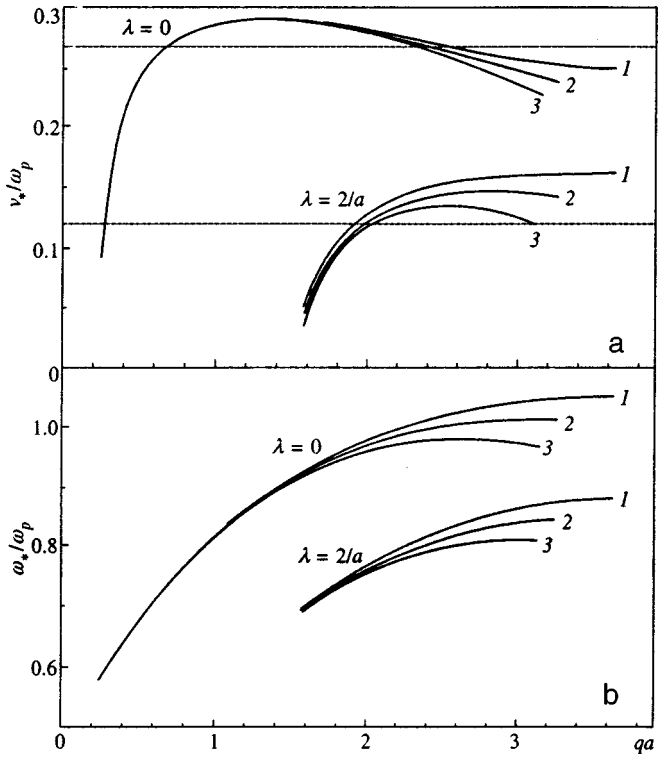


FIG. 1. Critical value of the coefficient of friction (a) and the frequency of the corresponding unstable oscillations (b) vs. the wave vector in various directions: $q_y/q_x=\cos(\pi/6)$ (*1*), $q_y/q_x=\cos(\pi/12)$ (*2*), $q_y=0$ (*3*) for Coulomb ($\lambda=0$) and screened ($\lambda=2/a$) interaction potentials. The dashed horizontal lines show the value of the coefficient of friction corresponding to melting of the crystal.

3. RESULTS AND DISCUSSION

The main feature of a dust crystal of particles is asymmetry in the interaction of the particles of the lower and upper layers. The particles of the lower layer are attracted to the ion clouds that extend downstream of the upper particles, but there is no attraction of the upper particles toward the lower ones. For such an open non-Hamiltonian system, the dynamical matrix of the first derivatives of forces $\mathbf{F}_{ik}$ is not Hermitian and in general has complex eigenvalues η. The eigenvalues λ and eigenvectors $\boldsymbol{\xi}$ of the system of Eqs. (1) for different wave vectors $\mathbf{q}$ were found in Ref. 22 for small deviations

$$\boldsymbol{\chi}_{ik}=\boldsymbol{\xi}(\mathbf{q})\exp(\lambda t+i\mathbf{q}\cdot\boldsymbol{\rho}_{ik})$$

of the particles from their positions $\boldsymbol{\rho}_{ik}$ in the crystal. Since

$$\lambda=(-\nu\pm\sqrt{\nu^2+4\eta})/2, \tag{4}$$

for complex values of η the crystal becomes unstable when the coefficient of friction drops below a certain critical value

$$\nu_*=\mathrm{Im}\,\eta/\omega_*,$$

where $\omega_*=\sqrt{-\mathrm{Re}\,\eta}$ is the frequency of unstable oscillations.[22] For the parameters of the dust crystal considered below, the dependences of ν_* and ω_* on the amplitude of the wave vector are given in Fig. 1 for various directions of $\mathbf{q}$. The most dangerous perturbation, with the maximum coefficient of friction ν_i, corresponds to $q_y/q_x=1/\sqrt{3}$ both for a Coulomb potential ($\nu_i\approx0.290\omega_p,\omega_*\approx0.89\omega_p$), and for a screened potential ($\nu_i\approx0.163\omega_p,\omega_*\approx0.88\omega_p$).

Four characteristic regimes of particle motion can be identified, depending on the value of the coefficient of friction in the dust crystal: (1) $\nu>\nu_i$ corresponds to chaotic motion in the crystal phase, with a mean particle energy several times as large as the gas temperature (Figs. 2a and 2b); (2) $\nu_c<\nu<\nu_i$ corresponds to almost harmonic oscillations of the particles (Fig. 2c); (3) $\nu_m<\nu<\nu_c$ corresponds to chaotic motion about the equilibrium positions in the crystal, with a mean energy several orders of magnitude higher than the gas temperature (Figs. 2d and 2e); (4) $\nu<\nu_m$ corresponds to Brownian particle motion after the crystal–liquid phase transition (Fig. 2f). The critical values ν_i and ν_m, corresponding to the threshold of development of instability and the boundary of the crystal–liquid phase transition, are shown in Fig. 3 by vertical lines. The transition from harmonic oscillations to chaotic motion has no sharp boundary.

3.1. Chaotic motion of particles at high gas pressure

Special calculations at zero gas temperature showed that, when $\nu>\nu_i$, a dust crystal is stable against the development of perturbations and, with the passage of time, the mean kinetic energy of the particles tends to zero. It is well known that in the heat bath comprised of the neutral gas surrounding the particles, any Hamiltonian system tends to equilibrium, and the mean kinetic energy of the particles equalizes with the temperature of the gas. In particular, for a single-layer crystal, this point was used as one validation of the solution of the Langevin equations. In this regard, note that the presence of Brownian particle motion was used in Ref. 24 to explain the experimentally measured low critical values of G_*, corresponding to melting of the crystal. In our opinion, such an explanation is incorrect, since the presence of a thermostat need not affect the equilibrium characteristics of a Hamiltonian system. In particular, our calculations yield $G_*\approx135$ for the Coulomb interaction in a single-layer crys-

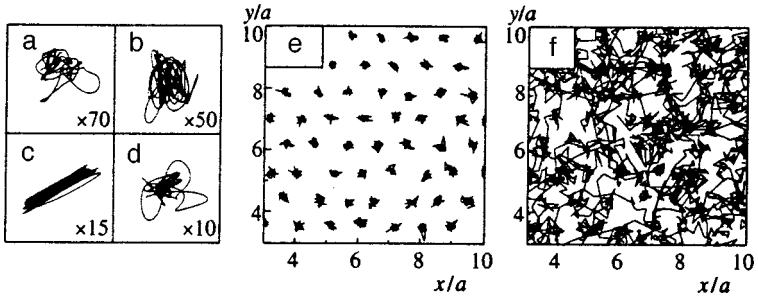


FIG. 2. Particle trajectories in a fragment of the upper layer of the crystal for a screened interaction potential with various coefficients of friction $\nu/\omega_p=0.21$ (a), 0.1625 (b), 0.1575 (c), 0.15 (d), 0.12 (e), and 0.115 (f).

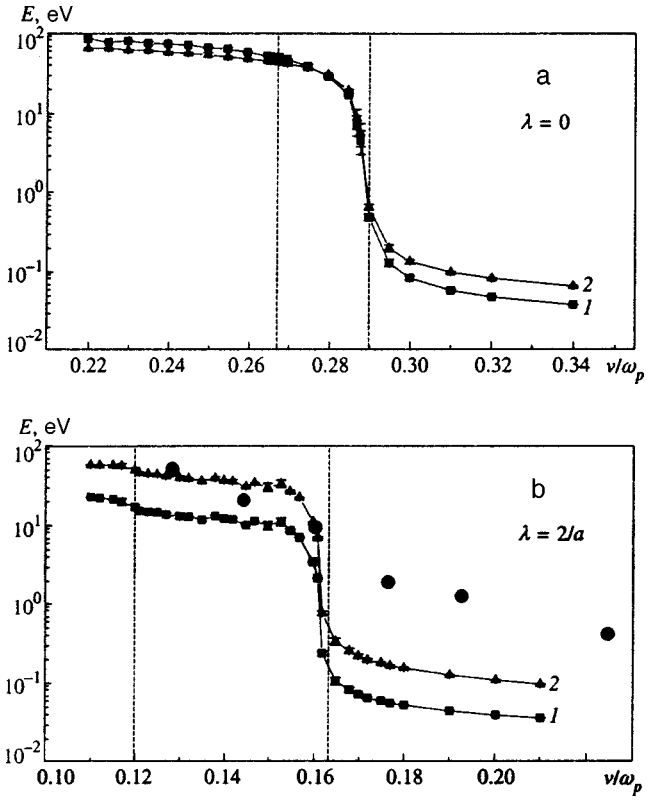


FIG. 3. Particle energies in the upper (*1*, circles) and lower (*2*) layers for Coulomb (a) and screened (b) interaction potentials. The circles correspond to experimental data, and the triangles and squares to calculations. The vertical dashed lines show the instability limits obtained in a linear analysis and the melting thresholds of the crystal.

tal, consistent with the data of other authors obtained using the Monte Carlo method or the method of molecular dynamics.

Energy is also transferred from the ion flux to the multilayer dust crystal for $\nu > \nu_i$, when the crystal is stable against the development of perturbations. The presence of complex eigenvalues of the dynamical matrix reduces the effective coefficient of friction, and the mean kinetic energy of the particles exceeds the gas temperature even for $\nu > \nu_i$ (Fig. 3). The experimentally measured energy of the particles also exceeds the gas temperature, but this difference lies within the measurement error. The particle speed distribution function is essentially Maxwellian (Fig. 4a), and the velocity distribution is isotropic. The mean particle energy in the lower layer is higher than in the upper layer, both for the screened and for the Coulomb interaction potentials, and this agrees with the experimental data[25] and linear analysis.[22] The motion of particles with a comparatively small amplitude of the deviation from the equilibrium position produces negligible broadening of the peaks of the pairwise correlation function, which correspond to the positions of the particles in an ideal hexagonal lattice (Fig. 5a). The presence of modes with complex eigenfunctions is most notable in the spectrum of the velocity autocorrelation function, where the height of the peak corresponding to plasma oscillations increases as the coefficient of friction decreases (Figs. 6 and 7a). The presence of a continuous spectrum of the velocity autocorrelation function and its exponential decrease over several vi-

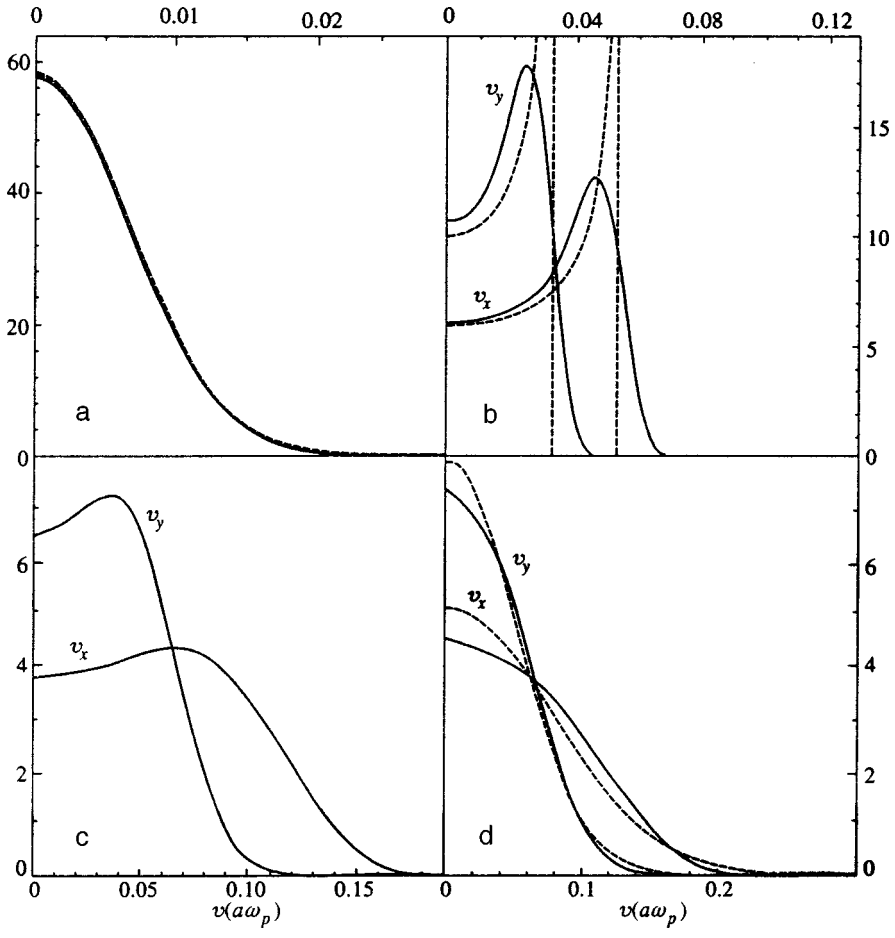


FIG. 4. Distribution function of particles with velocities v_x, v_y in the lower layer of a crystal with a screened interaction potential and various coefficients of friction $\nu/\omega_p = 0.165$ (a), 0.16125 (b), 0.155 (c) and 0.145 (d). The dashed curves correspond to a Maxwellian distribution (a and d) and the velocity distribution accompanying harmonic oscillations (b).

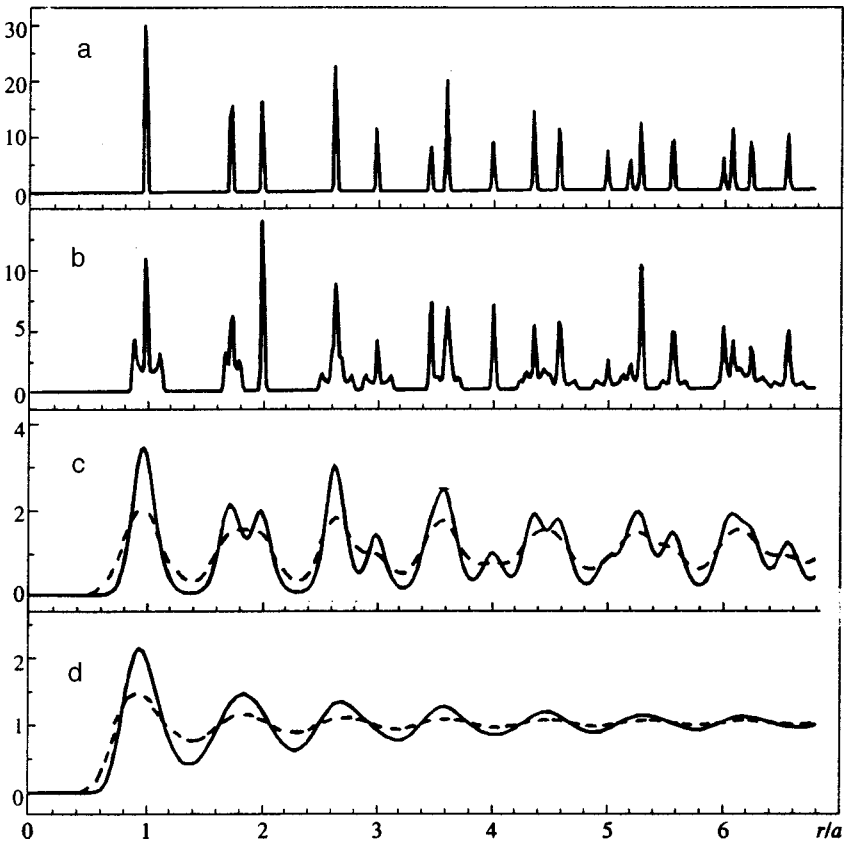


FIG. 5. Two-point correlation function for screened interaction potential witih various coefficients of friction $\nu/\omega_p = 0.21$ (a), 0.1575 (b), 0.125 (c), 0.115 (d). The solid and dashed curves correspond to the upper and lower layers.

brational periods (Fig. 8a) indicates that the particle motion is chaotic. Note that the velocity autocorrelation functions of particles in the upper and lower layers are qualitatively similar for all the regimes of motion.

3.2. Coherent oscillations in the crystal phase

When $\nu < \nu_i$, friction between particles in the gas can no longer stabilize the development of instability, and this leads to a sharp increase in the mean kinetic energy of the particles (Fig. 3). The most interesting feature of the particle motion is the existence of the coherent oscillations, which occur over a narrow range of the coefficient of friction, $\nu_c < \nu < \nu_i$. For the screened potential, $\nu_c \sim 0.155\omega_p$. During coherent oscillations, all particles move with identical phase and frequency ω. For harmonic particle motion, $v_i = v_m \sin(\omega t)$, the particle velocity distribution function is

$$f(v) = \frac{dt}{dv} = \frac{1}{\sqrt{v_m^2 - v^2}}, \tag{5}$$

where v_m is the maximum particle velocity. A comparison of Eq. (5) with the calculations is given in Fig. 4b. The presence of the coherent regime indicates that only one vibrational mode is excited in the crystal.

We consider a finite fragment of the crystal, in which the phonon spectrum is discrete. Therefore, just after the transition through the critical value of the coefficient of friction, only one mode is excited, with a frequency of $0.881\omega_p$. However, as the coefficient of friction decreases further, the number of unstable modes increases. Figure 9 shows the frequency distribution of the unstable modes obtained by linear analysis of the fragment of the crystal considered here. For a screened potential with $\nu = 0.1575\,\omega_p$, allowing for degeneracy, there are 26 unstable modes, whose frequencies are distributed over the range 0.847–0.882. The peak width in the spectrum of the velocity autocorrelation function in the coherent regime is clearly smaller and is determined by the finite integration time $(400/\omega_p)$ in Eq. (3).

Thus, as a result of nonlinear interaction of the various modes when nonlinearity is weak, only one mode is dominant. The situation is analogous in this respect to the lasing of a multimode laser.[33] The presence of the coherent regime shows up in the particle trajectories as essentially harmonic oscillations (Fig. 2c). These oscillations cause two additional maxima to appear near the equilibrium positions in the pairwise correlation function (Fig. 5b); their origin is quite analogous to the appearance of maxima in the velocity distribution function. The velocity autocorrelation function oscillates harmonically in time with essentially constant amplitude (Fig. 8b), and this also indicates that only one mode is dominant. A coherent regime of particle motion also exists for Coulomb interaction, but it is less pronounced.

The available experimental data have yet to confirm the existence of a coherent regime. Experimentally observed crystals differ from our model in two respects. First, the typical number of particles in an actual crystal, $\sim 10^4$, is much greater than the number of particles whose motion we are modeling (448). The spectrum of elementary excitations of an actual crystal is accordingly closer to continuous than is the spectrum of the crystal fragment that we are considering. However, using the analogy with a multimode laser, it stands to reason that the presence of a continuous spectrum should

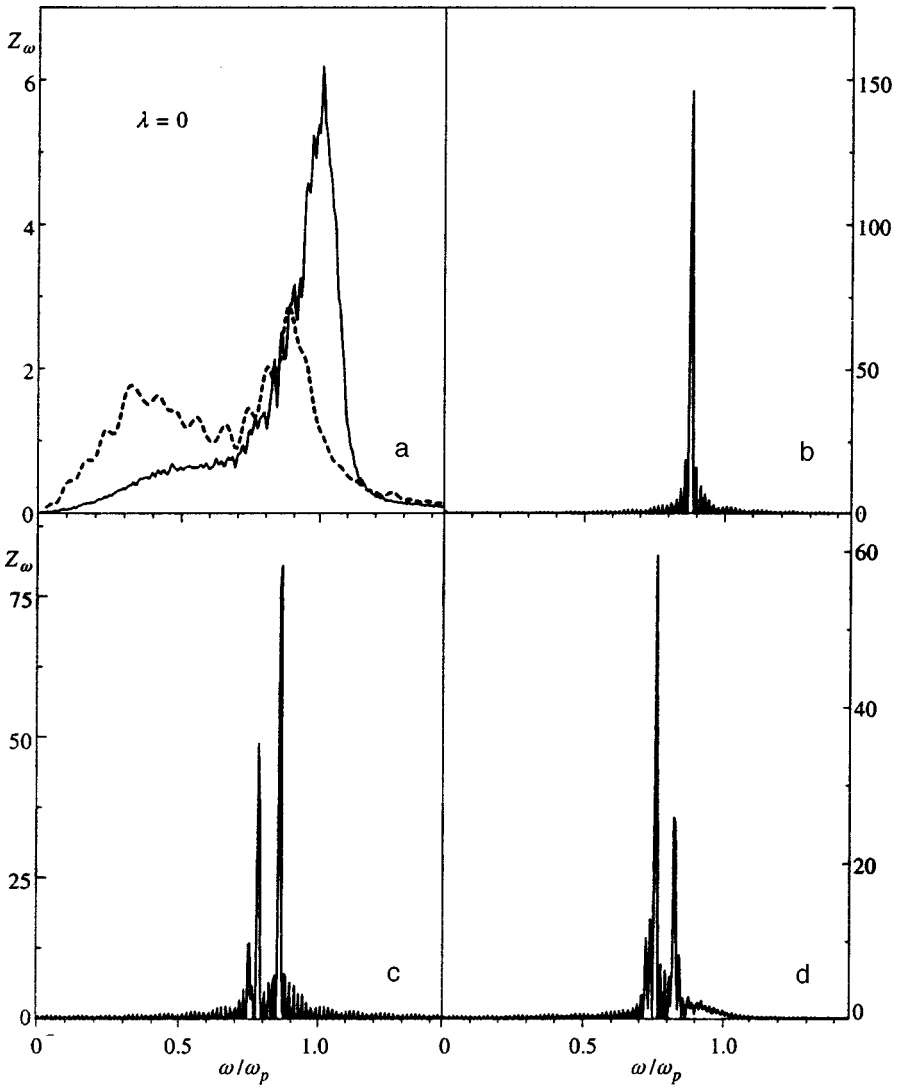


FIG. 6. Spectrum of the particle velocity autocorrelation function of the upper layer for a Coulomb interaction potential with various coefficients of friction $\nu/\omega_p = 0.32$ (a), 0.285 (b), 0.27 (c), 0.23 (d). The dashed curve (a) corresponds to the spectrum of the velocity autocorrelation function of a single-layer crystal with $G = 400$.

not radically affect the existence of the coherent regime. The second problem is more important—there are a large number of defects in an experimentally produced crystal (Fig. 10) even for low particle energies, when $G \gg G_*$, and there must be no defects in the ideal situation. The origin of the defects in the actual crystal is not completely clear. Since the crystal is limited in the plane of the electrodes, hexagonal symmetry must break down in the peripheral regions, and this inevitably causes defects to appear. However, calculations of large 2D Coulomb clusters confined by a parabolic potential indicate a significantly smaller percentage of defects.[34] It is possible that defects in an actual crystal are associated with the spread in particle size. The influence of defects on the particle heating mechanism and the crystal melting has yet to be studied. The experimental data (Fig. 3) show that as the gas pressure decreases, the mean particle energy increases in the actual crystal more smoothly than in the calculations. In this case, the number of defects also increases (Fig. 10), even at particle energies for which the calculations predict no defects. Note that using krypton as a carrier gas makes it possible to reduce the defect level of the crystal and to observe a more abrupt rise in particle energy with decreasing gas pressure.[35] The presence of defects can result in the localization of unstable modes and can substantially reduce intermodal interaction, and this in turn leads to suppression of the coherent regime.

3.3. Chaotic oscillations in the crystalline phase

As the coefficient of friction decreases, the maxima in the particle velocity distribution function gradually broaden (Fig. 4c) and a smooth transition occurs to an almost Maxwellian distribution (Fig. 4d), which remains anisotropic over the directions of the velocities right up to the point at which the crystal melts. The velocity distribution function scarcely differs at all from a Maxwellian just before melting, and becomes isotropic in velocity in the liquid phase. The velocity autocorrelation function decreases with time (Fig. 8d), and its spectrum broadens (Fig. 7c), which further suggests that the particle motion becomes chaotic. The peak in the spectrum of the velocity autocorrelation function corresponding to the frequencies of the unstable modes shifts toward lower frequencies as the coefficient of friction decreases, and this agrees with a linear analysis (Fig. 9b). Because the modes interact, two additional broadened peaks appear when $\lambda = 2/a$; their position approximately corre-

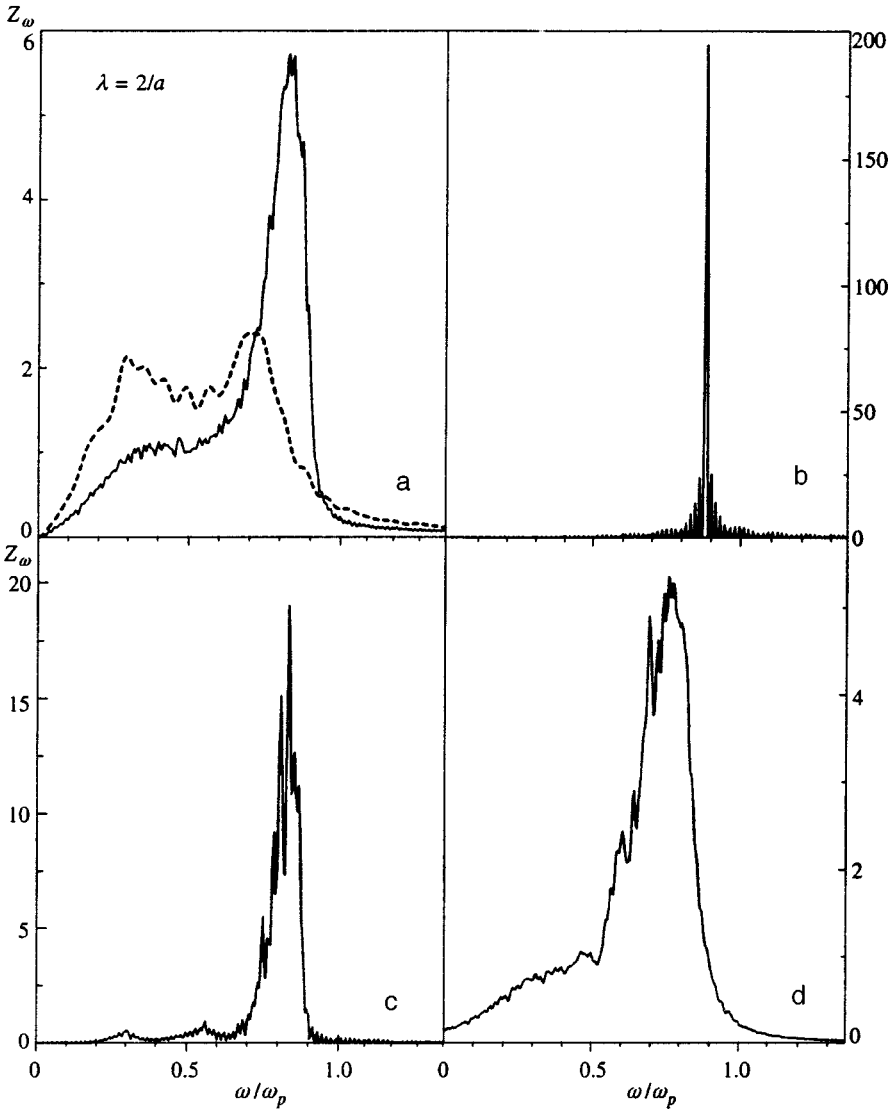


FIG. 7. Spectrum of the particle velocity autocorrelation function of the upper layer for a screened interaction potential with various coefficients of friction $\nu/\omega_p = 0.21$ (a), 0.1575 (b), 0.125 (c), 0.115 (d). The dashed curve (a) corresponds to the spectrum of the velocity autocorrelation function of a single-layer crystal with $G = 400$.

sponds to the subharmonics $\omega/3$ and $2\omega/3$ of the unstable modes (Fig. 7c). For Coulomb interaction with $\nu < \nu_c$, the spectra of the velocity autocorrelation function qualitatively differ from those described above. As the coefficient of friction decreases, a second narrow peak appears, whose amplitude gradually grows and becomes larger than that of the first peak. One of the reasons for the difference in the behavior of the spectra is probably the difference in the density of states of the unstable modes (Fig. 9). The distance between adjacent modes is less for $\lambda = 2/a$ than for $\lambda = 0$, and this must make it easier for the modes to overlap as a result of their interaction.

The mean kinetic energy of the particles continues to grow as the coefficient of friction decreases, with particles having higher energy (E_1) in the upper layer than in the lower (E_2) for the Coulomb potential. The opposite is true for the screened potential, $E_1 < E_2$, which has been observed in experiment.[21,22,25] just before the crystal melts, the peaks in the two-point correlation function, although they broaden, remain fairly distinct (Fig. 5c), especially for the upper layer, where the mean particle energy is lower when $\lambda = 2/a$.

3.4. Melting of the crystal

The melting of the crystal, which was identified in the calculations from the sharp jump in both the number of defects in the system (Fig. 10) and the Lindemann parameter (Fig. 11), occurs when $\nu \approx 0.12\,\omega_p$ for the screened potential and $\nu \approx 0.267\,\omega_p$ for the Coulomb potential. The calculations of Ref. 32 show that the critical value of the Lindemann parameter, corresponding to the melting of 2D crystals, remains approximately the same for different interaction potentials, $\gamma_* \approx 0.1$. For the screened potential, the energy of the particles and, accordingly, the Lindemann parameter is greater in the lower layer than in the upper. However, the crystal melts only when the Lindemann parameter reaches its critical value of about 0.09–0.11 in the upper layer, which is approximately the same for $\lambda = 0$ and $\lambda = 2/a$. Unlike the Coulomb potential, melting of the crystal when $\lambda = 2/a$ is accompanied by a jump in mean particle energy. After the crystal–liquid phase transition, the peak amplitudes of the pairwise correlation function decrease with distance, and

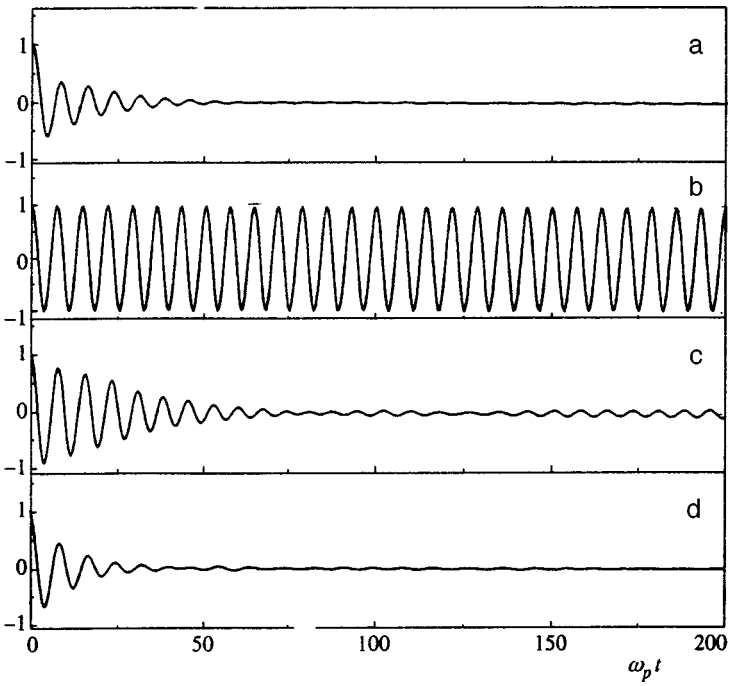


FIG. 8. Particle velocity autocorrelation function of the upper layer for a screened interaction potential with various coefficients of friction $\nu/\omega_p = 0.21$ (a), 0.1575 (b), 0.125 (c), 0.115 (d).

$g(r)$ reaches its asymptotic value $g(r\to\infty)=1$ when r reaches a small multiple of the interparticle distance. The spectrum of the velocity autocorrelation function at $\omega=0$ becomes nonzero, and suggesting particle self-diffusion in the liquid phase (Fig. 7d). Since there is no sharp change in particle kinetic energy after melting for the Coulomb interaction, the broadening of the spectrum of the velocity autocorrelation function and particle self-diffusion after melting is substantially less for $\lambda=0$ than for $\lambda=2/a$.

The limited size of the crystal fragment that we considered leads to no conclusions concerning the order of the phase transition, the presence of an intermediate hexatic phase, or other crucial questions in the theory of the melting of 2D crystals. To judge by recent numerical experiments,[36]

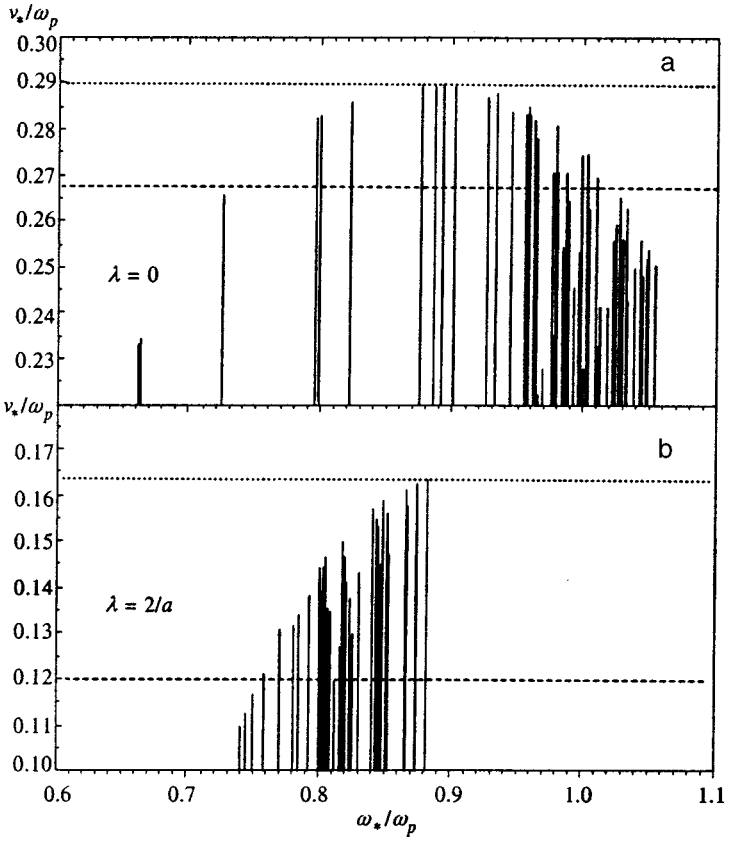


FIG. 9. Unstable modes of the 448-particle crystal fragment under consideration for Coulomb (a) and screened (b) interaction potentials. The upper horizontal dashed lines correspond to the instability threshold, and the lower to melting of the crystal.

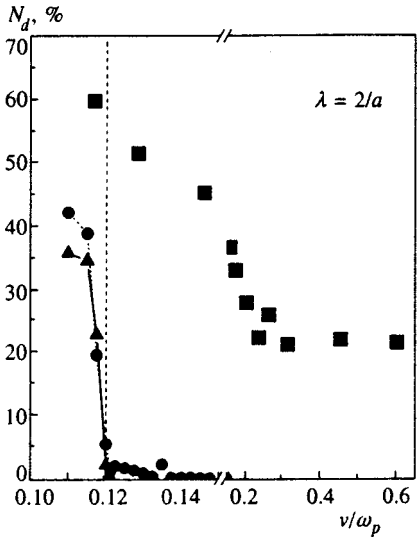


FIG. 10. Dependence on the coefficient of friction of the percentage of defects in the lower (circles) and upper (triangles) layers of the crystal for a screened interaction potential. Squares correspond to experimental results.

this would require the number of particles to be at least an order of magnitude greater. To validate the predicted location of the phase transition itself, we modeled the melting of a single-layer Coulomb crystal and obtained $G_* \approx 135$, which is close to calculations of other authors and the experimental data of Ref. 37. For a screened potential in a single-layer crystal, $G_* \approx 189$. A dust crystal melts when $G_* \approx 19$ and $G_* \approx 52$ for Coulomb and screened potentials, respectively. Thus, multilayer dust crystals are substantially more stable against melting than ordinary Coulomb crystals. Note that 3D or two-layer close-packed Coulomb crystals melt at higher G values than do single-layer ones.[38] The difference in the G_* values for dust crystals and ordinary crystals is probably associated with differing energy distributions over the vibrational modes (Figs. 6 and 7). The main part of the energy in dust crystals is carried by short-wavelength oscillations, whereas the melting mechanism of ordinary 2D crystals is dictated by long-wavelength phonons.[39–42]

4. CONCLUSION

A comparison of the experimental data with numerical modeling of the heating and melting of a crystal shows that the assumed mathematical model yields the correct order of magnitude for particle energies and critical gas pressures corresponding to instability and melting of the crystal. At the same time, certain theoretically predicted results—a sharp increase in particle energy as a result of the development of instability, and a coherent vibrational regime—have yet to be confirmed experimentally. This is probably associated with the large defect content of the crystal, which is observable even at low particle energies. For a more detailed comparison of theory with experiment, it is necessary in the experiment to reduce the defect content of the crystal and in the theoretical calculations to numerically model finite systems in which defects exist when the particles have zero kinetic energy.

The present calculations show that a dust crystal is more stable against melting than ordinary Coulomb crystals. The form of the interaction potential of the particles qualitatively affects the behavior of the autocorrelation functions of the velocity and particle motion after melting. Unlike a Coulomb interaction, the melting of a crystal with screened interaction is accompanied by an abrupt increase in the kinetic energy of the particles.

This work was carried out with financial support from the Russian Fund for Fundamental Research (RFBR) (Grant No. 96-02-19134-a), the RFBR–DFG (Grant No. 96-02-00241-G), and INTAS (Grant No. 94-740).

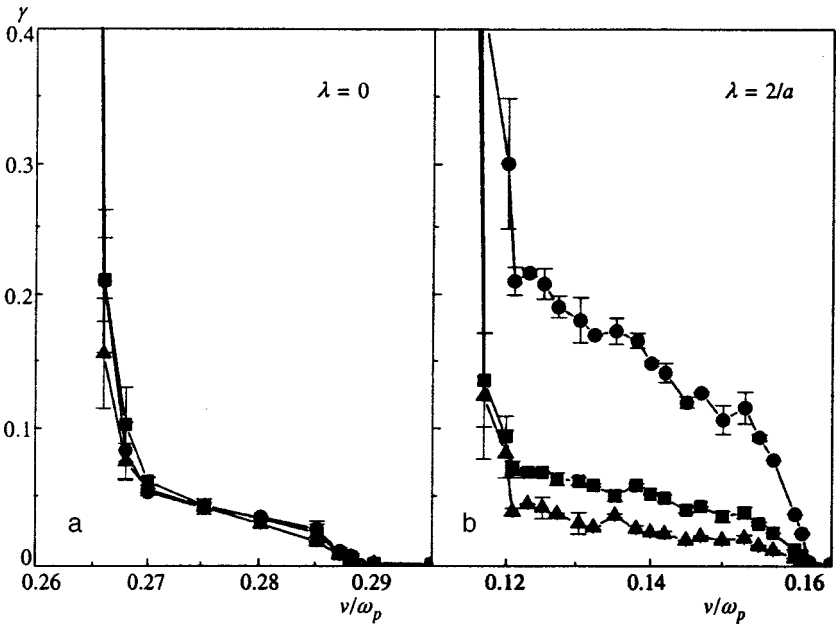


FIG. 11. Root-mean-square relative displacement of particles in the upper (squares) and lower (circles) layers of the crystal for Coulomb (a) and screened (b) interaction potential. Triangles correspond to the Lindemann parameter γ_i, which describes the displacement of upper and lower particles relative to one another.

*)E-mail: ischweig@isp.nsc.ru

[1] J. H. Chu and L. I, Phys. Rev. Lett. **72**, 4009 (1994).
[2] H. Tomas, G. E. Morfill, V. Demmel, J. Goree, B. Feuerbacher, and D. Möhlmann, Phys. Rev. Lett. **73**, 652 (1994).
[3] Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys., Part 2 **33**, L804 (1994).
[4] K. J. Strandburg, Rev. Mod. Phys. **60**, 161 (1988).
[5] R. T. Farouki and A. Hamaguchi, J. Chem. Phys. **101**, 9885 (1994).
[6] H. Ikezi, Phys. Fluids **29**, 1764 (1986).
[7] J. H. Chu and L. I, Physica A **205**, 183 (1994).
[8] J. H. Chu and L. I, J. Phys. D **27**, 296 (1994).
[9] A. Melzer, T. Trottenberg, and A. Piel, Phys. Rev. Lett. **195**, 301 (1994).
[10] T. Trotenberg, A. Melzer, and A. Piel, Plasma Sources Sci. Technol. **4**, 450 (1995).
[11] V. E. Fortov, A. P. Nefedov, O. F. Petrov, A. A. Samaryan, and A. V. Chernyshev, JETP Lett. **63**, 187 (1996); Phys. Rev. E **54**, R 2236 (1996); Zh. Éksp. Teor. Fiz. **111**, 467 (1997) [JETP **84**, 256 (1997)].
[12] V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ *et al.*, JETP Lett. **64**, 92 (1996).
[13] Y. Hayashi and K. Tachibana, J. Vac. Sci. Technol. A **14**, 506 (1996).
[14] J. Pieper, J. Goree, and R. Quinn, J. Vac. Sci. Technol. A **14**, 519 (1996).
[15] D. H. E. Dubin, Phys. Rev. Lett. **71**, 2753 (1993).
[16] V. A. Schweĭgert and M. S. Obrekht, Pis'ma Zh. Tekh. Fiz. **21**(10), 57 (1995) [Tech. Phys. Lett. **21**, 377 (1995)].
[17] G. Goldoni, V. A. Schweigert, and F. M. Peeters, Surf. Sci. **361/362**, 163 (1996).
[18] H. Totsuji, T. Kishimoto, Y. Inoue, C. Totsuji, and S. Nara, Phys. Lett. A **221**, 215 (1996); H. Totsuji, T. Kishimoto, and C. Totsuji, Phys. Rev. Lett. **78**, 3113 (1997).
[19] S. V. Vladimirov and M. Nambu, Phys. Rev. E **52**, 2172 (1995); M. Nambu, S. V. Vladimirov, and P. K. Shukla, Phys. Lett. A **203**, 40 (1995); S. V. Vladimirov and O. Ishihara, Phys. Plasmas **3**, 444 (1996).
[20] F. Melandso and J. Goree, Phys. Rev. E **52**, 5312 (1995); J. Vac. Sci. Technol. A **14**, 511 (1996).
[21] A. Melzer, V. A. Schweigert, I. V. Schweigert, A. Homann, S. Peters, and A. Piel, Phys. Rev. E **54**, R46 (1996).
[22] V. A. Schweigert, I. V. Schweigert, A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **54**, 4155 (1996).
[23] V. A. Schweigert, Pis'ma Zh. Tekh. Fiz. **21**(12), 69 (1995) [Tech. Phys. Lett. **21**, 476 (1995)].
[24] X. H. Zheng and J. C. Earnshaw, Phys. Rev. Lett. **75**, 4214 (1995).
[25] A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1995).
[26] H. Thomas and G. E. Morfill, Nature (London) **379**, 806 (1996).
[27] S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
[28] P. S. Epstein, Phys. Rev. **23**, 710 (1924).
[29] A. Y. Toukmaji and J. A. Board, Comput. Phys. Commun. **95**, 73 (1996).
[30] P. Ewald, Ann. Phys. (Leipzig) **64**, 253 (1921).
[31] H. Gould and J. Tobochnik, *Computer Simulation Methods, Part 2*, Addison–Wesley, Reading, Mass. (1987).
[32] V. M. Bedanov, G. V. Gadiyak, and Y. E. Lozovik, Phys. Lett. A **109**, 289 (1985).
[33] H. Haken, *Synergetics: an Introduction*, Springer-Verlag, Berlin (1977).
[34] V. M. Bedanov and F. M. Peeters, Phys. Rev. B **49**, 2667 (1994).
[35] H. Thomas, private communication.
[36] K. Chen, T. Kaplan, and M. Mostoller, Phys. Rev. Lett. **74**, 4019 (1995).
[37] C. C. Crimes and G. Adams, Phys. Rev. Lett. **42**, 795 (1979).
[38] G. Goldoni and F. M. Peeters, Phys. Rev. B **53**, 4591 (1996).
[39] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[40] B. I. Halperin and D. R. Nelson, Phys. Rev. Lett. **41**, 121 (1978).
[41] D. R. Nelson and D. B. Halperin, Phys. Rev. B **19**, 2457 (1979).
[42] A. P. Young, Phys. Rev. B **19**, 1855 (1979).

Translated by W. J. Manthey

Theory of nonlocal transport for small perturbations in a plasma

V. Yu. Bychenkov,*) V. N. Novikov, and V. T. Tikhonchuk

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia
(Submitted 25 February 1998)
Zh. Éksp. Teor. Fiz. **114**, 1691–1708 (November 1998)

A system of electron-transport equations for small perturbations in a plasma, which is suitable for an arbitrary relation between the electron mean free path and the density gradient length, is formulated. Electron–electron collisions are treated on the basis of the exact Landau collision integral, making it possible to obtain expressions for the potential elements of the Fourier components of electron fluxes, generalized forces, and all the transport coefficients in a plasma with an arbitrary ion charge for the first time. The transport coefficients found, viz., the electrical conductivity, the thermal diffusivity, the thermoelectric coefficient, and the ion convection coefficients, permit the quantitative description of an extensive list of small-scale processes in a completely ionized plasma. The suppression of heat transport and the damping of ion-sound waves in a current-free plasma over the entire range of spatial scales of the perturbations from the high-collisionality limit to the collisionless limit are examined as applications of the theory developed. © *1998 American Institute of Physics.* [S1063-7761(98)01011-7]

1. INTRODUCTION

Hydrodynamic equations provide an effective tool for investigating an extensive list of phenomena in the physics of continuous media. They are considerably simpler than the kinetic equations and faithfully describe processes on scales considerably larger than the mean free path of the particles. The familiar derivation of hydrodynamic equations from the kinetic theory by the Chapman–Enskog method involves expansion of the deviation of the distribution function from thermodynamic equilibrium to first order in the gradients of the hydrodynamic moments[1–3] and leads to local relations between the fluxes and generalized forces. However, this method is confined to the region of very weak spatial gradients, where the density gradient length of the hydrodynamic variables is hundreds of times greater than mean free paths of the particles. Considerably smaller-scale inhomogeneities are of direct interest for numerous important applied problems in plasma physics; therefore, the development of a transport theory that is free of such a constraint is needed. Such a theory leads to nonlocal relations between fluxes and generalized forces and is thus called a nonlocal transport theory. Its development is a pressing problem for numerous areas of plasma physics, such as laser fusion,[4] thermonuclear research in magnetically confined plasmas,[5] astrophysical plasmas,[6] and weakly ionized, low-temperature plasmas.[7]

The general approach to deriving generalized hydrodynamic equations involves writing a hierarchy of equations for the moments of the distribution function of the particles, truncating it at a certain level, and expressing the higher moments in terms of lower moments. The last step is indefinite, since it presumes knowledge of the complete distribution function, while it can be found only in an approximation. For this reason, all the known nonlocal transport theories[8–10] have very limited ranges of applicability and frequently contain empirical parameters. A new method for deriving nonlocal hydrodynamic equations, which, although they contain only low moments of the electron distribution function, are completely equivalent to the kinetic theory for small-amplitude perturbations in a plasma, was proposed and developed in Refs. 11 and 12. This was accomplished via a general solution of the kinetic equation for the electron distribution function with a Landau collision integral for collisions with ions and electrons. The method in Ref. 11 permits the systematic derivation of exact expressions for electron transport coefficients in the Fourier representation, but is restricted to the approximation of a large ion charge, $Z \gg 1$, since electron–electron collisions were taken into account only in the equation for the isotropic part of the distribution function. From the practical standpoint, this constraint is highly significant, since the quantitative description of the transport properties of a plasma is possible only when $Z \gtrsim 10$, while in most practical applications it is necessary to deal with ions having lower extents of ionization.

In this paper we take an important step in perfecting the theory of nonlocal transport, i.e., we find expressions for the electron transport coefficients for a plasma with ions of any charge. This is achieved by systematically taking into account the exact electron–electron collision integral in the Landau form, as was done in the classical theory of transport in a plasma.[2,3] The significant difference from the classical theory is that the ratio between the density gradient length of the perturbations and the electron mean free path can have an arbitrary value. Allowing for electron–electron collisions is a complicated problem not only for an analytical theory, but also for numerical simulation. There are only a few numerical programs for solving a kinetic equation in which electron–electron collisions are taken into account in some approximation,[13–15] while there has hitherto been no theoretical description of nonlocal transport over the entire range of

variation of the collisionality parameter with systematic allowance for electron–electron collisions. The closest thing to our approach in formulation is the work in Ref. 14 on the numerical simulation of the damping of small-amplitude ion-sound perturbations in a homogeneous plasma. This special case is a representative test for our general transport theory.

The method used in this paper to solve the kinetic equation for electrons is valid for an arbitrary ratio between the perturbation scale length $L\sim k^{-1}$ and the mean free path λ_{ei}, since it enables us to take into account the contributions of all the anisotropic components of the electron distribution function and to quantitatively describe the transition from ordinary collisional hydrodynamics to the collisionless limit. A previously proposed[11] procedure for solving the initial-value problem is used to find the general solution for the perturbed electron distribution function for the case of potential perturbations and to determine the relations between the electron fluxes and generalized hydrodynamic forces, which include the electric field, the plasma flow velocity, and the plasma-density and temperature gradients. In the spatial Fourier representation these transport relations have a form similar to the classical expressions,[2,3] but the transport coefficients are found to depend on the wave number (the collisionality parameter, $k\lambda_{ei}$). Such an approach allows us to unequivocally specify the nonlocal kinetic coefficients for a completely ionized plasma with an arbitrary ion charge for the first time. To illustrate the application of the nonlocal hydrodynamic equations obtained, nonlocal heat transport and the damping of ion-sound waves in a current-free plasma in the region of parameters of interest for modern experiments on the interaction of high-power laser radiation with plasmas are considered in this paper.

2. KINETIC DESCRIPTION OF NONLOCAL ELECTRON TRANSPORT

Let us consider a homogeneous plasma with a Maxwellian electron velocity distribution function $F_0(v_e)$, an electron density n_e, and an electron temperature T_e as the ground state. We assume that the plasma is completely ionized and contains ions with the charge number Z. Being interested in phenomena associated with electron transport, for simplicity we shall assume that the ions are cold and neglect the ion–ion collisions. We shall also neglect the transfer of energy in electron–ion collisions.

Let the initial perturbation of the Maxwellian electron distribution be given at $t=0$:

$$\delta f(t=0,\mathbf{r},\mathbf{v}_e)=\left[\frac{\delta n_0}{n_e}+\left(\frac{v_e^2}{2v_{Te}^2}-\frac{3}{2}\right)\frac{\delta T_0}{T_e}\right]F_0(v_e),$$

where $v_{Te}=\sqrt{T_e/m_e}$ and m_e are the electron thermal velocity and mass. We shall examine the linear response of the plasma to this perturbation, as well as to small-amplitude irrotational perturbations of the electric field $\mathbf{E}=-\nabla\phi$ and the ion velocity $\mathbf{u}_i$. In addition, δn_0 and δT_0 can be arbitrary functions of the coordinates $\mathbf{r}$, and ϕ and $\mathbf{u}_i$ can be arbitrary functions of the coordinates and time (provided $\operatorname{curl}\mathbf{u}_i=0$). Owing to the linearity of the response, it is sufficient to confine ourselves to consideration of spatially periodic perturbations of the density, temperature, electric field, and ion velocity associated with the wave vector $\mathbf{k}$. We linearize the kinetic equation for the electron distribution function f_e, taking into account the electron–electron collisions in the form of the Landau collision integral, $C_{ee}[f_e,f_e]$, and the electron–ion collisions in a representation which describes only scattering with respect to the angle θ between the wave vector $\mathbf{k}$ and the electron relative velocity vector $\mathbf{v}=\mathbf{v}_e-\mathbf{u}_i$:

$$C_{ei}[f_e]=\frac{1}{2}\nu_{ei}(v)\frac{\partial}{\partial\mu}(1-\mu^2)\frac{\partial f_e}{\partial\mu}. \tag{2.1}$$

Here $\nu_{ei}(v)=4\pi Zn_0e^4\Lambda/m_e^2v^3$ is the velocity-dependent frequency of electron–ion collisions, $-e$ is the electron charge, Λ is the Coulomb logarithm, and $\mu=\cos\theta$. As a result of plugging the expressions just presented into the kinetic equation for electrons and going over to a reference frame moving with the local ion velocity, we obtain the following equation for the spatial Fourier component of the deviation of the electron distribution function $\delta f(t,k,\theta,v)$ from the equilibrium function $F_0(v_e)$:

$$\frac{\partial\delta f}{\partial t}+ikv\mu\delta f-ikvu_i\mu^2\frac{\partial F_0}{\partial v}+ik\frac{e}{m_e}\phi\mu\frac{\partial F_0}{\partial v}$$
$$=C_{ei}[\delta f]+C_{ee}[\delta f,F_0]+C_{ee}[F_0,\delta f]. \tag{2.2}$$

Here $u_i(t)$ and $\phi(t)$ are the Fourier components of the ion velocity and the electric potential. We next expand the perturbation δf of the distribution function in Legendre polynomials which are eigenfunctions of the electron–ion collision operator (2.1) in the local reference frame moving with the ion velocity u_i:

$$\delta f=\sum_{l=0}^{\infty} i^l f_l(t,k,v)P_l(\mu),$$

$$C_{ee}[\delta f,F_0]+C_{ee}[F_0,\delta f]=\sum_{l=0}^{\infty} i^l C_{ee}^l P_l(\mu), \tag{2.3}$$

and we perform the Laplace transformation in time. As a result of the operations just described, the electron kinetic equation (2.2) is reduced to an infinite system of equations for the angular harmonics f_l of the electron distribution function:

$$pf_l+kv\frac{l}{2l-1}f_{l-1}-kv\frac{l+1}{2l+3}f_{l+1}$$
$$=-\frac{l(l+1)}{2}\nu_{ei}(v)f_l+C_{ee}^l+S_l. \tag{2.4}$$

Here the first term on the right-hand side corresponds to electron–ion collisions, and the second term describes electron–electron collisions in the form of Rosenbluth potentials:[16]

$$\frac{C_{ee}^l}{\nu_{ee}(v)}=\frac{l(l+1)}{6}f_l(I_2^0-3I_0^0-2J_{-1}^0)+v\frac{\partial}{\partial v}\Bigg[f_lI_0^0+\frac{v}{3}\frac{\partial f_l}{\partial v}(I_2^0+J_{-1}^0)\Bigg]+\frac{4\pi v^3}{n_e}f_lF_0+\frac{v^2}{2(2l+1)}\frac{\partial^2F_0}{\partial v^2}\Bigg[\frac{(l+1)(l+2)}{2l+3}(\delta I_{l+2}^l+\delta J_{-1-l}^l)-\frac{l(l-1)}{2l-1}(\delta I_l^l+\delta J_{1-l}^l)\Bigg]+\frac{v}{2(2l+1)}\frac{\partial F_0}{\partial v}\Bigg[\frac{l^2+3l-2}{2l-1}\delta I_l^l+\frac{l(l-1)}{2l-1}\delta J_{1-l}^l-\frac{(l+1)(l+2)}{2l+3}\delta I_{l+2}^l-\frac{l^2-l-4}{2l+3}\delta J_{-1-l}^l\Bigg]. \tag{2.5}$$

Here $\nu_{ee}(v)=\nu_{ei}(v)/Z$ is the velocity-dependent frequency of electron–electron collisions,

$$[I_m^0;\delta I_m^n]=\frac{4\pi}{n_ev^m}\int_0^v[F_0;f_n]v^{2+m}\,dv,$$

$$[J_m^0;\delta J_m^n]=\frac{4\pi}{n_ev^m}\int_v^\infty[F_0;f_n]v^{2+m}\,dv$$

are Rosenbluth potentials, and the source functions S_l are specified by u_i and ϕ, as well as by the Fourier components δn_0 and δT_0, and are nonzero only for $l\leqslant 2$:

$$S_0=\frac{\delta n_0}{n_e}F_0+\frac{3}{2}\frac{\delta T_0}{T_e}\left(\frac{v^2}{3v_{Te}^2}-1\right)F_0-\frac{i}{3}ku_i\frac{v^2}{v_{Te}^2}F_0,$$
$$S_1=\frac{e\phi}{T_e}kvF_0,\quad S_2=\frac{2}{3}iku_i\frac{v^2}{v_{Te}^2}F_0. \tag{2.6}$$

The system of equations (2.4) differs from the one studied in Refs. 11 and 12 through the inclusion of the electron–electron collision terms (2.5) in the equations with $l\geqslant 1$.

Below we shall confine ourselves to a study of fairly slow (quasistationary) processes taking place over characteristic times that are short compared with the electron–electron collision time, for which the first terms on the left-hand sides of (2.4) can be neglected. Bearing in mind the linearity of the kinetic problem being solved and following Refs. 11 and 12, we introduce the basis distribution functions φ_l, which do not depend on the amplitudes of the sources (2.6) in Eqs. (2.4), but reflect specific dependences on velocity, i.e., we seek a general solution for the angular harmonics of the perturbation of the electron distribution function in the form

$$f_l=\frac{e\phi}{T_e}F_0\delta_{l,0}+\frac{\delta n_0}{n_e}F_0\varphi_l^N+\frac{3}{2}\frac{\delta T_0}{T_e}F_0\varphi_l^T-iku_iF_0\varphi_l^R. \tag{2.7}$$

Here $\delta_{n,m}$ is the Kronecker delta, and the three sets of independent basis functions φ_l^A $(A=N,T,R)$ satisfy three similar systems of equations with different sources:

$$kv\left(\frac{l}{2l-1}\varphi_{l-1}^A-\frac{l+1}{2l+3}\varphi_{l+1}^A\right)=-\nu_{ei}(v)\frac{l(l+1)}{2}\varphi_l^A+\frac{1}{F_0}\tilde{C}_{ee}^l+S_l^A, \tag{2.8}$$

where $S_0^N=1$, $S_0^T=v^2/3v_{Te}^2-1$, and $S_0^R=-S_2^R/2=v^2/3v_{Te}^2$ are unitary sources corresponding to perturbations of the density (N), temperature (T), and ion velocity (R) of unit amplitude, and the electron–electron collision contributions $\tilde{C}_{ee}^l$ are specified by Eq. (2.5) after the replacement of f_l by $F_0\varphi_l^A$.

3. INITIAL-VALUE PROBLEM FOR PERTURBATION OF THE ELECTRON DISTRIBUTION FUNCTION

The perturbation of the electron distribution function described by the angular harmonics (2.7) is created by inducing forces proportional to $\mathbf{E}=-i\mathbf{k}\phi$ and $\mathbf{u}_i$ and also depends on the perturbations of the initial density (δn_0) and the temperature (δT_0). Therefore, at first glance, the expression (2.7) does not permit consideration of the hydrodynamic moments δn and δT of the electron distribution function as independent variables, which is necessary for a general formulation of a theory of electron transport. In particular, δn and δT appear in the classical high-collisionality theory. However, as was shown in Ref. 11, obtaining a general solution for the perturbation of the electron distribution function allows a procedure for introducing the hydrodynamic moments δn_e and δT_e as independent variables owing to the linearity of the kinetic equation for δf. The solution of Eqs. (2.4) and (2.8) in general form specifies the perturbation of the electron distribution function at an arbitrary time t as a function of the four quantities $\mathbf{E}$, $\mathbf{u}_i$, δn_0, and δT_0. Therefore, for any time the hydrodynamic moments of the perturbation of the distribution function can be calculated:

$$\delta n_e=4\pi\int_0^\infty dv\,v^2f_0,$$

$$\delta T_e=\frac{4\pi m_e}{3n_e}\int_0^\infty dv\,v^2(v^2-3v_{Te}^2)f_0, \tag{3.1}$$

which can be represented as linear combinations of the initial perturbations δn_0 and δT_0. Accordingly, the initial density and temperature perturbations can be eliminated by expressing them in terms of the instantaneous perturbations and then solving the system of two linear algebraic equations. Thus, the electron distribution function can be expressed in terms of its instantaneous hydrodynamic moments δn_e and δT_e as independent variables.

In the language of the Laplace–Fourier transforms of the angular harmonics of the electron distribution function, the latter statement corresponds to the following relation:

$$f_l = i\frac{eE^*}{kT_e}F_0\frac{J_T^N\varphi_l^T - J_T^T\varphi_l^N}{D_{NT}^{NT}} + \frac{\delta T_e}{T_e}F_0\frac{(J_T^N+J_N^N)\varphi_l^T-(J_T^T+J_N^T)\varphi_l^N}{D_{NT}^{NT}} + \frac{e\phi}{T_e}F_0\delta_{l,0} + iku_iF_0\left(\frac{D_{NT}^{RT}}{D_{NT}^{NT}}\varphi_l^N + \frac{D_{TN}^{RN}}{D_{NT}^{NT}}\varphi_l^T - \varphi_l^R\right), \quad (3.2)$$

which was written using the moment matrix of the isotropic components of the basis functions φ_0^A

$$J_B^A = \frac{4\pi}{n_e}\int_0^\infty v^2\,dv\;\varphi_0^A F_0 S_0^B. \quad (3.3)$$

Here we have introduced the effective electric field

$$E^* = -ik\phi + ikT_e(\delta n/n_e + \delta T/T_e)/e$$

and used the notation

$$D_{AB}^{CD} = J_A^C J_B^D - J_A^D J_B^C.$$

The expression (3.2) was written in terms of the hydrodynamic moments and the basis solutions of Eq. (2.8) and can be used to find the anisotropic correction to the distribution function f_1 and to close the system of hydrodynamic equations.

4. SOLUTION OF THE KINETIC EQUATIONS FOR THE BASIS FUNCTIONS AND NONLOCAL QUASIHYDRODYNAMIC TRANSPORT EQUATIONS

The kinetic equations for the basis functions (2.8) must be solved in order to determine the electron distribution function. They generally form an infinite system, and it is useful to have a definite procedure for its truncation and closure. Such a procedure can easily be established, if it is taken into account for $l \gg 1$ that the first term on the right-hand side in the expression for the lth angular harmonic of the electron–electron collision integral (2.5) is considerably greater than all the other terms. For this reason, beginning at a certain number $l_{\max} \gg 1$, all the equations for the harmonics of the basis functions take on the following simple form:

$$kv\left(\frac{l}{2l-1}\varphi_{l-1}^A - \frac{l+1}{2l+3}\varphi_{l+1}^A\right) = -\nu^*(v)\frac{l(l+1)}{2}\varphi_l^A, \quad l > l_{\max}, \quad (4.1)$$

where we have used the following notation for the renormalized collision frequency:

$$\nu^* = \nu_{ei} + \nu_{ee}(I_0^0 + 2J_{-1}^0/3 - I_2^0/3).$$

Using (4.1) to rewrite the formal relations

$$\varphi_l^A = -\frac{l}{2l-1}\frac{kv}{\nu_l(v)}\varphi_{l-1}^A,$$

$$\nu_l(v) = \frac{1}{2}l(l+1)\nu^*(v) - kv\frac{l+1}{2l+3}\frac{\varphi_{l+1}^A}{\varphi_l^A}, \quad (4.2)$$

we arrive at the recurrence formula

$$\nu_{l-1}(v) = \frac{1}{2}l(l+1)\nu^*(v) + \frac{l^2}{4l^2-1}\frac{k^2v^2}{\nu_l(v)}. \quad (4.3)$$

Since the condition $\nu_l \simeq \nu_{l+1}$ holds for large values of l, we obtain the following approximate expression for the effective collision frequency ν_l:

$$\nu_l(v) = \frac{1}{2}\nu^*(v)H_l\left(\frac{kv}{\nu^*(v)}\right), \quad H_1(x) = \frac{1}{2} + \sqrt{\frac{1}{4}+\frac{y^2}{l^2}}. \quad (4.4)$$

Thus, to find the basis functions it is sufficient to solve a finite number of equations of the form (2.8) with $l \leqslant l_{\max}$ after substituting $\varphi_{l_{\max}+1}^A = -(kv/2\nu_{l_{\max}})\varphi_{l_{\max}}^A$ into the last of them.

We shall seek the solution of the system of equations (2.8), expanding the basis functions φ_l^A in Sonine–Laguerre polynomials:

$$\varphi_l^A = \sum_{n=0}^\infty c_{ln}^A L_n^{1/2}\left(\frac{v^2}{2v_{Te}^2}\right). \quad (4.5)$$

Substituting the expansion (4.5) into the original equations (2.8), we obtain a system of linear equations for the coefficients c_{ln}^A. This system was solved by the matrix-inversion method, which was adapted to the *MATHEMATICA* software applications package.[17] The calculations were performed for $l_{\max}=10$, at which the error associated with closure of the infinite system of equations was less than 1–2%.

In practice, the use of 55–60 Sonine–Laguerre polynomials is sufficient for describing electron transport over the entire range of values of the collisionality parameter up to $\lambda_{ei} \leqslant 10^3$, where the results already coincide with the results in the collisionless limit, which correspond to Z-independent transport coefficients. Such a calculation of the coefficients c_{ln}^A is possible using any up-to-date personal computer. We also note that, as was shown in Ref. 18, the results of the theory of classical high-collisionality transport ($k\lambda_{ei} < 10^{-2}$) are reproduced to within 1–2% when four or five Sonine–Laguerre polynomials are used ($n_{\max}=4$) in the first two equations for the symmetric (f_0) and first anisotropic (f_1) additions ($l_{\max}=2$).

The description of nonlocal transport is based on the use of the solutions found for the basis functions φ_1^A in the expression (3.2) for the first ($l=1$) anisotropic component of the electron distribution function. In this case the expressions for the electric current j and the electron heat flux q_e have the form

$$j = -ie\frac{4\pi}{3}\int_0^\infty dv\,v^3 f_1,$$

$$q = -iT_e\frac{4\pi}{3}\int dv\,v^3\left(\frac{5}{2} - \frac{v^2}{v_{Te}^2}\right)f_1. \quad (4.6)$$

Since, according to (3.2), the function f_1 is already expressed in terms of hydrodynamic variables, closure of the hydrodynamic equations, which is the main problem of transport theory, is solved in the method under consideration in a natural manner.

TABLE I. Values of the transport coefficients σ/σ_0, α/α_0, χ/χ_0, β_j, and β_q (from top to bottom).

$Z\backslash k\lambda_{ei}$	0.01	0.03	0.1	0.3	1	3	10
	0.999	0.994	0.950	0.770	0.394	0.153	0.0516
	0.996	0.975	0.827	0.454	0.0700	−0.0179	−0.0129
1	0.995	0.964	0.775	0.432	0.165	0.0686	0.0257
	0.000303	0.00268	0.0260	0.147	0.469	0.738	0.890
	0.000549	0.00471	0.0386	0.148	0.266	0.233	0.142
	0.998	0.986	0.905	0.672	0.320	0.148	0.0363
	0.991	0.940	0.703	0.322	0.0465	0.00597	−0.0100
2	0.987	0.904	0.604	0.279	0.0972	0.0414	0.0144
	0.000571	0.0049	0.0418	0.194	0.512	0.734	0.899
	0.00131	0.0104	0.0670	0.196	0.285	0.227	0.145
	0.995	0.970	0.853	0.595	0.272	0.105	0.0334
	0.979	0.879	0.580	0.238	0.0351	−0.00466	−0.00557
4	0.961	0.799	0.443	0.180	0.0592	0.0233	0.00866
	0.000923	0.00745	0.0563	0.227	0.538	0.764	0.899
	0.00251	0.0175	0.0895	0.219	0.285	0.231	0.142
	0.990	0.949	0.805	0.541	0.243	0.0945	0.0304
	0.955	0.801	0.481	0.188	0.0301	−0.00146	−0.00352
8	0.913	0.671	0.321	0.121	0.0385	0.0150	0.00565
	0.00127	0.00952	0.0658	0.246	0.552	0.770	0.901
	0.00379	0.0226	0.0980	0.218	0.271	0.218	0.135
	0.983	0.926	0.768	0.508	0.227	0.0881	0.0284
	0.915	0.717	0.406	0.157	0.0275	0.000386	−0.00234
16	0.840	0.546	0.236	0.0852	0.0266	0.0104	0.00398
	0.00150	0.0106	0.0700	0.255	0.560	0.775	0.903
	0.00460	0.0238	0.0939	0.202	0.248	0.200	0.126
	0.971	0.900	0.737	0.486	0.217	0.0844	0.0273
	0.863	0.637	0.348	0.136	0.0258	0.00154	−0.00160
32	0.747	0.436	0.177	0.0625	0.0194	0.00756	0.00298
	0.00159	0.0108	0.0703	0.256	0.563	0.779	0.905
	0.00471	0.0219	0.0828	0.178	0.221	0.180	0.115

Substituting f_1 from (3.2) into the relations (4.6), we can express the Fourier components of the electric current j and the heat flux q_e in terms of generalized hydrodynamic forces: the Fourier components of the effective electric field E^*, the temperature gradient $ik\delta T$, and the plasma flow velocity u_i. We can thus write

$$j=\sigma E^*+\alpha i\,\delta T_e+\beta_j e n_e u_i,$$
$$q_e=-\alpha T_e E^*-\chi i\,\delta T_e-\beta_q n_e T_e u_i, \tag{4.7}$$

where σ is the electrical conductivity, α is the thermoelectric coefficient, χ is the thermal diffusivity, and the $\beta_{j,q}$ are the ion convective transport coefficients. We note that these coefficients are functions of the wave number k, and so the transport relations have the form of convolution integrals in coordinate space. The transport coefficients are defined by the expressions

$$\sigma=\frac{e^2 n_e}{k^2 T_e}\frac{J_T^T}{D_{NT}^{NT}},\quad \alpha=-\frac{e n_e}{k^2 T_e}\frac{J_T^N+J_T^T}{D_{NT}^{NT}},\quad \beta_j=1-\frac{D_{NT}^{RT}}{D_{NT}^{NT}},$$
$$\beta_q=\frac{D_{NT}^{RT}+D_{NT}^{RN}}{D_{NT}^{NT}},\quad \chi=\frac{n_e}{k^2}\frac{2J_T^N+J_T^T+J_N^N}{D_{NT}^{NT}}. \tag{4.8}$$

At first glance, along with the moments J_A^B of the isotropic parts of the basis functions φ_0^A, the relations (4.8) should also contain similar moments of the first anisotropic components φ_1^A as a result of the integration in (4.6). However, the latter can be eliminated by employing the direct integration of Eq. (2.8) for $l=0$, taking into account the conservation of particle number and energy in electron–electron collisions:

$$\int_0^\infty d^3v\; C_{ee}^0=0,\quad \int_0^\infty d^3v\; v^2 C_{ee}^0=0.$$

The transport equations (4.7) reflect the Onsager symmetry properties: the coefficient α is the same in the expressions for the electric current and the heat flux, in agreement with the equalities[11,12] $J_A^B=J_B^A$ for an arbitrary collisionality parameter $k\lambda_{ei}$.

5. NONLOCAL TRANSPORT COEFFICIENTS

Let us now consider the dependence of the nonlocal transport coefficients on the collisionality parameter. For this purpose, we plug the solutions of the kinetic equations for the basis functions φ_0^A into the relations (3.3) and (4.8) and calculate the transport coefficients σ, α, χ, β_j, and β_q. The results of these calculations for various values of $k\lambda_{ei}$ and Z are presented in Table I.

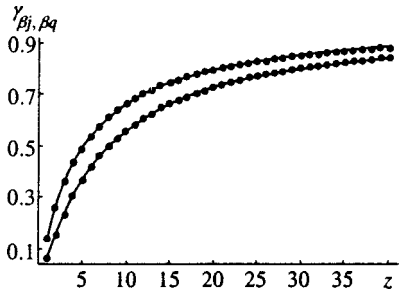


FIG. 1. Dependence of the coefficients $\gamma_{\beta j}$ and $\gamma_{\beta q}$, which characterize ion convective transport in the long-wavelength limit, on ion charge. The solid curves correspond to the approximate formulas (5.5) for $\gamma_{\beta j}$ (upper curve) and $\gamma_{\beta q}$ (lower curve).

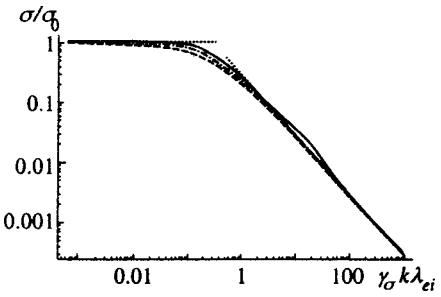


FIG. 2. Dependence of the electrical conductivity σ/σ_0 on the collisionality parameter $\gamma_\sigma(Z)k\lambda_{ei}$ for plasmas with $Z=1$ (solid curve), $Z=4$ (dot-dashed curve), and $Z=64$ (dashed curve). The dotted lines correspond to the classical high-collisionality asymptote and the collisionless limit.

In the high-collisionality limit $\sqrt{Z}k\lambda_{ei}\ll 1$ the longitudinal electrical conductivity σ, the thermoelectric coefficient α, and the thermal diffusivity χ are given by the classical expressions[2,3,18]

$$\sigma_0=\gamma_\sigma(Z)\frac{32e^2n_e\lambda_{ei}}{3\pi m_e v_{Te}},\quad \alpha_0=\gamma_\alpha(Z)\frac{16en_e\lambda_{ei}}{\pi m_e v_{Te}},$$

$$\chi_0=\gamma_\chi(Z)\frac{200n_e v_{Te}\lambda_{ei}}{3\pi}, \tag{5.1}$$

whereas in the short-wavelength limit ($k\lambda_{ei}\gg 1$) they all have similar asymptotes, which are inversely proportional to the wave number and do not depend on the ion charge:[11,12]

$$\sigma=\frac{5e^2n_e v_{Te}}{\sqrt{8\pi}kT_e},\quad \alpha=-\frac{en_e v_{Te}}{\sqrt{2\pi}kT_e},\quad \chi=\frac{4n_e v_{Te}}{\sqrt{2\pi}k}. \tag{5.2}$$

The functions $\gamma_\sigma(Z)$, $\gamma_\alpha(Z)$, and $\gamma_\chi(Z)$ in the relations (5.1), which are equal to unity at $Z\gg 1$, take into account the difference between the exact values of the classical transport coefficients and the corresponding values in the model of a Lorentz plasma, which corresponds to the complete neglect of electron–electron collisions. These functions were calculated and tabulated in Refs. 2 and 18. It is convenient to interpolate them using the following simple expressions[3]

$$\gamma_\sigma(Z)=\frac{0.87+Z}{2.2+Z},\quad \gamma_\alpha(Z)=\frac{0.25+Z}{3.6+Z},\quad \gamma_\chi(Z)=\frac{0.13+Z}{4.7+Z}, \tag{5.3}$$

which reproduce the exact values to within an error of no moe than a few percent.

The ion convective transport coefficients β_j and β_q in Eqs. (4.7) are related if the higher ($l>1$) angular harmonics of the electron distribution function are taken into account, and they do not appear in the classical high-collisionality theory.[2,3] In the long-wavelength limit they are proportional to k^2:

$$\beta_j=22(\gamma_{\beta j}(Z)k\lambda_{ei})^2,\quad \beta_q=87(\gamma_{\beta q}(Z)k\lambda_{ei})^2,\quad k\lambda_{ei}\ll 1. \tag{5.4}$$

In the collisionless limit β_j tends to unity, while β_q falls off as $(\ln k)/k$. We established the dependences of $\gamma_{\beta j}$ and $\gamma_{\beta q}$ on ion charge. They are shown in Fig. 1 and are described well (to within a few percent) by the approximate formulas

$$\gamma_{\beta j}(Z)=\frac{-0.19+Z}{4.9+Z},\quad \gamma_{\beta q}(Z)=\frac{-0.5+Z}{7+Z}, \tag{5.5}$$

which are also shown in Fig. 1. We note that going over to the limit $Z\gg 1$ in formulas (5.4), which corresponds to $\gamma_{\beta j,\beta q}=1$, gives the result of Ref. 11 in the long-wavelength approximation.

The transport coefficients listed in Table I, obtained here for the first time, quantitatively determine the nonlocal transport properties of a plasma in the case of an arbitrary value of Z, in analogy to the classical transport coefficients in the case of local hydrodynamics, which are specified by the functions $\gamma_\sigma(Z)$, $\gamma_\alpha(Z)$, and $\gamma_\chi(Z)$. The almost universal character of the dependence of the normalized (to the classical value) electrical conductivity σ/σ_0 on $\gamma_\sigma(Z)k\lambda_{ei}$, which is illustrated by the curves in Fig. 2, is noteworthy. Similar plots for the thermoelectric coefficient α/α_0 and for the thermal diffusivity χ/χ_0 are shown in Figs. 3 and 4, respectively. The corresponding plots for the ion convective transport coefficients are presented in Fig. 5. In the region $1\leqslant k\lambda_{ei}$ the thermoelectric coefficient α changes sign at a value of k

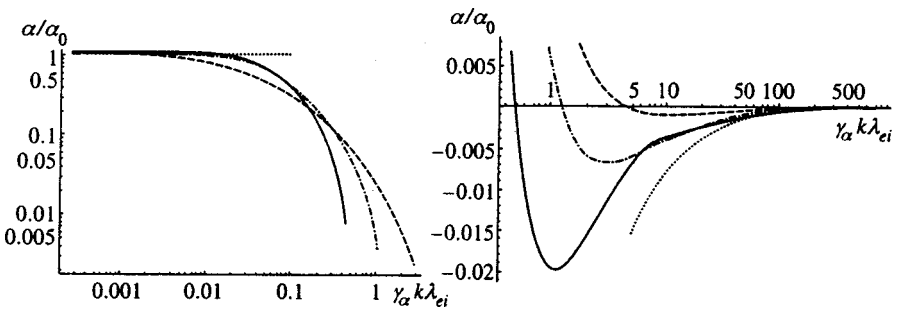


FIG. 3. Dependence of the thermoelectric coefficient α/α_0 on the collisionality parameter $\gamma_\alpha(Z)k\lambda_{ei}$ for plasmas with $Z=1$ (solid curve), $Z=4$ (dot-dashed curve), and $Z=64$ (dashed curve). The dotted lines correspond to the classical high-collisionality asymptote and the collisionless limit.

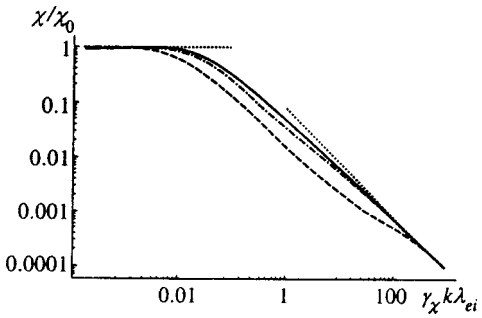


FIG. 4. Dependence of the thermal diffusivity χ/χ_0 on the collisionality parameter $\gamma_\chi(Z)k\lambda_{ei}$ for plasmas with $Z=1$ (solid curve), $Z=4$ (dot-dashed curve), and $Z=64$ (dashed curve). The dotted lines correspond to the classical high-collisionality asymptote and the collisionless limit.

which depends on Z. For example, α vanishes at $k\lambda_{ei}=1.8$ and 4.6 for $Z=1$ and 64, respectively. The position $k_{\max}$ of the maximum of β_q, which can be as large as 0.2–0.3 and corresponds to $k\lambda_{ei}\approx 1$, scarcely depends on the ion charge number Z. Figures 2–4 reveal the monotonic approach of the electron transport coefficients to the collisionless asymptotes, which are plotted as dashed lines in them.

TABLE II. Values of transport coefficients κ/κ_0 and β and the ion-sound damping rate Γ_s/kc_s (from top to bottom).

$Z\backslash k\lambda_{ei}$	0.01	0.03	0.1	0.3	1	3	10
	0.995	0.965	0.791	0.482	0.211	0.0889	0.0325
1	0.000337	0.00287	0.0227	0.0874	0.207	0.294	0.299
	0.265	0.0904	0.0323	0.0169	0.0119	0.0109	0.0108
	0.987	0.906	0.626	0.327	0.132	0.0574	0.0187
2	0.000800	0.00622	0.0376	0.112	0.218	0.200	0.369
	0.181	0.0632	0.0262	0.0161	0.0125	0.0116	0.0119
	0.961	0.801	0.465	0.218	0.0841	0.0339	0.0122
4	0.00152	0.0102	0.0477	0.119	0.210	0.268	0.306
	0.132	0.0497	0.0244	0.0169	0.0139	0.0130	0.0126
	0.910	0.669	0.338	0.149	0.0559	0.0226	0.00828
8	0.00227	0.0126	0.0492	0.112	0.186	0.233	0.265
	0.105	0.0454	0.0260	0.0192	0.0161	0.0148	0.0140
	0.832	0.540	0.247	0.105	0.0389	0.0159	0.00599
16	0.00271	0.0127	0.0439	0.0953	0.156	0.196	0.227
	0.0905	0.0471	0.0296	0.0229	0.0196	0.0173	0.0159
	0.735	0.428	0.183	0.0759	0.0283	0.0117	0.00456
32	0.00270	0.0110	0.0357	0.0765	0.126	0.160	0.191
	0.0872	0.0501	0.0354	0.0294	0.0239	0.0210	0.0185

6. TRANSPORT IN A CURRENTLESS PLASMA AND DAMPING OF ION-SOUND WAVES

As in the classical high-collisionality case, one of the important applications of the quasihydrodynamic equations is the case of no electric current, $j=0$, which describes quasineutral plasma motions. The generalized Ohm's law then permits elimination of the ambipolar electric field

$$E^*=-ik\delta T_e\alpha/\sigma-en_eu_i\beta_j/\sigma \tag{6.1}$$

from the expression for the electron heat flux, which is determined in this case by two transport coefficients:

$$q_e=-\kappa ik\delta T_e-\beta n_eT_eu_i, \quad \text{where } \kappa=\chi-\alpha^2T_e/\sigma,$$

$$\beta=\beta_q-e\alpha\beta_j/\sigma. \tag{6.2}$$

The values of the thermal diffusivity κ and the ion convective transport coefficient β are given in Table II. Both coefficients depend sensitively on ion charge. When the inhomogeneity is weak $k\lambda_{ei}\to 0$, κ transforms into the Spitzer–Härm thermal conductivity:[18]

$$\kappa_0=\gamma_\kappa(Z)128n_ev_{Te}\lambda_{ei}/3\pi.$$

This coefficient is determined by the function $\gamma_\kappa(Z)$, which was tabulated in Ref. 18 and is interpolated well by the analytic expression[14]

$$\gamma_\kappa(Z)=(0.24+Z)/(4.2+Z).$$

The ion convective coefficient β is negligibly small under the conditions of classical transport theory:

$$\beta\propto k^2\lambda_{ei}^2.$$

The dependence of γ_β, which determines the magnitude of the ion convective contribution to the heat flux, on Z practically (to within several percent) coincides with $\gamma_{\beta q}(Z)$, given by (5.5), i.e., $\gamma_\beta(Z)\simeq\gamma_{\beta q}(Z)$. In the collisionless limit, $k\lambda_{ei}\to\infty$, the thermal conductivity has an asymptote similar to (5.2), and β reaches a constant value:

$$\kappa=\frac{18n_ev_{Te}}{5\sqrt{2\pi}k}, \quad \beta=0.4. \tag{6.3}$$

The dependence of the thermal conductivity κ and the ion convective transport coefficient β on the wavelength in plasmas with various values of Z is presented in Fig. 6. The figure also displays straight lines corresponding to the high-collisionality and collisionless asymptotes.

Another application of the theory developed above, which is of practical importance, is the description of the damping of ion-sound waves in the region of intermediate values of the collisionality parameter, where neither the formula from the high-collisionality approximation for the damping rate Γ_s

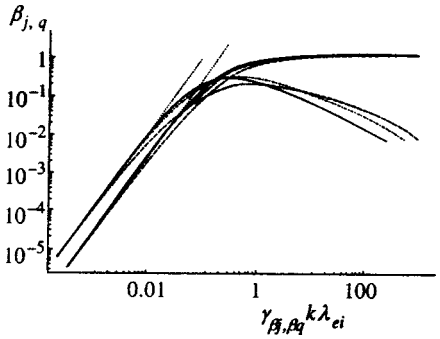


FIG. 5. Dependence of the ion convective transport coefficients β_j and β_q on the collisionality parameters $\gamma_{\beta j}k\lambda_{ei}$ and $\gamma_{\beta q}k\lambda_{ei}$, respectively, for plasmas with $Z=1$ (solid curve), $Z=4$ (dot-dashed curve), and $Z=64$ (dashed curve). The dotted lines are the high-collisionality asymptotes (5.4).

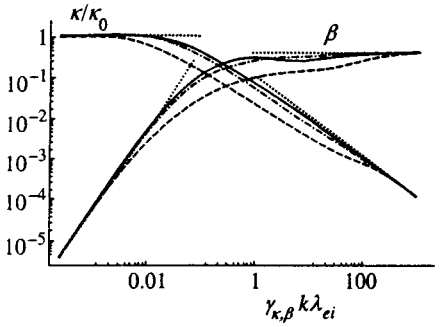


FIG. 6. Dependence of the thermal conductivity κ/κ_0 and the ion convective transport coefficient β on the perturbation wavelength for current-free plasmas with $Z=1$ (solid curve), $Z=4$ (dot-dashed curve), and $Z=64$ (dashed curve). The dotted lines are the high-collisionality and collisionless asymptotes.

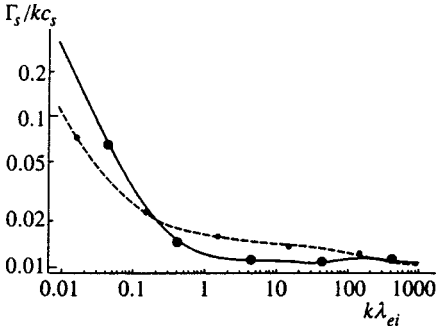


FIG. 7. Wave-number dependence of the damping rate of ion-sound waves for a plasma with $A/Z=2$ (A is the atomic number) when $Z=1$ (solid curve) and $Z=8$ (dashed curve) in comparison to results from numerical simulation[14] (large points — $Z=1$, small points — $Z=8$).

$$\Gamma_s/kc_s=3\pi c_s/256\gamma_\kappa(Z)v_{Te}k\lambda_{ei}$$

($c_s=\sqrt{ZT_e/m_i}$ is the speed of ion sound) nor the collisionless formula

$$\Gamma_s/kc_s=\sqrt{\pi/8}c_s/v_{Te},$$

which corresponds to Landau damping by electrons, is applicable.

In order to describe the damping of ion-sound waves, we utilize the conservation laws fpr particle number and energy to write the relations for the spatial Fourier components of the first two moments of the kinetic equation (2.2) under the assumption of a quasineutral plasma ($j=0$)

$$\frac{\partial\delta n_e}{\partial t}+iku_in_e=0,\quad \frac{\partial\delta T_e}{\partial t}+\frac{2}{3n_e}(ikq_e+n_eT_eu_i)=0, \tag{6.4}$$

as well as the equation of motion for the ions

$$\frac{\partial\delta u_i}{\partial t}=Z\frac{e}{m_i}E^*-ikc_s^2\left(\frac{\delta n_e}{n_e}+\frac{\delta T_e}{T_e}\right)+\frac{Z}{m_in_e}R_{ie}, \tag{6.5}$$

in which the effective electric field E^* is defined by (6.1) and R_{ie} is the force of friction of the ions against electrons:

$$R_{ie}=im_e\frac{4\pi}{3}\int_0^\infty dv\,v^3\nu_{ei}(v)f_1. \tag{6.6}$$

Like the electron fluxes (4.7), the friction force can be expressed using moments of the basis functions:

$$R_{ie}=-\rho_Nen_eE^*+\rho_Tn_eik\delta T_e-\rho_Rm_en_eu_iv_{Te}/\lambda_{ei},$$

$$\rho_N=\sqrt{\frac{\pi}{2}}\frac{1}{k\lambda_{ei}}\frac{J_T^NL_R^T-J_T^TL_R^N}{D_{NT}^{NT}},$$

$$\rho_T=\sqrt{\frac{\pi}{2}}\frac{1}{k\lambda_{ei}}\frac{(J_T^N+J_N^N)L_R^T-(J_T^T+J_N^T)L_R^N}{D_{NT}^{NT}}, \tag{6.7}$$

$$\rho_R=\sqrt{\frac{\pi}{2}}\,kv_{Te}\left(\frac{D_{NT}^{RT}}{D_{NT}^{NT}}L_R^N+\frac{D_{TN}^{RN}}{D_{NT}^{NT}}L_R^T-L_R^R\right),$$

$$L_R^A=\frac{4\pi}{n_e}v_{Te}^2\int_0^\infty dv\,\varphi_1^AF_0.$$

If $Z\gg1$, the relations $\rho_N=1-\beta_j$ and $\rho_T=\beta_q$ hold, and ρ_R transforms into the coefficient β_r from Ref. 11.

Considering perturbations that are $\propto\exp(-i\omega t)$, from the system of equations (6.4), (6.5), and (6.7) supplemented by the expression for the heat flux (6.2) we obtain a dispersion equation, whose weakly damped solution $\omega=kc_s-i\Gamma_s$ describes an ion-sound wave with a damping rate specified by the formula

$$\frac{\Gamma_s}{kc_s}=\frac{n_ec_s}{2k}\left[\frac{(1-\beta)(1-\rho)}{\kappa}+\frac{e^2\beta_j(1-\rho_N)}{T_e\sigma}+\frac{\rho_R}{n_ev_{Te}\lambda_{ei}}\right], \tag{6.8}$$

where $\rho=\rho_T-e\alpha(1-\rho_N)/\sigma$ is an analog of β given by (6.2), in the case of a plasma with finite Z and coincides with the latter when $Z\gg1$. The values of the ion-sound damping rate are given in Table II, and its wave-number dependence is shown in Fig. 7. This figure also presents the results obtained by direct numerical solution of the Fokker–Planck equation for electrons,[14] which, as can be seen from the figure, agree very closely with ours. Even the small deviation from the nonmonotonic decrease in Γ_s/kc_s with increasing k at $100\lesssim k\lambda_{ei}$ seen in Fig. 2 in Ref. 14 for $Z=1$ is fully reproduced by (6.8) (see Fig. 7). Similar variation in the monotonicity of the decreases of the electrical conductivity and β, which determine the ion-sound damping rate, can also be seen at large $k\lambda_{ei}$ in Figs. 2 and 6.

Thus, comparing (6.8) with the results of direct numerical solution of the Fokker–Planck equation[14] and bearing in mind the single-valued relation (6.8) between Γ_s and the nonlocal transport coefficients, we obtain convincing evidence that the theory developed above quantitatively describes nonlocal transport in a plasma. Formula (6.8) describes the smooth transition from the high-collisionality hydrodynamic expression for the ion-sound damping rate to collisionless Landau damping, demonstrating the complete analogy between the proposed quasihydrodynamic equations and the kinetic model.

7. DISCUSSION OF RESULTS

The investigations performed above have shown that there is a broad region of values of the collisionality parameter λ_{ei}/L and Z in which the familiar results in the high-

collisionality and collisionless limits cannot be used to describe transport processes in plasmas, this region being broader, the larger is the charge number. The deviations from the familiar limiting cases are manifested to a lesser extent in the electrical conductivity, and they are displayed more strongly in the thermal diffusivity and the thermal conductivity. We can write the following approximate condition for the collisionality parameter, which specifies the region in which the thermal conductivity deviates significantly from the values predicted by the classical collisional and collisionless theories:

$$\frac{1}{20\sqrt{Z}} \lesssim \frac{\lambda_{ei}}{L} \lesssim 20\sqrt{Z}. \tag{7.1}$$

The region for the electrical conductivity is somewhat smaller: $1/5\sqrt{Z} \lesssim \lambda_{ei}/L \lesssim 5\sqrt{Z}$. We also note that because the thermoelectric coefficient has different signs in the high-collisionality and collisionless limits, it is impossible to obtain even a rough estimate in the transition region by interpolating the known results for these limiting cases.

For modern thermonuclear research the intermediate-collisionality regime, for which the classical transport theories are inapplicable, is typical in practice. It is easy to see[15] that the condition (7.1) generally holds in a near-wall Tokamak plasma.

More examples can be cited for the case of thermonuclear investigations based on lasers in plasmas with electron temperatures from one to several kiloelectron volts. For example, the typical electron mean free path in the critical density region for a neodymium laser irradiating a solid-state target is $\sim 10^{-3}$ cm, while the characteristic spatial scale for the decrease in the electron temperature inside a target is essentially always less than 10^{-2} cm. Even in experiments with ultrashort laser pulses, during which the plasma does not manage to expand and its typical density remains close to the density of the solid, the condition (7.1) can, nevertheless, hold, since the strong skin effect of the laser field creates very large gradients ($L \sim 10^{-5}$ cm). One more interesting example of the appearance of nonlocal heat transport is associated with high-Z plasmas. As was noted above, although the electron mean free path falls off with increasing Z, the onset of the deviation of the thermal conductivity from the Spitzer value[18] is earlier specifically for plasmas with a high extent of ionization of the ions. For this reason, even in the case of fairly smooth irregularities $L > 10^{-2}$ cm a hot plasma ($T_e \sim 3-5$ keV) of hohlraum targets ($n_e \sim 10^{21}$ cm^{-3}) should be regarded as a significantly nonlocal medium, as was pointed out in Ref. 22.

The solution of the problem of controlled laser fusion has been associated with the use of various techniques for smoothing the laser beam.[23] Its structure is characterized by fluctuating irregularities in the intensity of the laser field. The corresponding typical correlation length (the radius of the hot spots) amounts to several wavelengths of the laser radiation and is comparable in order of magnitude to the electron mean free path. Thus, a plasma interacting with such a speckled laser beam has an intermediate-collisionality regime, which differs significantly from the high-collisionality and collisionless regimes. The relaxation of hot spots and the generation of nonthermal fluctuations in such a plasma should have a highly nonlocal character.[24]

Finally, nonlocal effects should also be taken into account in the parametric processes accompanying the interaction of laser radiation with a plasma. They are highly significant for filamentation and stimulated Mandelstam–Brillouin scattering. For example, it was shown in Ref. 25 in reference to filamentation instability that the use of the classical thermal conductivity lowers the filamentation level unacceptably and leads to underestimates of its aftereffects for controlled laser fusion. For this reason, a quantitative description of the nonlocal thermal conductivity of a laser plasma is needed.

8. CONCLUSION

Summarizing the foregoing material, we note that, just as a systematic quantitative theory of electron transport in high-collisionality plasmas was formulated more than 30 years ago, a similar theory, which enables us to quantitatively describe nonlocal transport under the conditions of an arbitrary ratio between the mean free path of the particles and the characteristic spatial scale of potential perturbations, has been developed in this paper. The generalization to the case of rotational perturbations is obvious and can be accomplished in accordance with the method described in Ref. 12.

The ''price'' for the quantitative description of all the nonlocal transport coefficients is the approximation of small perturbations. However, it can be expected that the region of applicability of the theory will be broader than that implied by the constraints imposed. In particular, Vidal *et al.*[19] compared the results of the direct numerical solution of the kinetic equation and the results previously obtained by numerical methods using quasihydrodynamic equations with nonlocal thermal conductivity in Ref. 25. It was discovered in solving the problem of the inverse bremsstrahlung heating of a plasma by laser radiation that the results of the kinetic and quasihydrodynamic approaches are also in good agreement for real (not small) density and temperature fluctuations.

We also note that a comparison of the nonlocal theory of heat transport for small perturbations with the results of the experiment in Ref. 21 (where the perturbations were not small) also exhibited good agreement. This allows us to hope that the nonlocal transport equations obtained above will be of practical use in simulating small-scale processes in plasmas. We stress once again that the quasihydrodynamic equations of nonlocal hydrodynamics obtained are completely equivalent to the kinetic description and are suitable for studying transport over a broad range of values of the collisionality parameter from the high-collisionality region of classical transport to the collisionless limit.

The theory presented has been compared with the numerical kinetic calculations in Ref. 14, and it displays good quantitative agreement with the latter. This demonstrates the advantages of the new model of nonlocal transport developed here, which, unlike kinetic calculations, does not require large numerical resources.

The nonlocal quasihydrodynamic equations derived can

be used to study plasma instabilities and fluctuations, for which transport effects are important. One significant result of our theory is the expression found for the ion-sound damping rate, which is suitable for use at arbitrary values of the charge number of the plasma and the collisionality parameter.

This work was performed with partial financial support from the International Scientific-Technical Center (Project No. 310-96) and the Russian Fund for Fundamental Research (Grant No. 96-02-16165-a).

*)E-mail: bychenk@sci.lebedev.ru

[1] E. M. Lifshitz, L. P. Pitaevskii, and L. D. Landau, *Physical Kinetics*, Pergamon Press, Oxford (1981).
[2] S. I. Braginskiĭ, in *Reviews of Plasma Physics, Vol. 1*, M. A. Leontovich (Ed.), Consultants Bureau, New York (1963) [Russ. original, Atomizdat, Moscow (1963), Vol. 1, p. 183].
[3] R. Balescu, *Transport Processes in Plasmas, Vol. 1*, Elsevier, Amsterdam (1988).
[4] V. Yu. Bychenkov and V. T. Tikhonchuk, in *Nuclear Fusion by Inertial Confinement*, G. Velarde, Y. Ronen, and J. M. Martinez-Val (Eds.), CRC Press (1993).
[5] Z. Chang and J. D. Callen, Phys. Fluids B **4**, 1167 (1992).
[6] P. B. Snyder, G. W. Hammet, and W. Dorland, Phys. Plasmas **4**, 3974 (1997).
[7] E. Furkal, A. I. Smolyakov, and A. Hirose, Bull. Am. Phys. Soc. **42**, 1822 (1997).
[8] J. F. Luciani, P. Mora, and J. Virmont, Phys. Rev. Lett. **51**, 1664 (1983).
[9] J. R. Albritton, E. A. Williams, I. B. Bernstein, and K. P. Swartz, Phys. Rev. Lett. **57**, 1887 (1986).
[10] S. I. Krasheninnikov, Phys. Rev. B **5**, 74 (1992).
[11] V. Yu. Bychenkov, W. Rozmus, V. T. Tikhonchuk, and A. V. Brantov, Phys. Rev. Lett. **75**, 4405 (1995).
[12] A. V. Brantov, V. Yu. Bychenkov, V. T. Tikhonchuk, and V. Rozmus, Zh. Éksp. Teor. Fiz. **110**, 1301 (1996) [JETP **83**, 716 (1996)].
[13] J. P. Matte and J. Virmont, Phys. Rev. Lett. **49**, 1936 (1982).
[14] E. M. Epperlein, Phys. Plasmas **1**, 109 (1994).
[15] O. V. Batishchev, S. I. Krasheninnikov, P. J. Catto, A. A. Batishcheva, D. J. Sigmar, X. Q. Xu, J. A. Byers, T. D. Rognlien, R. H. Cohen, M. M. Shoucri, and I. P. Shkarofskii, Phys. Plasmas **4**, 1672 (1997).
[16] I. P. Shkarofsky, T. W. Johnston, and M. A. Bachynski, *The Particle Kinetics of Plasmas*, Addison-Wesley, Reading, MA (1966).
[17] S. Wolfram, MATHEMATICA, A System for Doing Mathematics by Computer, Addison-Wesley, Redwood City (1988).
[18] L. Spitzer and R. Härm, Phys. Rev. **89**, 977 (1953).
[19] E. Vidal, J. P. Matte, M. Casanova, and O. Larroche, Phys. Plasmas **2**, 1412 (1995).
[20] E. M. Epperlein, Phys. Rev. Lett. **65**, 2145 (1990).
[21] D. S. Montgomery, O. N. Landen, R. P. Drake, K. G. Estabrook, H. A. Baldis, S. H. Batha, K. S. Bradley, and R. J. Procassini, Phys. Rev. Lett. **73**, 2055 (1994).
[22] J. S. De Groot, K. G. Estabrook, S. H. Glenzer, W. L. Kruer, and J. P. Matte, Bull. Am. Phys. Soc. **42**, 1993 (1997).
[23] A. J. Schmitt, Phys. Fluids **31**, 3079 (1988).
[24] V. K. Senecha, A. V. Brantov, V. Yu. Bychenkov, and V. T. Tikhonchuk, Phys. Rev. E **57**, 978 (1998).
[25] E. M. Epperlein and R. W. Short, Phys. Fluids B **4**, 2211 (1992).

Translated by P. Shelnitz

Theory of light-induced drift of a binary gas mixture in a capillary

V. G. Chernyak*) and E. A. Vilisova

Ural State University, 620083 Ekaterinburg, Russia
(Submitted 21 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 1709–1722 (November 1998)

The drift of a binary gas mixture in capillaries induced by resonant light is studied theoretically. The surface and bulk mechanisms of flow of the mixture components are analyzed for arbitrary values of the Knudsen number and the rate parameter (the ratio of the radiative decay rate of an excited molecular level to the intermolecular collision rate). Finally, the theoretical results are compared with the experimental data. © *1998 American Institute of Physics.* [S1063-7761(98)01111-1]

1. INTRODUCTION

The phenomenon of light-induced drift in bulk, predicted by Gel'mukhanov and Shalagin,[1] involves the occurrence of a directed flow of a gas that is in a mixture with a buffer gas and absorbs light selectively with respect to the velocities of the molecules. Since the interaction of excited particles of the absorbing gas and the particles of the buffer gas differs from that of unexcited particles and buffer-gas particles, the components of the mixture flow in opposite directions. The kinetic theory of this phenomenon in the approximation of unrestricted, spatially uniform gas was developed by Dykhne and Starostin.[2]

The wall of the vessel containing the gas mixture may also play the role of the buffer gas. In view of the difference in the interactions of excited and unexcited particles of the absorbing gas and the boundary between the phases, a surface light-induced drift develops.[3]

The momentum conservation law implies that in the case of unrestricted gas the hydrodynamic (average-mass) flow of the mixture is zero. But when there is an interface between the phases, the total momentum of the gas mixture changes, stimulating its macroscopic flow. In particular, in the Knudsen regime, where the mean free path of the molecules is much longer than the capillary radius, the buffer gas is at rest and the drift of the absorbing gas is responsible for the average-mass flow of the gas mixture. Hence in the theory of gas-mixture drift in a capillary it is advisable to determine the flows of the absorbing and buffer gases.

The kinetic theory of bulk and surface light-induced drift of a binary gas mixture for arbitrary Knudsen numbers (Kn) was developed in Ref. 4. The main limitation of this theory is the assumption that the radiative decay constant Γ_m for an excited level of a molecule of the absorbing gas is much smaller than the intermolecular collision rate γ_n. The results of Ref. 4 were obtained in the linear approximation in the small parameter $\Gamma_{mn}=\Gamma_m/\gamma_n$. This approximation usually works for a molecular gas ($\Gamma_m \sim 10^4$ Hz). For an atomic gas ($\Gamma_m \sim 10^7$–10^8 Hz) the approximation is valid only at high pressures, 10^4 Pa and higher.

Atutov *et al.*[5] experimentally studied light-induced drift of sodium vapor in a mixture with inert gases. In their experiments the rate parameter Γ_{mn} varied from approximately 0.3 at high pressures to roughly 150 at low pressure. Therefore it is not possible to compare the theoretical results of Ref. 4 with the experimental data of Ref. 5 over a broad range of Knudsen numbers. Such a comparison would be of interest, on the one hand, for verifying theoretical models and, on the other, for establishing the numerical values of transport and accommodation characteristics of excited molecules.

In the present paper we develop a theoretical model for the bulk and surface mechanisms of light-induced drift of a binary gas mixture in a capillary for arbitrary values of the Knudsen number and the rate parameter Γ_{mn}.

2. STATEMENT OF THE PROBLEM

Let us consider a binary gas mixture that fills a capillary whose length L is much greater than its radius R_0 (Fig. 1). The flow profiles depend solely on the radial coordinate, and end effects can be ignored.

A traveling light wave propagating along the axis z of the capillary is absorbed by the particles of the active component of the mixture in an electronic (for atoms) or vibrational–rotational (for molecules) transition from the ground state n to an excited state m. The frequency ω of the monochromatic light is offset from, ω_{mn}, the center of the absorption line, by $\Omega=\omega-\omega_{mn} \ll \omega, \omega_{mn}$. Because of the Doppler effect, the only particles that interact with the light are those of the absorbing component whose projections of the velocity $\mathbf{v}$ on the wave vector $\mathbf{k}$ are close to the resonance value, for which $\mathbf{k}\cdot\mathbf{v}=\Omega$. The excited and unexcited particles of the absorbing component have the same mass, $m_m=m_n=m_1$, but different effective diameters, $d_m \neq d_n$. Hence a binary gas mixture in which one component interacts with resonant light can be considered a three-component mixture.

The velocity distribution functions f_m and f_n of the excited and unexcited particles of the absorbing gas at $\Omega \neq 0$ become asymmetric about $v_z=0$ because of the occurrence of a Bennett peak and dip,[6] respectively, near the resonant velocity $v_z=\Omega/k$. Hence along the capillary oppositely directed macroscopic fluxes of excited (J_m) and unexcited (J_n) particles develop, which give rise to a total flux of the

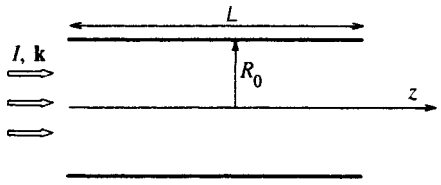


FIG. 1. Geometry of the problem.

absorbing component, $J_1 = J_m + J_n$ (light-induced drift) if the interactions of the excited and unexcited with the capillary surface and the buffer gas are different. Here and in what follows the label 1 refers to the characteristics of the absorbing gas and the label 2, to the characteristics of the buffer gas.

The distribution functions of the excited and unexcited particles of the gas (f_m and f_n, respectively) and of the buffer gas (f_2) satisfy the following system of kinetic equations:

$$\mathbf{v}\cdot\nabla f_m = \tfrac{1}{2}\kappa(\mathbf{v})\Gamma_m(f_n - f_m) - \Gamma_m f_m + S_m,$$

$$\mathbf{v}\cdot\nabla f_n = -\tfrac{1}{2}\kappa(\mathbf{v})\Gamma_m(f_n - f_m) + \Gamma_m f_m + S_n,$$

$$\mathbf{v}\cdot\nabla f_2 = S_2, \tag{1}$$

here

$$\kappa(\mathbf{v}) = \frac{4|G_{mn}|^2\Gamma}{\Gamma_m[\Gamma^2 + (\Omega - \mathbf{k}\cdot\mathbf{v})^2]}, \quad G_{mn} = \frac{E_0 d_{mn}}{2\hbar},$$

$$S_m = S_{mm} + S_{mn} + S_{m2}, \quad S_n = S_{nm} + S_{nn} + S_{n2},$$

$$S_2 = S_{2m} + S_{2n} + S_{22},$$

where Γ_m is the radiative decay constant, Γ is the homogeneous halfwidth of the absorption line, S_{ij} is the Boltzmann collision integrals for the particles of the ith and jth species, E_0 is the electric field amplitude, d_{mn} is the dipole moment of the n–m transition, $\hbar$ is Planck's constant, and $\kappa(\mathbf{v})$ the saturation parameter, which characterizes the probability of induced transitions and is proportional to the light intensity I.

For the boundary conditions we take the specular–diffuse reflection model. When the interaction is elastic, a fraction ε_i of the particles of the ith species are diffusely scattered at each point of the capillary surface with a Maxwellian velocity distribution, while the remaining particles $1-\varepsilon_i$ are specularly reflected. Then the distribution functions satisfy the boundary conditions

$$f_i^+(\mathbf{v}) = \varepsilon_i f_i^s(\mathbf{v}) + (1-\varepsilon_i) f_i^-(\mathbf{v} - 2(\mathbf{v}\cdot\mathbf{n})\mathbf{n}),$$

$$\mathbf{v}\cdot\mathbf{n} > 0, \quad f_i^s = n_i^s\left(\frac{m_i}{2\pi k_B T}\right)^{3/2}\exp\left(-\frac{v^2}{\overline{v_i}^2}\right),$$

$$\overline{v_i} = \left(\frac{2k_B T}{m_i}\right)^{1/2}, \quad i = n,m,2, \tag{2}$$

where $\mathbf{n}$ is the inner normal to the capillary wall, m_i is the mass of the particles of the ith species, T is the gas temperature, k_B is the Boltzmann constant, the superscripts $+$, s, and $-$ denote, respectively, the reflected particles, the particles diffusely emitted by the surface, and the particles incident on the surface, and n_i^s stands for the number density of diffusely scattered particles of the ith species.

Consider the case where the saturation parameter is small, $\kappa(\mathbf{v}) \ll 1$, which usually corresponds to moderate light intensities I. Then the density of the excited particles is much lower than that of the unexcited ones, $n_m \ll n_n$, and the distribution functions differ only slightly from equilibrium Maxwellian distributions:

$$f_i = f_{i0}[1 + h_i(\tilde{\mathbf{r}}, \mathbf{v})],$$

$$f_{i0} = n_{i0}\left(\frac{m_i}{2\pi k_B T}\right)^{3/2}\exp\left(-\frac{v^2}{\overline{v_i}^2}\right), \quad i = n,m,2, \tag{3}$$

where n_{i0} is the equilibrium number density of the particles of the ith species, $\tilde{\mathbf{r}}$ is the radius vector in a plane perpendicular to the capillary axis z, and $h_i(\tilde{\mathbf{r}}, \mathbf{v})$ are unknown perturbation functions.

For an optically thin medium, the dependence of the perturbation functions on the longitudinal coordinate z can be ignored. We also assume that the light intensity is uniform over the capillary cross section. In this way we exclude the possibility of light-induced striction or expulsion of particles by light.

We limit ourselves to the case where the density of the absorbing particles is much less than that of the buffer particles, $n_1 \ll n_2$. Here we can ignore collisions between particles of the absorbing gas and can assume that the effect of light-induced drift is due solely to the bulk (buffer) and surface (accommodation) mechanisms. Then the system of kinetic equations (1), reduced to dimensionless form and linearized with respect to the perturbations of the distribution functions and with respect to the small parameters n_m/n_n, n_m/n_2, and n_n/n_2, takes the form

$$\mathbf{c}_{m\perp}\cdot\frac{\partial h_m}{\partial \mathbf{r}} - \frac{n_n}{n_m}R_m\Gamma_{mm}\frac{\kappa(\mathbf{v})}{2} + R_m\Gamma_{mm}h_m = L_m,$$

$$\mathbf{c}_{n\perp}\cdot\frac{\partial h_n}{\partial \mathbf{r}} + R_n\Gamma_{mn}\frac{\kappa(\mathbf{v})}{2} - \frac{n_m}{n_n}R_n\Gamma_{mn}h_m = L_n,$$

$$\mathbf{c}_{2\perp}\cdot\frac{\partial h_2}{\partial \mathbf{r}} = L_2, \tag{4}$$

where

$$\mathbf{c}_i = \frac{\mathbf{v}}{\overline{v_i}}, \quad c_{i\perp}^2 = c_{ir}^2 + c_{i\varphi}^2, \quad \mathbf{r} = \frac{\tilde{\mathbf{r}}}{R_0},$$

$$R_i = \frac{R_0\gamma_{i2}}{\overline{v_i}}, \quad \Gamma_{mi} = \frac{\Gamma_m}{\gamma_{i2}}, \quad i = n,m,2,$$

$\mathbf{c}_{i\perp}$ is the vector component of the dimensionless velocity of the particles of the ith species in the cross section of the capillary, γ_{ij} is the effective rate of elastic collisions of the particles of the ith species with the particles of the jth species, and R_i is the rarefaction parameter for the ith component, which is inversely proportional to the Knudsen number (see Eq. (7)).

We use the second approximations for the linearized collision integrals L_i (Ref. 7), retaining the terms that are odd in c_{iz}, which contribute to longitudinal gas flow. Ignoring subtle effects like isothermal heat transfer, we obtain

$$L_i = R_i[-h_i + 2c_{iz}(1-\varphi_{i2}^{(1)})u_i + 4c_{ir}c_{iz}(1-\varphi_{i2}^{(3)})\pi_{irz}],$$

$$i = n,m,$$

$$L_2 = R_2\left\{-h_2 + 2c_{2z}\left[u_2 + \sqrt{\frac{m_1}{m_2}}\left(\frac{n_m}{n_2}\psi_{m2}^{(1)}u_m + \frac{n_n}{n_2}\psi_{n2}^{(1)}u_n\right)\right] + 4c_{2r}c_{2z}\left[(1-\psi_{22}^{(3)}+\psi_{22}^{(4)}) + \frac{n_m}{n_2}\psi_{m2}^{(4)}\pi_{mrz} + \frac{n_n}{n_2}\psi_{n2}^{(4)}\pi_{nrz}\right]\right\}, \quad (5)$$

where

$$u_i = \frac{U_i}{\overline{v}_i} = \int c_{iz}E_i h_i d\mathbf{c}_i,$$

$$\pi_{irz} = \frac{P_{irz}}{2p_i} = \int c_{ir}c_{iz}E_i h_i\, d\mathbf{c}_i,$$

$$E_i = \pi^{-3/2}\exp(-c_i^2), \quad i = n,m,2,$$

$$\varphi_{j2}^{(k)} = \frac{\nu_{j2}^{(k)}}{\gamma_{j2}}, \quad \psi_{j2}^{(k)} = \frac{\nu_{j2}^{(k)}}{\gamma_{22}}, \quad j = m,n, \quad (6)$$

with U_i, P_{irz}, and p_i the partial velocity, stress tensor, and pressure of the ith component, respectively. The expressions for the rates $\nu_{j2}^{(k)}$ in terms of particle masses and Chapman–Cowling integrals are given in the Appendix.

The product $C = R_i\Gamma_{mi}$ in the system of equations (4) is a parameter that is independent of the gas pressure and is determined solely by the microscopic characteristics of the absorbing gas, its temperature, and the capillary radius: $C = R_0\Gamma_m/\overline{v}_1$. This parameter characterizes the ratio of the time for an excited particle to traverse the capillary cross section to the lifetime of this particle in the excited state. Estimates of C for atomic gases ($\Gamma_m \sim 10^7 - 10^8$ Hz) in the SI system of units yield $C \approx 10^5 R_0$. For molecular gases ($\Gamma_M \sim 10^4$ Hz) estimates yield $C \approx 10R_0$, with R_0 measured in meters. Assuming that the range of possible values of the capillary radius R_0 is $(0.1\text{–}2)\times 10^{-3}$ m, we can say that $C \approx 10-250$ for light-induced drift for atomic gases and $C \approx 0.001-0.1$ for molecular gases. Then the rate parameter $\Gamma_{mn} = C/R$ for molecular gases with $R > 0.1$ can be considered small. Hence the theory of light-induced drift of single-component gas developed in Ref. 8, in which Γ_{mn} was assumed to be much smaller than unity, is in satisfactory agreement with the experimental data on light-induced drift of CH_3F in a capillary for all values of the Knudsen number. However, for atomic gases the rate parameter Γ_{mn} over a broad range of pressures cannot be assumed small, so that for $\Gamma_{mn} \ll 1$ the model proposed in Ref. 4 of light-induced drift of a binary gas mixture does not provide a qualitative description of the experimental data on the drift of sodium vapor in helium of Ref. 5.

Let us select the effective collision rate γ_{j2} in the form $\gamma_{j2} = \nu_{j2}^{(1)}$, $j = m,n$, and the rate γ_{22}, by analogy with the Bhatnagar–Gross–Krook model for a single-component gas, in the form $\gamma_{22} = \nu_{22}^{(3)} - \nu_{22}^{(4)}$. Then for molecules in the hard-sphere model the rarefaction parameter $R \equiv R_n$ is linked to the Knudsen number as follows:

$$R = \frac{8}{2\sqrt{\pi}}\frac{m_2}{m_1+m_2}\frac{1}{\mathrm{Kn}}, \quad \mathrm{Kn} = \frac{l_n}{R_0},$$

$$l_n = \frac{1}{\pi n_2 d_{n2}^2}\sqrt{\frac{m_2}{m_1+m_2}}, \quad \frac{R_m}{R} = \left(\frac{d_{m2}}{d_{n2}}\right)^2,$$

$$\frac{R_2}{R} = \frac{3}{5}\left(\frac{d_2}{d_{n2}}\right)^2\sqrt{\frac{2(m_1+m_2)}{m_2}}, \quad d_{i2} = \frac{d_i+d_2}{2}, \quad (7)$$

with $i = n,m$, where l_n is the mean free path of the unexcited particles of the absorbing gas in the gas mixture.

The boundary conditions for the perturbation functions h_i follow from Eqs. (2) and (3):

$$h_i^+(\mathbf{r}_0,\mathbf{c}) = (1-\varepsilon_i)h_i^-(\mathbf{r}_0,\mathbf{c}) + \varepsilon_i\frac{n_i^s - n_{i0}}{n_{i0}},$$

$$\mathbf{r}_0 = \frac{\tilde{\mathbf{r}}_0}{R_0}, \quad |\tilde{\mathbf{r}}_0| = R_0, \quad i = m,n,2. \quad (8)$$

The second term on the right-hand side of Eq. (8) can be dropped, since it contributes nothing to the macroscopic velocity and stress tensor in (6).

We consider the case of almost diffuse scattering of particles by the capillary surface, i.e.,

$$1 - \varepsilon_i \ll 1, \quad i = m,n,2. \quad (9)$$

The resulting flows of the absorbing and buffer gases averaged over the capillary cross section are given by the following formulas:

$$J_1 = J_m + J_n = 2\overline{v}_1\int_0^1 (n_n u_n + n_m u_m) r\, dr,$$

$$J_2 = 2\overline{v}_2\int_0^1 n_2 u_2 r\, dr. \quad (10)$$

For numerical calculations it is convenient to use the dimensionless quantities G_1 and G_2, which are related to the flows J_1 and J_2 by

$$J_i = \frac{n_1 R_0 \Gamma_m \kappa}{2\sqrt{\pi}} G_i, \quad i = 1,2,$$

$$\kappa = \int_{-\infty}^{\infty} c_{1z}\exp(-c_{1z}^2)\,\kappa(\mathbf{v})\, dc_{1z}. \quad (11)$$

The quantity κ can be calculated numerically for all values of the parameters $\Gamma/(k\overline{v}_1)$ and $\Omega/(k\overline{v}_1)$. A discussion of the main aspect of this problem can be found in Ref. 9.

3. SOLUTION OF THE KINETIC EQUATIONS

We will use the integral–moment method to solve the system of equations (4). The method is based on the trans-

formation of the integro-differential kinetic equation for the distribution function into a system of integral equations for the moments of this function.

Allowing for the boundary conditions (8), we formally integrate the inhomogeneous linear equations (4) along the direction of the vector $\mathbf{c}_{i\perp}$ of the dimensionless particle velocity.[4] Then, using the above expressions for the perturbation functions h_i and the definitions (6) for the macroscopic quantities, we arrive at three systems of integral equations for the dimensionless velocities u_i and stress tensors π_{irz} of the absorbing ($i=m,n$) and buffer ($i=2$) components of the gas mixture. We allow for the fact that the fraction of specularly reflected particles is small. Then to first order in the small parameter $1-\varepsilon_i$ we have

(1) for the excited particles of the absorbing gas,

$$u_m(r)=\frac{R_m}{\pi}\int_\Sigma D_0\,d\mathbf{r}', \tag{12}$$

$$\pi_{mrz}(r)=\frac{R_m}{\pi}\int_\Sigma D_1\frac{\mathbf{e}\cdot\mathbf{r}}{r}d\mathbf{r}', \tag{13}$$

where

$$D_j=\frac{C\kappa}{2R_m\sqrt{\pi}}\frac{n_n}{n_m}\left(\frac{T_j}{|\mathbf{r}-\mathbf{r}'|}-\frac{K_{jm}}{|\mathbf{r}_N-\mathbf{r}'|}\right)+2(1-\varphi_{m2}^{(3)})\pi_{mrz}(r')$$
$$\times\left(\frac{T_{j+1}\mathbf{e}}{|\mathbf{r}-\mathbf{r}'|}+\frac{K_{j+1,m}\mathbf{e}_0}{|\mathbf{r}_N-\mathbf{r}'|}\right)\cdot\frac{\mathbf{r}'}{r'},\quad j=0,1; \tag{14}$$

(2) for the particles of the absorbing gas in the ground state,

$$u_n(r)=-Su_m(r)+\frac{1}{\pi}\int_\Sigma Q_0\,d\mathbf{r}', \tag{15}$$

$$\pi_{nrz}(r)=-S\pi_{mrz}(r)+\frac{1}{\pi}\int_\Sigma Q_1\frac{\mathbf{e}\cdot\mathbf{r}}{r}d\mathbf{r}', \tag{16}$$

where

$$Q_j=\frac{C\kappa}{2\sqrt{\pi}}\left[\frac{n_n}{n_m}S\left(\frac{T_j}{|\mathbf{r}-\mathbf{r}'|}-\frac{P_j}{|\mathbf{r}_N-\mathbf{r}'|}\right)-\frac{T_j}{|\mathbf{r}-\mathbf{r}'|}+\frac{K_{jn}}{|\mathbf{r}_N-\mathbf{r}'|}\right]$$
$$+2R_n(1-\varphi_{n2}^{(3)})\pi_{nrz}(r')\left(\frac{T_{j+1}\mathbf{e}}{|\mathbf{r}-\mathbf{r}'|}-\frac{K_{j+1,n}\mathbf{e}_0}{|\mathbf{r}_N-\mathbf{r}'|}\right)\cdot\frac{\mathbf{r}'}{r'}$$
$$+2SR_m(1-\varphi_{m2}^{(3)})\pi_{mrz}(r')\left(\frac{T_{j+1}\mathbf{e}}{|\mathbf{r}-\mathbf{r}'|}-\frac{P_{j+1}\mathbf{e}_0}{|\mathbf{r}_N-\mathbf{r}'|}\right)\cdot\frac{\mathbf{r}'}{r'},$$
$$j=0,1; \tag{17}$$

(3) for the buffer gas,

$$u_2(r)=\frac{R_2}{\pi}\int_\Sigma Z_0\,d\mathbf{r}', \tag{18}$$

$$\pi_{2rz}(r)=\frac{R_2}{\pi}\int_\Sigma Z_1\frac{\mathbf{e}\cdot\mathbf{r}}{r}d\mathbf{r}', \tag{19}$$

where

$$Z_j=\left[u_2(r')+\sqrt{\frac{m_1}{m_2}}\left(\frac{n_m}{n_2}\psi_{m2}^{(1)}u_m(r')+\frac{n_n}{n_2}\psi_{n2}^{(1)}u_n(r')\right)\right]$$
$$\times\left(\frac{T_j}{|\mathbf{r}-\mathbf{r}'|}-\frac{K_{j2}}{|\mathbf{r}_N-\mathbf{r}'|}\right)+2\left[\frac{n_m}{n_2}\psi_{m2}^{(4)}\pi_{mrz}(r')\right.$$
$$\left.+\frac{n_n}{n_2}\psi_{n2}^{(4)}\pi_{nrz}(r')\right]$$
$$\times\left(\frac{T_{j+1}\mathbf{e}}{|\mathbf{r}-\mathbf{r}'|}-\frac{K_{j+1,2}\mathbf{e}_0}{|\mathbf{r}_N-\mathbf{r}'|}\right)\cdot\frac{\mathbf{r}'}{r'},\quad j=0,1. \tag{20}$$

In (12)–(20) we used the notation

$$\mathbf{e}=\frac{\mathbf{r}-\mathbf{r}'}{|\mathbf{r}-\mathbf{r}'|},\quad \mathbf{e}_0=\frac{\mathbf{r}_N-\mathbf{r}'}{|\mathbf{r}_N-\mathbf{r}'|},$$

$$S=\frac{C}{(\tau_m-R_n)}\frac{n_n}{n_m},\quad \tau_m=R_m(1+\Gamma_{mm}),$$

$$T_p(t)=\int_0^\infty x^p\exp\left(-x^2-\frac{t}{x}\right)dx,$$

$$K_{pi}=\frac{\mathbf{r}_M\cdot\mathbf{r}}{r^2}(1-\varepsilon_i)T_p[q(y_1+y_2)],\quad i=m,n,2,$$

$$P_p=\frac{\mathbf{r}_M\cdot\mathbf{r}}{r^2}\{(1-\varepsilon_n)T_p[R_n(y_1+y_2)]$$
$$+(\varepsilon_n-\varepsilon_m)T_p(\tau_m y_1+R_n y_2)\},$$

$$y_1=|\mathbf{r}_N-\mathbf{r}'|,\quad y_2=|\mathbf{r}_M-\mathbf{r}|.$$

The argument t of the functions T_p in Eqs. (12)–(14) is $t=\tau_m|\mathbf{r}-\mathbf{r}'|$, in Eqs. (15)–(17) $t=R_n|\mathbf{r}-\mathbf{r}'|$, and in Eqs. (18)–(20) $t=R_2|\mathbf{r}-\mathbf{r}'|$. The argument q of the functions K_{pi} in Eqs. (12)–(14) is $q=\tau_m$, in Eqs. (15)–(17) $q=R_n$, and in Eqs. (18)–(20) $q=R_2$. Integration in (12)–(20) is over the cross-sectional area Σ of the capillary.

To solve Eqs (12)–(20), which are linear Fredholm integral equations of the second kind, we employ the Bubnov–Galerkin method,[10] since this method allows us the determine the coordinate-averaged flows J_1 and J_2 without calculating the velocity and stress profiles. The method requires specifying the approximating expressions for the unknown quantities.

The values of u_i and π_{irz} in the Knudsen regime (Kn $\gg 1$) are determined by the absolute terms of Eqs. (12)–(20). Hence we can expect rapid convergence of the Bubnov–Galerkin method if the trial functions are chosen on the basis of the form of the macroparameters in the hydrodynamic regime (Kn$\ll 1$):

$$\tilde{u}_i=a_{1i}+a_{2i}r^2,\quad \tilde{\pi}_{irz}=a_{3i}r,\quad i=m,n,2, \tag{21}$$

where a_{ki} are unknown constants that depend on the Knudsen number, the accommodation coefficients ε_i, and the molecular parameters. It has been found[4,11] that approximations of the form (21) produce sufficiently accurate results (with

an error less than about 3%) for fluxes averaged over the cross section of the capillary at all values of the Knudsen number.

Substituting (21) in Eqs. (12)–(20) and requiring that the resulting expressions be orthogonal to each basis function (1 and r^2 for (12), (15), and (18), and r for (13), (16), and (19)), we obtain a system of linear algebraic equations for the unknowns a_{1i}, a_{2i}, and a_{3i}. Here the orthogonality condition for two arbitrary functions f and g has the form

$$2\pi\int_0^1 f(r)g(r)r\,dr=0.$$

Since there is very little difference between effective diameters of excited and unexcited particles of the absorbing gas, we have

$$\frac{\Delta d}{d_{n2}}\ll 1, \quad \Delta d=d_{m2}-d_{n2}. \tag{22}$$

After linearizing the expressions (11) in the parameters $1-\varepsilon_i$ and $\Delta d/d_{n2}$ for the flow of the absorbing gas (light-induced drift) we find that

$$J_1=\frac{n_1R_0\Gamma_m\kappa}{2\sqrt{\pi}}\left(G_1^{(1)}\Delta\varepsilon+G_1^{(2)}\frac{\Delta d}{d_{n2}}\right), \quad \Delta\varepsilon=\varepsilon_n-\varepsilon_m, \tag{23}$$

where $G_1^{(1)}$ and $G_1^{(2)}$ are the kinetic coefficients characterizing, respectively, the surface and bulk mechanism of light-induced drift. The solution of the system of integral–moment equations (12)–(20) by the Bubnov–Galerkin method with the trial functions (21) yields

$$G_1^{(1)}=A_1-\frac{\Phi}{\pi(\pi/2+\Phi D)}\left(B_1-\frac{C_1\Phi}{\pi/2+\Phi W}\right), \tag{24}$$

$$G_1^{(2)}=A_2-\frac{\Phi}{\pi(\pi/2+\Phi D)}\left(B_2-\frac{C_2\Phi}{\pi/2+\Phi W}\right),$$

$$\Phi=\frac{4}{5}\frac{m_2}{m_1+m_2}-1. \tag{25}$$

The quantities A_1, A_2, B_1, B_2, C_1, C_2, D, and W depend solely on the rarefaction parameter R and the rate parameter Γ_{mn}. The dependence of the kinetic coefficients $G_1^{(1)}$ and $G_2^{(2)}$ on the molecular masses of the components of the gas mixture is totally determined by the factor Φ.

For the buffer gas we have

$$J_2=\frac{n_1R_0\Gamma_m\kappa}{2\sqrt{\pi}}\left(G_2^{(1)}\Delta\varepsilon+G_2^{(2)}\frac{\Delta d}{d_{n2}}\right). \tag{26}$$

Here the kinetic coefficients $G_2^{(1)}$ and $G_2^{(2)}$, which characterize, respectively, the surface and bulk mechanisms of the buffer-gas flow, depend on the rarefaction parameter R, the rate parameter Γ_{mn}, the mass ratio m_1/m_2 and effective-diameter ratio d_n/d_2 of the particles of the absorbing and buffer gases.

It is possible to obtain analytic expressions for the kinetic coefficients only for large or small Knudsen numbers. For atomic gases ($C\gg 1$) these expressions are written below.

1. An almost free-molecule regime (Kn$\gg 1$ or $R\ll 1$):

$$G_1^{(1)}=\frac{1}{C}(1-1.505R)+\cdots, \tag{27}$$

$$G_1^{(2)}=-1.505\frac{R}{C}+\cdots, \tag{28}$$

$$G_2^{(1)}=\Phi_1\left(1.505\frac{R}{C}+3.394\Phi_2\frac{R^2}{C}\ln R\right)+\ldots, \tag{29}$$

$$G_2^{(2)}=(1.505\Phi_1-0.113\Phi_3)\frac{R}{C}$$

$$+3.394\Phi_1\Phi_2\frac{R^2}{C}\ln R+\cdots, \tag{30}$$

where

$$\Phi_1=\sqrt{\frac{m_1}{m_2}}, \quad \Phi_2=\frac{\sqrt{1+(m_1/m_2)}}{(1+d_n/d_2)^2}, \quad \Phi_3=\frac{m_1}{m_1+m_2}.$$

2. A hydrodynamic regime with slipping (Kn$\ll 1$ or $R\gg 1$):

$$G_1^{(1)}=\frac{1}{\sqrt{\pi}\,R(R+C)}+\cdots, \tag{31}$$

$$G_1^{(2)}=-\frac{1}{R+C}+\frac{2+\Gamma_{mn}}{\sqrt{\pi}(R+C)^2}+\cdots, \tag{32}$$

$$G_2^{(1)}=\left[\Phi_1\left(\frac{1}{2}+3\sqrt{2}\right)+\Phi_5\right]\frac{1}{R+C}+\cdots, \tag{33}$$

$$G_2^{(2)}=\left\{\Phi_1\left(\frac{1}{2}+3\sqrt{2}\,\Phi_4\left[1+\frac{\Phi-2}{(\Phi-1)(1+\Gamma_{mn})}\right]\right)\right.$$

$$\left.+\Phi_5\left(1-\frac{1}{(\Phi-1)(1+\Gamma_{mn})}\right)\right\}\frac{1}{R+C}+\cdots, \tag{34}$$

where

$$\Phi_4=\frac{(1+m_1/m_2)^{3/2}}{(3+5m_1/m_2)(1+d_n/d_2)^2}, \quad \Phi_5=\frac{m_1/m_2}{3+5m_1/m_2}.$$

Analytic expressions for the kinetic coefficient in the case where $C\ll 1$ are given in Ref. 4.

4. DISCUSSION OF THE RESULTS AND COMPARISON WITH EXPERIMENTAL DATA

The directions of the surface components of the flows of the absorbing and buffer gases (J_1 and J_2) are determined by the signs of the difference of the accommodation coefficients of the unexcited and excited particles of the absorbing gas, $\Delta\varepsilon=\varepsilon_n-\varepsilon_m$, and the offset between the light frequency and the center of the absorption line, $\Omega=\omega-\omega_{mn}$. If we have $\Delta\varepsilon>0$, the directions of the surface components of the absorbing and buffer gases at $\Omega>0$ coincide with the direction of light propagation, but at $\Omega<0$ they are opposite to the direction of light propagation. The fact that the surface components of J_1 and J_2 have the same directions is consistent with momentum conservation in the gas–wall interaction.

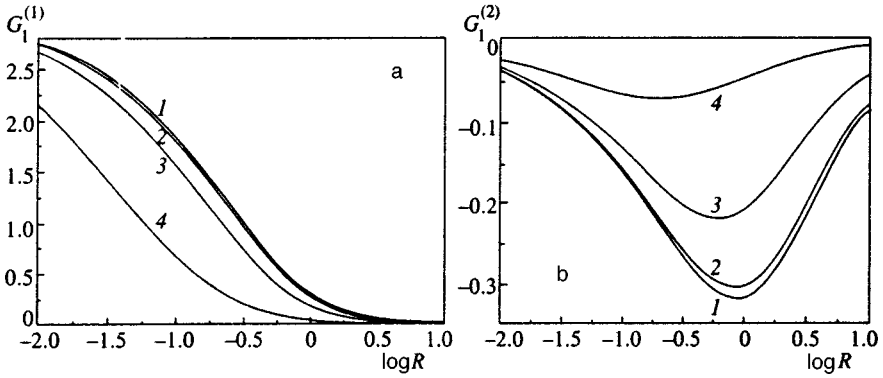


FIG. 2. Kinetic coefficients $G_1^{(1)}$ (a) and $G_1^{(2)}$ (b) as functions of the rarefaction parameter R ($m_1/m_2=1$): curves *1*, $\Gamma_{mn}=0.01$; curves *2*, $\Gamma_{mn}=0.1$; curves *3*, $\Gamma_{mn}=1$; and curves *4*, $\Gamma_{mn}=10$.

The directions of the bulk components of J_1 and J_2 are determined by the signs of the difference of the effective diameters of the excited and unexcited particles of the absorbing gas, Δd, and the offset Ω. If $\Delta d>0$ holds, the direction of the bulk component of the absorbing gas flow for $\Omega<0$ is the same as the direction of light propagation, while the direction of the bulk component of the buffer gas flow is opposite to the direction of light propagation, in accordance with momentum conservation in intermolecular collisions. For $\Omega>0$ the two gas flows reverse directions.

Figures 2a and 2b depict the kinetic coefficients $G_1^{(1)}$ and $G_1^{(2)}$ as functions of the rarefaction parameter R for different values of the rate parameter Γ_{mn}. We see that as the free-molecule regime ($R\ll 1$) is replaced by the hydrodynamic regime ($R\gg 1$), the kinetic coefficient $G_1^{(1)}$, which characterizes surface light-induced drift, monotonically decreases from a fixed value (Eq. (27)) to zero (Eq. (31)).

Equations (31)–(34) imply that in the slip regime ($R\gg 1$) the surface light-induced drift is an effect of second order in the Knudsen number ($G_1^{(1)}\sim \mathrm{Kn}^2$) and hence cannot be described by the theory of a flat Knudsen layer. The curvature of the boundary surface must be taken into account.

For a fixed value of R, the value of $G_1^{(1)}$ decreases with increasing Γ_{mn}: as Γ_{mn}^{-1} in the free-molecule regime, and as $(\Gamma_{mn}+1)^{-1}$ in the hydrodynamic regime. The reason is the decrease in the relative number of excited particles colliding with the capillary wall. In the $\Gamma_{mn}\to\infty$ limit all the molecules have time to go over to their ground state before they reach the capillary wall, so that there is no surface light-induced drift.

The kinetic coefficient $G_1^{(2)}$, which characterizes the bulk light-induced drift, is a nonmonotonic function of the rarefaction parameter R. As R increases, $G_1^{(2)}$ grows in absolute value in the Knudsen regime (Eq. (28)), reaches its maximum value in the intermediate regime at $R\sim 1$, and decreases as R^{-1} (Eq. (32)) in the hydrodynamic regime. At a fixed value of R the absolute value of $G_1^{(2)}$ decreases with increasing Γ_{mn}: as Γ_{mn}^{-1} in the free-molecule regime, and as $(\Gamma_{mn}+1)^{-1}$ in the slip regime. The explanation lies in the fact that with R fixed the fraction of the excited particles colliding with buffer-gas particles diminishes with increasing Γ_{mn}. When $\Gamma_{mn}\gg 1$ holds, all the excited particles within one mean free path have time to go to their ground state, which means there is no bulk light-induced drift.

Numerical calculations have shown that for $\Gamma_{mn}\leqslant 1$ the kinetic coefficients $G_1^{(1)}$ and $G_1^{(2)}$ are weakly dependent on Γ_{mn}. Hence for $\Gamma_{mn}\leqslant 1$ we can use the results of Ref. 4.

Figures 3a and 3b depict the kinetic coefficients $G_i^{(k)}$ as functions of the rarefaction parameter for a mixture of sodium vapor and helium at $C=R\Gamma_{mn}\approx 60$, a value used in the experiment described in Ref. 5. The method used in that experiment (see Ref. 5) consisted in the following. Sodium vapor mixed with helium was placed in a narrow capillary and the effect of 50-mW laser light on the D_2-line of sodium was studied. The capillary length was 40 cm and the inner diameter was 1 mm, which agrees with the adopted theoretical model. The gas pressure in the capillary was varied from 20 Pa to 13 kPa, which increased the rarefaction parameter R from 0.4 to 240 and decreased the rate parameter Γ_{mn} from 150 to 0.25.

The components of the gas mixture began to flow when laser light was applied, and the light frequency was selected to maximize the drift velocity of the fluorescing sodium va-

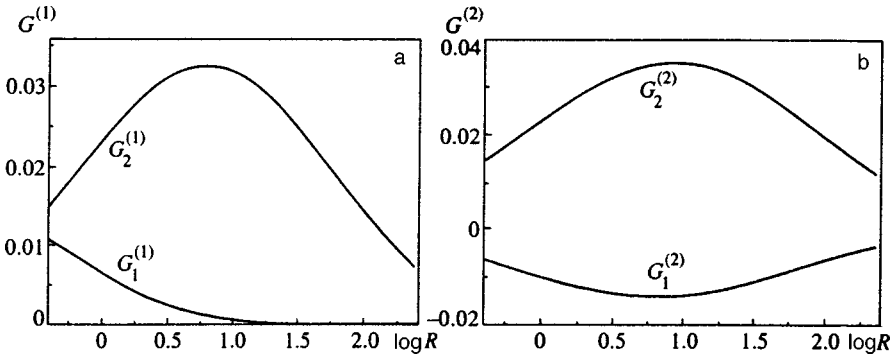


FIG. 3. Kinetic coefficients $G_1^{(1)}$ and $G_2^{(1)}$ (a), and $G_1^{(2)}$ and $G_2^{(2)}$ (b) as functions of the rarefaction parameter R for the Na–He gas mixture.

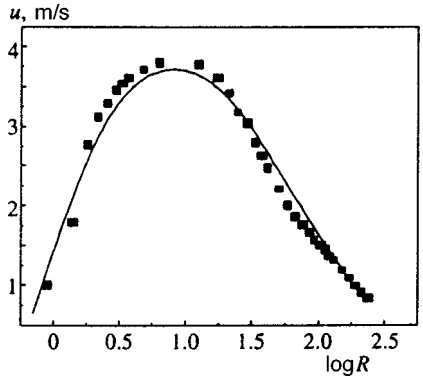


FIG. 4. Drift velocity of Na vapor in He as a function of the rarefaction parameter R; comparison of theoretical results (solid curve) and the experimental data of Ref. 5.

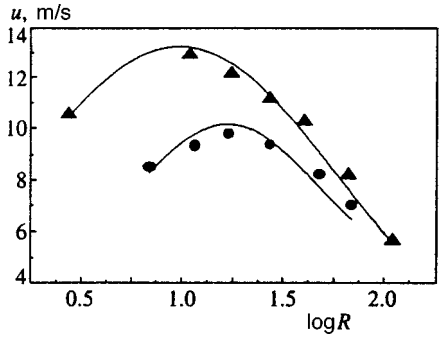


FIG. 5. Comparison of theoretical results (solid curves) with the experimental data of Ref. 12: ▲—the Na–Xe gas mixture; ●—the Na–Ar gas mixture.

por. The variation in sodium vapor density along the capillary was measured with a differential photodetector. Then the shift in sodium vapor density was transformed into the velocity u_{expt} of light-induced drift via the theoretical relationships derived in Ref. 5.

What was found was that u_{expt} depends on pressure nonmonotonically: as the pressure grows the value of u_{expt} increases from zero to its maximum value, reaches a plateau and remains constant in the 110–670 Parange, and then decreases to zero. Probably, the plateau develops because in the theoretical model of Ref. 5 the diffuse flow of the absorbing gas is calculated by Fick's law, $J_d=-D\nabla n$ (D is the diffusion coefficient), valid only in the hydrodynamic regime ($R>0$). But the plateau region corresponds to intermediate values of the Knudsen number. In Ref. 11 Fick's law was generalized to the entire range of Knudsen numbers: $J_d=-SD\nabla n$.

Calculations for the Na–He mixture have shown that the correction factor S increases from 0.4 at $R\sim 0.4$ to 1 at $R\gg 1$. The R-dependence of experimental values of the velocity of light-induced drift corrected for the factor S is shown in Fig. 4. The same figure also depicts the theoretical curve corresponding to the derived formula

$$u=\frac{R_0\Gamma_m\kappa}{2\sqrt{\pi}}\left(G_1^{(1)}\Delta\varepsilon+G_1^{(2)}\frac{\Delta d}{d_{n2}}\right). \tag{35}$$

The experimental values of $\Delta\varepsilon$ and $\Delta d/d_{n2}$ reconstituted from the condition of best agreement between the theory and the experiment of Ref. 5 for the Na–He mixture are

$$\Delta\varepsilon=-3.07\times 10^{-3},\quad \frac{\Delta d}{d_{n2}}=-4.05\times 10^{-3}.$$

The minus in the numerical values of $\Delta\varepsilon$ and $\Delta d/d_{n2}$ results from the drift of sodium vapor is opposite to the direction of light propagation. The theoretical curve describes fairly accurately the behavior of the experimental data of Ref. 5 in the entire range of gas pressures.

The discrepancy between the theoretical results and experimental data can be explained by the fact that $\kappa(\mathbf{v})\sim 1$ and that the light intensity is nonuniform for the cross section of the capillary (the diameter of the light beam amounted to half the inner diameter of the capillary). Possibly, allowing for these factors would change the resulting values of $\Delta\varepsilon$ and $\Delta d/d_{n2}$.

The results of experiments in light-induced drift of sodium vapor in the inert gases Xe and Ar can be found in Ref. 12. The researchers used the same experimental method as in Ref. 5. Figure 5 shows the results of the theory compared with the experimental data of Ref. 12 on the sodium-vapor drift velocity. Below we list the values of the parameters obtained from the condition of best agreement between theory and experiment.

(1) for the Na–Ar mixture,

$$\Delta\varepsilon=-7.33\times 10^{-2},\quad \frac{\Delta d}{d_{n2}}=-1.36\times 10^{-2};$$

(2) for the Na–Xe mixture,

$$\Delta\varepsilon=-1.79\times 10^{-2},\quad \frac{\Delta d}{d_{n2}}=-1.48\times 10^{-2}.$$

We see that the theory provides a satisfactory description of the experiment at all pressures in the capillary.

APPENDIX

The rates $\nu_{j2}^{(k)}$ obey the following expressions:

$$\nu_{j2}^{(1)}=\frac{16}{3}\frac{m_{12}}{m_1}n_2\Omega_{j2}^{(1,1)},$$

$$\nu_{j2}^{(3)}=\frac{16}{5}\frac{m_{12}^2}{m_1m_2}n_2\left(\frac{10}{3}\Omega_{j2}^{(1,1)}+\frac{m_2}{m_1}\Omega_{j2}^{(2,2)}\right),$$

$$\nu_{j2}^{(4)}=\frac{16}{5}\frac{m_{12}^2}{m_1m_2}n_2\left(\frac{10}{3}\Omega_{j2}^{(1,1)}-\Omega_{j2}^{(2,2)}\right),\quad j=m,n,$$

where $m_{12}=m_1m_2/(m_1+m_2)$ is the reduced mass of the colliding molecules, and $\Omega_{j2}^{(l,r)}$ are the Chapman–Cowling integrals, which for the hard-sphere model have the form

$$\Omega_{j2}^{(l,r)}=\left(\frac{k_BT}{2\pi m_{12}}\right)^{1/2}\frac{(r+1)!}{2}\left[1-\frac{1+(-1)^l}{2(l+1)}\right]\pi d_{j2}^2.$$

*)E-mail: Vladimir.Chernyak@usu.ru

[1] F. Kh. Gel'mukhanov and A. M. Shalagin, JETP Lett. **29**, 711 (1979).
[2] A. M. Dykhne and A. N. Starostin, Zh. Éksp. Teor. Fiz. **79**, 1211 (1980) [Sov. Phys. JETP **52**, 612 (1980)].
[3] M. A. Vaksman and A. V. Gaĭner, Zh. Éksp. Teor. Fiz. **89**, 41 (1985) [Sov. Phys. JETP **62**, 23 (1985)].
[4] V. G. Chernyak and E. A. Vilisova, Zh. Éksp. Teor. Fiz. **107**, 125 (1995) [JETP **80**, 65 (1995)].
[5] S. N. Atutov, I. M. Ermolaev, and A. M. Shalagin, Zh. Éksp. Teor. Fiz. **92**, 1215 (1987) [Sov. Phys. JETP **65**, 679 (1987)].
[6] S. G. Rautian, G. I. Smirnov, and A. M. Shalagin, *Nonlinear Resonances in the Spectra of Atoms and Molecules* [in Russian], Nauka, Novosibirsk (1979).
[7] F. J. McCormack, Phys. Fluids **16**, 2095 (1973).
[8] V. G. Chernyak, E. A. Vintovkina, and I. V. Chermyaninov, Zh. Éksp. Teor. Fiz. **103**, 1571 (1993) [JETP **76**, 768 (1993)].
[9] V. R. Mironenko and A. M. Shalagin, Izv. Akad. Nauk SSSR, Ser. Fiz. **45**, 995 (1981).
[10] S. G. Mikhlin, *Variational Methods in Mathematical Physics* [in Russian], Nauka, Moscow (1970).
[11] V. G. Chernyak and E. A. Vilisova, Zh. Éksp. Teor. Fiz. **110**, 150 (1996) [JETP **83**, 80 (1996)].
[12] S. N. Atutov, St. Lesyak, S. P. Pod'yachev, and A. M. Shalagin, Preprint No. 321, Institute of Automation and Electrometry, Siberian Branch of the USSR Academy of Sciences, Novosibirsk (1986).

Translated by Eugene Yankovsky

Dispersion of the speed of sound and absorption in the vicinity of the liquid–gas critical point: renormalization-group calculation in the two-loop approximation

L. Ts. Adzhemyan,*) A. N. Vasil'ev,†) and A. V. Serdyukov‡)

St. Petersburg State University, 199034 St. Petersburg, Russia
(Submitted 23 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1723–1741 (November 1998)

The renormalization group method to second order in the ε expansion is used to calculate the singular parts of the absorption and dispersion of the speed of sound on the critical isochor above T_c. We express the investigated quantities in terms of the response function to temperature variations in the *H* model of Halperin, Hohenberg, and Siggia. Results are compared with the experimental data. © *1998 American Institute of Physics.* [S1063-7761(98)01211-6]

1. INTRODUCTION

The liquid–gas critical point belongs to the class of phase transitions which conserve the order parameter. Dynamic critical effects are manifested for such systems through the anomalous behavior of the hydrodynamic modes. The general physical picture is well known. The kinetics of the order parameter are determined by the diffusion mode, whose relaxation time grows without limit as the critical point is approached (critical slowing down). For the dynamics of this mode only its interaction with the transverse components of the velocity fluctuation field is important; the remaining hydrodynamic modes have no appreciable effect on the kinetics of the order parameter, but are themselves substantially affected by the critical fluctuations. This leads, in particular, to a vigorous growth of the attenuation of sound, and also to the appearance of low-frequency dispersion of the speed of sound and the attenuation factor.

A consistent statistical theory of dynamic critical phenomena in liquids can be based on the equations of nonlinear stochastic hydrodynamics.[1,2] Dropping the unimportant modes from these equations leads to the so-called *H* model, which describes the kinetics of the order parameter.[1,3] The renormalization-group (RG) method in conjunction with the ε expansion has proven to be an effective method for studying this model, as is has been in studying phase-transition thermodynamics. The dynamic similarity hypothesis and the calculation of the corresponding critical index were both based on this approach.[3]

Use of the RG method to describe sound waves in the critical region requires that we identify the main mechanism of interaction of sound waves with fluctuations of the order parameter in the nonlinear equations of hydrodynamics. Under conditions in which the sound wavelength λ_s substantially exceeds the correlation radius of the order parameter r_c (such conditions are always satisfied for experimentally attainable values of r_c), this interaction is realized via adiabatic temperature oscillations in the sound wave, which influence the most sensitive parameter of the statistical distribution of fluctuations of the order parameter—the deviation of the temperature from critical. This mechanism has long been known, and is reflected in a number of phenomenological theories. A basis for this mechanism within the framework of the statistical approach, itself based on a consideration of the complete set of equations of stochastic hydrodynamics, was worked out in Refs. 4 and 5 (an analogous problem was considered in Ref. 6 in application to the dynamics of liquid crystals).

According to Refs. 4 and 5, the dispersion relation for the acoustic mode on the critical isochor has the form

$$k^2=\omega^2[c_\infty^{-2}+AR(\omega,\tau)], \quad \tau=(T-T_c)/T_c, \tag{1}$$

where ω is frequency and τ is the reduced temperature, c_∞ and A are frequency- and temperature-independent constants, and T_c is the critical temperature. The function $R(\omega,\tau)$ is given by the relations

$$R(\omega,\tau)=D^{st}(\tau)+i\omega D^{ret}(\omega,\tau),$$

$$D^{st}\equiv D(t,t)=C, \quad D^{ret}(t,t')=\theta(t-t')D(t,t'), \tag{2}$$

where

$$D(t,t')=\int d\mathbf{x}'\langle F(\mathbf{x},t)F(\mathbf{x}',t')\rangle, \tag{3}$$

$$F(\mathbf{x},t)=[\psi^2(\mathbf{x},t)-\langle\psi^2(\mathbf{x},t)\rangle]/2, \tag{4}$$

and ψ is the order parameter field. The averaging in Eq. (3) is within the framework of the *H* model, i.e., with account of sound waves. Functions of the type (3) depend only on the difference $t-t'$, and in transforming to the ω representation, Fourier transforms are always taken with respect to $t-t'$.

The quantity $C\equiv D(t,t)$ in Eqs. (2) is the singular part of the specific heat at constant volume, and the expression in brackets in Eq. (1) can be interpreted (to within a multiplicative factor) as the dynamic generalization of the specific heat. This quantity plays a fundamental role in the phenomenological theory of sound propagation in critical systems proposed in Refs. 7 and 8 and developed by a number of other authors (see, e.g., Refs. 9 and 10). Relations (1)–(4) can be considered a statistical generalization of the concept of the dynamic specific heat.

For a given frequency ω, we seek a solution of the dispersion relation (1) in the form $k=\omega c^{-1}+i\alpha$, hence we seek the positive parameters $c(\omega)$ (the speed of sound) and $\alpha(\omega)$

(the attenuation factor). Instead of α, one frequently finds the dimensionless quantity $\alpha_\lambda \equiv \alpha\lambda_s = 2\pi c\alpha/\omega$, the attenuation per unit wavelength. Usually $\alpha_\lambda \ll 1$, so it follows from Eq. (1) that

$$c^{-2}(\omega) = c_\infty^{-2} + A\,\mathrm{Re}\,R(\omega),$$
$$\alpha_\lambda(\omega)c^{-2}(\omega) = \pi A\,\mathrm{Im}\,R(\omega). \tag{5}$$

The quantity $R(\omega)$ tends to zero as $\omega\to\infty$; therefore in relations (5)

$$R(\infty) = 0, \quad c(\infty) = c_\infty, \quad \alpha_\lambda(\infty) = 0. \tag{6}$$

Thus, the τ- and ω-independent constant c_∞ is in essence the limiting speed of sound as $\omega\to\infty$.

According to the dynamic scaling hypothesis the asymptotic limit of the function $R(\omega,\tau)$ in the critical region has the following form:

$$R(\omega,\tau) = C_1\tau^{-\alpha}\Phi(C_2\bar{\omega}), \quad \bar{\omega} \equiv \omega\tau^{-z\nu}, \tag{7}$$

where $C_{1,2}$ are nonuniversal constants that depend on the material, $\bar{\omega}$ is the reduced frequency (a multiple of the ratio of the speed of sound ω to the characteristic fluctuation frequency $\omega_{fl}\sim\tau^{z\nu}$), and $\Phi(\bar{\omega})$ is the universal scaling function. The well-known static indices in the expression on the right-hand side of Eq. (7), ν and $\alpha = 2 - d\nu$ (d is the dimensionality of the space), have been studied in detail within the framework of the renormalization group, all the way out to the five-loop approximation. The Borel sum of their ε expansions coincides with the experimental values $\alpha\simeq 0.11$, $\nu\simeq 0.63$ (Ref. 11). The accuracy to which the dynamic characteristics in Eq. (7) have so far been calculated is significantly lower. The dynamic index z (critical dimension of the frequency) in the H model was calculated in the two-loop approximation in Ref. 3. Its experimental value is $z\simeq 3.07$ (Ref. 12).

The experimental data (see, e.g., Refs. 13–15) confirm the validity of the scaling law (7) for the quantities (5) in the critical region, and enable one to assess the behavior of the universal function $\Phi(\bar{\omega})$ over a wide range of frequencies $\bar{\omega}$. On the theoretical plane, the scaling function Φ was calculated in the one-loop approximation in Ref. 16 (however, the derivation of the full dispersion relation in Ref. 16 is incorrect), and an effort was made in Ref. 17 to extend the one-loop approximation to a wider neighborhood of the critical point by taking corrections to scaling into account.

In the present paper, we present a detailed derivation of representation (7) within the framework of the RG technique, together with a calculation in the two-loop approximation of its constituent scaling function Φ, and compare our results with the experimental data. A brief preliminary report on this work was given at the conference ''Renormalization Group-96'' (Dubna, 1996), the proceedings of which were published in the International Journal of Modern Physics B.

2. RENORMALIZATION OF THE R FUNCTION IN THE H MODEL

The H model can be renormalized in the standard way in the language of field theory.[18] Thus, the correlation functions (Green's functions) can be defined as the functional averages of the fields of the H model with weight $\exp(S)$, and the unrenormalized action functional has the form

$$S = \int d\mathbf{x}\int dt\{-\lambda_0\psi'\nabla^2\psi' + \psi'[-\partial_t\psi - (\mathbf{v}\nabla)\psi + \lambda_0\nabla^2(-\nabla^2\psi + \tau_0\psi + g_{10}\psi^3/6)] - \lambda_0^{-1}g_{20}^{-1}\mathbf{v}'\nabla^2\mathbf{v}' + \mathbf{v}'[-a\partial_t\mathbf{v} + \lambda_0^{-1}g_{20}^{-1}\nabla^2\mathbf{v} + \psi\nabla\nabla^2\psi]\}, \tag{8}$$

where ψ is the order parameter field, $\mathbf{v}$ is the transverse velocity field, and ψ' and $\mathbf{v}'$ are auxiliary fields. The quantities λ_0 and $\nu_0 = (a\lambda_0 g_{20})^{-1}$ are the unrenormalized thermal conductivity and kinematic viscosity, g_{20} is the unrenormalized charge of the intermode interaction, g_{10} is the static charge, $\tau_0 \propto T - T_c$ is the deviation of the temperature from critical, and $a = \rho_c/k_B T_c$. The propagators of the model (8) in the momentum–time $(\mathbf{k},t)$ representation have the form

$$\langle\psi\psi\rangle = (k^2+\tau_0)^{-1}\exp\{-\varepsilon_k|t-t'|\}, \quad \langle\psi'\psi'\rangle = 0,$$
$$\langle\psi\psi'\rangle = \theta(t-t')\exp\{-\varepsilon_k(t-t')\},$$
$$\langle v_i v_j\rangle = P_{ij}(\mathbf{k})\nu_0\lambda_0 g_{20}\exp\{-\nu_0 k^2|t-t'|\}, \quad \langle v_i' v_j'\rangle = 0,$$
$$\langle v_i v_j'\rangle = P_{ij}(\mathbf{k})\nu_0\lambda_0 g_{20}\theta(t-t')\exp\{-\nu_0 k^2(t-t')\}, \tag{9}$$

where $\varepsilon_k \equiv \lambda_0 k^2(k^2+\tau_0)$ and $P_{ij}(\mathbf{k}) = \delta_{ij} - k_i k_j/k^2$ is the transverse projector.

The H model (8) is logarithmic in the dimensionality $d=4$ (Refs. 3 and 18), and we will consider it in the dimensional regularization $d = 4-\varepsilon$ using the minimal subtraction scheme in the renormalization. Ultraviolet (UV) divergences in the dimensional regularization show up as poles in ε, and renormalization in the minimal subtraction scheme reduces to removing these poles without altering any of the other (non-pole) contributions.

In studying the critical behavior of the Green's functions in the H model, we are interested in their infrared (IR) asymptotic limit $\tau_0\to 0$ in the regime $\omega\sim k^4\sim\tau_0^2$. It is well known (see, e.g., Refs. 3 and 18) that in such a regime the contribution $a\mathbf{v}'\partial_t\mathbf{v}$ in Eq. (8) is IR-negligible, and can be discarded. For brevity, we call the H model in the limit $a\to 0$ the H_0 model. Perturbation-theory diagrams with propagators (9) admit the limit $a\to 0$ (i.e., $\nu_0\to\infty$). Using the inequality

$$\lim_{\nu_0\to\infty}\nu_0\exp(-\nu_0 k^2 t) = 2\delta(t)$$

one could take the limit in the propagators themselves, but certain diagrams in this case would require redefinition, due to the presence of θ functions in the propagators $\langle\mathbf{v}\mathbf{v}'\rangle$ and $\langle\psi\psi'\rangle$. Therefore, to avoid confusion in the calculation of the H_0 model diagrams, they must be understood to be H model diagrams in which one subsequently takes the limit $a\to 0$.

Following the authors of Refs. 3 and 18 and all subsequent work along these lines, we assume that the renormalization of the H_0 model is multiplicative and agrees with the statics of the problem, i.e., the dynamic and static renormalization constants for the objects figuring in both the dynamics and statics of the problem in the minimal subtraction

scheme coincide. (Strictly speaking, none of the published work known to us contains a formal proof of this statement. But it is plausible, and all calculations carried out in the H model are consistent with it.)

Thus, we assume that the renormalized action of the H_0 model can be derived from the unrenormalized action (8) with $a=0$ by the standard multiplicative renormalization of the fields and unrenormalized parameters $e_0=\{\tau_0,\lambda_0,g_0\}$:

$$S_R(\phi,e,\mu)=S(Z_\phi\phi,e_0), \quad \phi\equiv\{\psi,\psi',\mathbf{v},\mathbf{v}'\},$$
$$\tau_0=\tau Z_\tau, \quad \lambda_0=\lambda Z_\lambda, \quad g_{i0}=g_i\mu^\varepsilon Z_{g_i}, \quad i=1,2, \quad (10)$$

where $e\equiv\tau,\lambda,g$ are the renormalized parameters, and μ, the renormalized mass, is an additional parameter of the renormalized theory. Agreement with the statics means that in the minimal subtraction scheme the dynamic renormalization constants of the field ψ and parameters τ, g_1 are the same as in the simple static ψ^4 model, and $Z_{\mathbf{v}}=1$, since the field $\mathbf{v}$ enters into the static action corresponding to expression (8) in the form of an additive term $\sim\mathbf{v}^2$ and is not renormalized. In the minimal subtraction scheme all renormalization constants have the form

$$Z=1+Z^{(1)}(g)/\varepsilon+Z^{(2)}(g)/\varepsilon^2+\ \ldots\ , \quad (11)$$

and the expansion of $Z^{(n)}(g)$ begins with contributions of order g^n.

Let us turn now to the function $R(\omega)$ defined by relations (2) and (3). Invoking the fluctuation–dissipation theorem[19] for the component operators in the H model, it can be shown that $R(\omega)$ coincides with the Fourier transform of the correlator

$$R(t,t')=\int d\mathbf{x}\langle F(\mathbf{x},t)F'(\mathbf{x}',t')\rangle, \quad (12)$$

where the component operator F is defined in Eq. (4) and the second operator

$$F'(t,\mathbf{x})=-\lambda_0\psi'\Delta\psi. \quad (13)$$

Representation (12) reveals the physical meaning of $R(t,t')$ to be the response function of the quantity $\langle F(\mathbf{x},t)\rangle$ to the strictly time-dependent temperature variation $\delta\tau(t')$ in the model (8), and the first of equalities (2) with R from Eq. (12) is a consequence of the fluctuation–dissipation theorem. In the calculation of the function R in the two-loop approximation, we start out with representation (12), calculating within the framework of the H_0 model, since we are interested in IR-negligible corrections.

Let us discuss briefly the renormalization of the response function and its constituent operators. Operator (4) in the static ψ^4 model is renormalized multiplicatively (by subtracting out the quantity $\langle\psi^2\rangle$ in Eq. (4)), i.e., $F=Z_FF_R$ with the constant $Z_F=Z_\tau^{-1}$ (Ref. 11). From general principles of correspondence between the dynamics and statics it follows that the operator F can be renormalized in exactly the same way in the dynamic H_0 model, and with the same constant $Z_F=Z_\tau^{-1}$ in the minimal subtraction scheme. Operator (13) has meaning only in a dynamic context. By analyzing the diagrams of the H_0 model with insertion of the single operator F', it is not hard to show that it too can be renormalized multiplicatively (it is not mixed with anything), i.e., $F'=Z_{F'}F'_R$. From relation (2) for the response function (12) with $t\neq t'$ rewritten in the t representation, it follows that $Z_{F'}=Z_F$. Thus renormalization of the component operators F and F' is given by the relations

$$F=Z_FF_R, \quad F'=Z_{F'}F'_R, \quad Z_F=Z_{F'}=Z_\tau^{-1}, \quad (14)$$

where Z_τ is the well-known renormalization constant of the parameter τ in the static ψ^4 model. Note that the constant $Z_{F'}$ for operator (13) is the product of Z_λ from Eqs. (10) and the renormalization constant of the monomial $\psi'\Delta\psi$.

In general, renormalization of correlators of the types (3) and (12) with two component operators requires, besides the substitution $F\to F_R$ of the operators themselves, the introduction of an additive, local (multiple of $\delta(t-t')$ in the t representation) ''counterterm per operator pair.'' It is not hard to show from dimensional arguments that for the correlator (3) in the dynamics such a counterterm is not required; therefore $D_R(\omega)=Z_F^{-2}D(\omega)$ and similarly for $D^{ret}(\omega)$ in (3). But for the static object $D^{st}=C$ a ''counterterm per pair'' is needed.[11] It follows from (2) and (14) that a counterterm per pair is also needed in that situation for the response function $R(\omega)$, and that it is exactly the same one as for the specific heat C. The foregoing can be summarized in the following way:

$$C_R=Z_F^{-2}C+\Delta C, \quad R_R=R'+\Delta C,$$
$$R'\equiv Z_F^{-2}R, \quad D_R^{ret}=Z_F^{-2}D^{ret}. \quad (15)$$

Here $Z_F=Z_\tau^{-1}$ according to (14), and ΔC, the ''counterterm per pair'' in the specific heat, is an ω-and τ-independent constant containing poles in ε:

$$\Delta C=\mu^\varepsilon b(g_1), \quad b(g_1)=b^{(1)}(g_1)/\varepsilon+b^{(2)}(g_1)/\varepsilon^2+\ \ldots\ . \quad (16)$$

The function $b(g_1)$ is known from an analysis of the renormalization of the specific heat in the static ψ^4 model.[11]

In an approximation linear in the charges g, which is sufficient for the development that is to follow, the quantities $Z^{(1)}$ in Eq. (11) for the independent renormalization constants and $b^{(1)}$ in Eq. (16) are given by[11,18]

$$Z_\phi^{(1)}=0\,\forall\,\phi, \quad Z_{g_1}^{(1)}=\frac{3u_1}{2}, \quad Z_{g_2}^{(2)}=\frac{19u_2}{24},$$
$$Z_\tau^{(1)}=-Z_F^{(1)}=\frac{u_1}{2}, \quad Z_\lambda^{(1)}=-\frac{3u_2}{4},$$
$$b^{(1)}=-\frac{1}{16\pi^2}+O(g_1^2), \quad (17)$$

in which $u_i\equiv g_i/8\pi^2$ for both charges $g_{1,2}$.

3. RENORMALIZATION-GROUP EQUATIONS: RG REPRESENTATION OF THE RESPONSE FUNCTION

In this section, on the basis of the renormalization-group equations we obtain the RG representation of the response function, which makes it possible to justify the scaling law (7) in the IR asymptotic limit and obtain working formulas with which to calculate the scaling function Φ Eq. (7).

The RG equation for the renormalized response function R_R is somewhat peculiar, due to the additive counterterm ΔC in its renormalization (15). In what follows it will be convenient to use the RG equations for R_R and the ''incompletely renormalized'' function R' in Eq. (15). These equations can be obtained by requiring that the unrenormalized function R be independent of the renormalized mass μ (this parameter is absent from the unrenormalized theory and appears only in the renormalization formulas (10), and through them in the renormalized Green's functions). Acting on both sides of the second and third of Eqs. (15) with the operator $\tilde{\mathscr{D}}_\mu$ (here and below $\mathscr{D}_a \equiv a\partial_a$, and $\tilde{\mathscr{D}}_\mu$ is the operation $\mu\partial_\mu$ for fixed values of the unrenormalized parameters), we obtain the desired RG equations:

$$(\mathscr{D}_{RG}+2\gamma_F)R_R=\mu^{-\varepsilon}\gamma_0, \quad (\mathscr{D}_{RG}+2\gamma_F)R'=0. \tag{18}$$

Here

$$\mathscr{D}_{RG}=\mathscr{D}_\mu+\beta\partial_g-\gamma_\tau\mathscr{D}_\tau-\gamma_\lambda\mathscr{D}_\lambda \tag{19}$$

is the operator $\tilde{\mathscr{D}}_\mu$ in renormalized variables with summation over the two charges $g=g_{1,2}$ in the term $\beta\partial_g \equiv \Sigma_i \beta_i \partial_{g_i}$. The relations (18) and (19) contain the RG functions

$$\gamma_0 \equiv \mu^\varepsilon Z_F^{-2}\tilde{\mathscr{D}}_\mu(Z_F^2\Delta C), \quad \gamma_a \equiv \tilde{\mathscr{D}}_\mu \ln Z_a \quad \forall a\neq 0,$$

$$\beta_i \equiv \tilde{\mathscr{D}}_\mu g_i = -g_i(\varepsilon+\gamma_{g_i}). \tag{20}$$

The RG functions γ_a with $a=g,\tau,\lambda,F$ are the anomalous dimensionalities of the quantities a, γ_0 is the ''vacuum RG function,'' β_i are the beta functions of the charges g_i, and all of them are UV-finite, i.e., they have no poles in ε. In the minimal subtraction scheme all of the functions γ are in general independent of ε and can be calculated with the help of the relations $\gamma_0=-(1+\mathscr{D}_g)b^{(1)}$ and $\gamma_a=-\mathscr{D}_g Z_a^{(1)}$, in which $\mathscr{D}_g=\mathscr{D}_{g_1}+\mathscr{D}_{g_2}$, and $Z_a^{(1)}$ and $b^{(1)}$ are the coefficients of $1/\varepsilon$ in Eqs. (11) and (16). Expressions for these quantities with the accuracy needed for further analysis are given in Eqs. (17). It follows from Eqs. (17) and (20) that

$$\beta_1=g_1\left(-\varepsilon+\frac{3u_1}{2}\right), \quad \beta_2=g_2\left(-\varepsilon+\frac{19u_2}{24}\right), \tag{21}$$

$$\gamma_\tau=-\gamma_F=-\frac{u_1}{2}, \quad \gamma_\lambda=\frac{3u_2}{4}, \quad \gamma_0=\frac{1}{16\pi^2}+O(g_1^2) \tag{22}$$

with $u_i \equiv g_i/8\pi^2$ for the charges $g_{1,2}$. The equality $\gamma_\tau=-\gamma_F$ is exact; it follows from the definition of γ_a in Eq. (20) and relation (14) between the renormalization constants Z_F and Z_τ. The RG functions β_1, γ_τ, and γ_0 depend only on the static charge g_1, while the functions β_2 and γ_λ depend on both charges (their lack of a dependence on g_1 in Eqs. (21) and (22) is a consequence of the approximation).

The beta functions (21) have an IR-stable fixed point with coordinates

$$u_{1*}=2\varepsilon/3+O(\varepsilon^2), \quad u_{2*}=24\varepsilon/19+O(\varepsilon^2), \tag{23}$$

in which $\beta_i(g_*)=0$, and the eigenvalues of the matrix $\omega_{ik} \equiv \partial\beta_i/\partial g_k|_{g=g_*}$ are positive (the latter implies IR-stability). The existence of an IR-stable fixed point is a necessary condition for critical scaling.

We turn now to the RG equation (18) for R' and make the substitutions

$$R'(\tau,\omega,\mu,\lambda,g)=\mu^\varepsilon Q(\tilde{\tau},w,g), \quad w\equiv\omega/2\lambda\mu^4, \quad \tilde{\tau}\equiv\tau/\mu^2, \tag{24}$$

where Q is a dimensionless function of the dimensionless arguments w and $\tilde{\tau}$. Equation (18) for R', allowing for the equality $\gamma_F=-\gamma_\tau$, can be rewritten in the form ($\mathscr{D}_a \equiv a\partial_a$)

$$[\beta\partial_g-(2-\gamma_\tau)\mathscr{D}_{\tilde{\tau}}-(4-\gamma_\lambda)\mathscr{D}_w-(\varepsilon+\gamma_\tau)]Q=0. \tag{25}$$

The solution of Eq. (25) has the form

$$Q(\tilde{\tau},w,g)=Q(1,\bar{w},\bar{g})\exp\left\{-\int_1^{\tilde{\tau}}\frac{d\tau_1}{\tau_1}\,\frac{\varepsilon+2\gamma_\tau(\bar{g}_1(\tau_1))}{2+\gamma_\tau(\bar{g}_1(\tau_1))}\right\}, \tag{26}$$

where $\bar{g}$ and $\bar{w}$ are invariants corresponding to g and w, given by

$$\mathscr{D}_{\tilde{\tau}}\bar{g}_i=\beta_i(\bar{g})/[2+\gamma_\tau(\bar{g})], \quad \bar{g}_i|_{\tilde{\tau}=1}=g_i, \tag{27}$$

$$\mathscr{D}_{\tilde{\tau}}\bar{w}=-\bar{w}[4-\gamma_\lambda(\bar{g})]/[2+\gamma_\tau(\bar{g})], \quad \bar{w}|_{\tilde{\tau}=1}=w. \tag{28}$$

Since the fixed point (23) is IR-stable, we have for the invariant charges defined by relations (27) the IR-asymptotic limit $\bar{g}_i \to g_{i*}$ as $\tilde{\tau}\to 0$. Therefore, the IR-asymptotic limit of the invariant $\bar{w}$ defined by Eq. (28) can be obtained by setting $\bar{g}=g_*$ in this equation. This yields

$$\bar{w}_{IR}=Cw\tilde{\tau}^{-z\nu}, \tag{29}$$

where C is a critical-dimensionless nonuniversal factor, and

$$1/\nu \equiv \Delta_\tau \equiv 2+\gamma_\tau^*, \quad \text{and} \quad z\equiv\Delta_\omega\equiv 4-\gamma_\lambda^* \tag{30}$$

are the critical dimensionalities of the variables τ and ω ($1/\nu$ and z are the standard notation for them, $\gamma^* \equiv \gamma(g_*)$). The asymptotic limit of the exponential in Eq. (26) as $\tilde{\tau}\to 0$ has the form $C'\tilde{\tau}^{-\alpha}$, where $\alpha \equiv (\varepsilon+2\gamma_\tau^*)/(2+\gamma_\tau^*)=2-d\nu$ is the critical index of the specific heat. Substituting the resulting asymptotic limits into Eq. (26) yields the IR-asymptotic limit of the function Q:

$$Q_{IR}(\tilde{\tau},w,g)=C'\tilde{\tau}^{-\alpha}Q(1,Cw\tilde{\tau}^{-z\nu},g_*). \tag{31}$$

This proves the validity of representation (7) for the IR-asymptotic limit of the unrenormalized function R, since it differs from Q only by an ω- and τ-independent factor: $R=\mu^{-\varepsilon}Z_F^{-2}Q$ according to Eqs. (15) and (24).

Since $g_*\sim\varepsilon$, relation (31) could also serve to obtain the ε expansion of the scaling function Φ in Eq. (7). This, however, is hindered by the remaining uneliminated poles in ε in the function $R'=\mu^{-\varepsilon}Q$. These poles are absent from the renormalized function $R_R=R'+\Delta C$, in terms of which the function $Q(1,w,g_*)$ on the right-hand side of Eq. (31) can be expressed in the form

$$Q(1,w,g_*)=\mu^\varepsilon R_R(\tau=\mu^2,\omega=2\lambda w\mu^4,g=g_*)-b_*, \tag{32}$$

with Eqs. (15), (24), and (31) taken into account, where $b_* \equiv b(g_*)$. This constant can be extracted from the right-

hand side of Eq. (32) as a common factor. This leads only to a change in the nonuniversal amplitude C' in Eq. (31).

The ε expansion of the function R_R in Eq. (32) starts off with contributions of order unity, while the quantities b_* start off with contributions of order $1/\varepsilon$. To calculate the coefficients of the ε expansion of b_* directly from Eq. (16), infinite sums are required (e.g., in the coefficient of $1/\varepsilon$ all terms of the type g_1^{n-1}/ε^n contribute for $g_1=g_{1*}\sim\varepsilon$). It is possible to avoid this difficulty, as the function $b(g_1)$ in the counterterm ΔC can be expressed in terms of the vacuum RG function γ_0 defined in Eq. (20), which is independent of ε. Indeed, it follows from Eqs. (18) that the counterterm $\Delta C=R_R-R'$ satisfies exactly the same RG equation as the function R_R in Eqs. (18). Hence, taking the explicit form of the counterterm $\Delta C=\mu^{-\varepsilon}b(g_1)$ and the operator $\mathscr{D}_{RG}$ (19) into account, we obtain

$$(\beta_1\partial_{g_1}-\varepsilon-2\gamma_\tau)b(g_1)=\gamma_0(g_1). \tag{33}$$

Setting $g_1=g_{1*}$ here, we find

$$b_*=-\gamma_0^*/(\varepsilon+2\gamma_\tau^*)=-\nu\gamma_0^*/\alpha, \tag{34}$$

where ν and α are the static indices defined above, and $\gamma_0^*\equiv\gamma_0(g_{1*})$. Taking Eqs. (27), (31), (32), and (34) into account, we obtain representation (7) for the unrenormalized function R in Eq. (17) with scaling function

$$\Phi(w)=1+\frac{\alpha\mu^\varepsilon}{\nu\gamma_0^*}R_R(\tau=\mu^2,\omega=2\lambda w\mu^4,g=g_*) \tag{35}$$

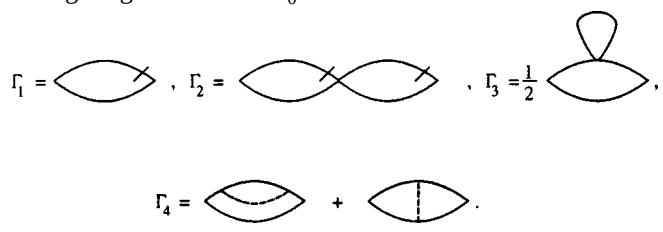
and some nonuniversal constants $C_{1,2}$. Equation (35) is fundamental to a calculation of the ε expansion of the scaling function.

4. CALCULATION OF THE SCALING FUNCTION

We want to calculate the scaling function (35) to accuracy ε^2, which corresponds to the two-loop approximation for R_R. The coefficient $\alpha/\nu\gamma_0^*$ in Eq. (35) is a quantity of order ε ($\alpha=\varepsilon/6-29\varepsilon^2/324+O(\varepsilon^3)$, $\nu=1/2+\varepsilon/12+O(\varepsilon^2)$ (Ref. 11), and γ_0^* can be found from Eqs. (22) to the required accuracy); hence

$$\frac{\alpha}{\nu}=\frac{\varepsilon}{3}\left(1-\frac{19}{27}\varepsilon\right)+O(\varepsilon^3),\quad \gamma_0^*=\frac{1}{16\pi^2}+O(\varepsilon^2). \tag{36}$$

The function R_R is given to two-loop accuracy by the following diagrams of the H_0 model:



(37)

The solid lines depict propagators of the fields ψ and ψ', the dashed lines depict propagators of the fields $\mathbf{v}$ and $\mathbf{v}'$. The slashes through three of the lines correspond to the auxiliary fields ψ' and $\mathbf{v}'$. The lines can be put into correspondence with the propagators of the basic theory obtained from Eq. (8) by replacing all of the unrenormalized parameters with renormalized ones: $\tau_0\to\tau$, $\lambda_0\to\lambda$, $g_{i0}\to g_i\mu^\varepsilon(i=1,2)$. The diagrams for Γ_3 and Γ_4 are a symbolic notation for the sum of diagrams with all possible arrangements of slashes at the ends of the lines (three diagrams for Γ_3, ten for Γ_4; a detailed breakdown is given in the Appendix).

The diagrams can be calculated for a given external frequency ω and zero external momentum. After integrating over times for the diagrams in Eqs. (37) in the H_0 model (the limit $\nu_0\sim a^{-1}\to\infty$ in Eqs. (9)) we obtain the following explicit expressions:

$$\Gamma_1=\int\mathscr{D}\mathbf{k}\ \frac{\lambda k^2}{\alpha_k(2\varepsilon_k-i\omega)},\quad \Gamma_2=-g_1\mu^\varepsilon\Gamma_1^2,$$

$$\Gamma_3=g_1\mathscr{T}(\tau)\partial_\tau\Gamma_1,\quad \mathscr{T}(\tau)=-\frac{1}{2}\tau\mu^\varepsilon\int\mathscr{D}\mathbf{q}\frac{1}{q^2\alpha_q},$$

$$\Gamma_4=-2\lambda g_2\mu^\varepsilon\int\mathscr{D}\mathbf{k}\int\mathscr{D}\mathbf{q}\times\frac{[k^2q^2-(\mathbf{kq})^2]i\omega(\varepsilon_k-\varepsilon_q)^2}{|\mathbf{k}+\mathbf{q}|^4\alpha_k(2\varepsilon_k-i\omega)^2\alpha_q(2\varepsilon_q-i\omega)^2}, \tag{38}$$

in which

$$\alpha_k\equiv k^2+\tau,\quad \varepsilon_k\equiv\lambda k^2(k^2+\tau),\quad \mathscr{D}\mathbf{k}\equiv d^d\mathbf{k}/(2\pi)^d. \tag{39}$$

Expressions (38) correspond to the ''basic theory,'' while the function R_R in Eq. (35) is the sum of the corresponding renormalized quantities. The one-loop diagram Γ_1 can be renormalized, according to (15), by adding the one-loop counterterm (which is known from Eqs. (16) and (17)) per operator pair, : $\Gamma_1^R=\Gamma_1+\Delta C_1$, where $\Delta C_1=-\mu^{-\varepsilon}/16\pi^2\varepsilon$. To renormalize the two-loop diagrams $\Gamma_{2,3,4}$, it is necessary to take into account, first, the contributions of the counterterms of all of the divergent one-loop subgraphs and, second, the two-loop counterterm per operator pair ΔC_2 for the diagram Γ_2 (for $\Gamma_{3,4}$ the counterterm per pair is not required since they contain no overall ''surface divergence,'' the dimensional factors τ in Γ_3 and ω in Γ_4 having been separated out). The ''counterterms per subgraph'' for the sum of all two-loop diagrams can be found by analyzing the corrections issuing from renormalization of all quantities in the one-loop graph Γ_1. To do so, it is necessary to make the substitutions $\tau\to\tau Z_\tau,\lambda\to\lambda Z_\lambda$ (to our level of accuracy, corrections due to field renormalization are not required, since in Eqs. (17) $Z_\phi^{(1)}=0\forall\ \phi$), multiply Γ_1 by $Z_F^{-2}=Z_\tau^2$ (which corresponds to taking account of renormalization of the component operators), and then extract all correction terms of first order in the charge from the resulting expression $Z_\tau^2\Gamma_1$. The sum of these corrections has the form

$$\varepsilon^{-1}[2Z_\tau^{(1)}\Gamma_1+Z_\tau^{(1)}\mathscr{D}_\tau\Gamma_1+Z_\lambda^{(1)}\mathscr{D}_\lambda\Gamma_1] \tag{40}$$

with the quantities $Z_a^{(1)}$ known from Eqs. (17).

The first term in expression (40) corresponds to one-loop renormalization of the component operators; the second, to one-loop renormalization of the parameter τ; and the third, to one-loop renormalization of the parameter λ. In our case,

these terms can be uniquely ''distributed'' among the diagrams $\Gamma_{2,3,4}$, thereby determining the renormalized value of each of them individually. Indeed, from the form of the diagrams it is clear that one-loop renormalization of the component operators is possible only in the graphs $\Gamma_{2,4}$, but is obviously lacking in Γ_3. The first term in expression (40), in which $Z_\tau^{(1)}$ depends only on g_1, is responsible for it.

Hence it is clear that in the graphs Γ_4, which depend only on the charge g_2, the one-loop subdivergences from the renormalization of the component operators should cancel, so that the first term in (40) takes part only in the renormalization of Γ_2. The second term, corresponding to renormalization of τ, takes part only in the renormalization of Γ_3, since divergences of both subgraphs Γ_2 obviously correspond to renormalization of the component operators, and Γ_4 does not depend on the charge g_1. Finally, the third term in (40), corresponding to renormalization of λ, in which $Z_\lambda^{(1)}$ depends only on g_2 (see Eqs. (17)), takes part only in the renormalization of Γ_4, since the graphs Γ_2 and Γ_3 depend only on g_1. Thus, Γ_4^R is the sum of Γ_4 and the third term of (40), Γ_3^R is the sum of Γ_3 and the second term of (40), and Γ_2^R is the sum of Γ_2 and first term of (40) plus the two-loop contribution ΔC_2, the counterterm per operator pair.

In conclusion, we may add that the quantities $\Gamma_{2,3}^R$ can be found very simply: the expressions for $\Gamma_{2,3}$ in (38) are expressions of the type $A\cdot B$, and renormalization of the product reduces to renormalization of the factors A and B. Explicit expressions for the renormalized quantities are given in the Appendix.

5. GENERAL PROPERTIES OF THE SCALING FUNCTION $R(\omega)$; TRANSFORMATION TO THE NORMALIZED FUNCTION

Analyzing the diagrams of the unrenormalized response function $R(\omega)$ (their contributions differ from expressions (38) only in the values of the parameters), it is not hard to convince oneself that $R(\omega)$ is analytic in the upper half of the complex ω plane (as is any response function[20]) and real for purely imaginary values of ω on the upper half-axis, and also that $\mathrm{Im}R(\omega)>0$ for $\omega>0$. In addition it is easy to verify that for $\tau>0$ the first two terms of its Taylor series expansion about zero in ω (i.e., the contributions of order unity and ω) exist, and that for $\omega>0$ the limit $\tau\to 0$ is finite (higher terms of the expansions under discussion do not exist, due to IR divergences in the coefficients). Regarding the relation (7) between $R(\omega)$ and the scaling function $\Phi(w)$, it follows from what has been said that for $\Phi(w)$ the first two terms of the expansion about zero in w exist, and in the limit $w\to\infty$ this function has an asymptotic limit $\sim w^{-\alpha/z\nu}$, which ensures the existence of a finite limit for $R(\omega)$ as $\tau\to 0$. Also, taking into account the aforementioned real-valuedness and positivity, we conclude that for this function

$$\Phi(w)=_{w\to 0}B'+iC'w+\ldots,$$
$$\Phi(w)=_{w\to\infty}A'(-iw)^{-\kappa}+\ldots, \tag{41}$$

where $\kappa\equiv\alpha/z\nu$ and A', B', C' are positive coefficients.

Knowledge of the asymptotic limits (41) can be used to determine the normalization of the scaling function $\Phi(w)$. Indeed, by virtue of the uncertainty in the amplitude factors $C_{1,2}$ in Eq. (7), the scaling function $\Phi(w)$ can be replaced by any other function $f(w)=b_1\Phi(b_2w)$ with arbitrary positive constants $b_{1,2}$. This arbitrariness can be used to replace $\Phi(w)$ by the normalized function $f(w)$, for which two of the three coefficients A, B, and C can be determined in the asymptotic limit analogous to (41). The normalization procedure is useful because in contrast to $\Phi(w)$, the normalized function $f(w)$ does not depend on the subtraction scheme used in the renormalization. Moreover, an appropriate choice of the normalization substantially simplifies the form of $f(w)$. We employ the normalization for which $B=1$, $C=\alpha/z\nu\equiv\kappa$, and the third coefficient A remains unknown and remains to be calculated. Thus, for the normalized scaling function $f(w)$ we have

$$f(w)=b_1\Phi(b_2w), \tag{42}$$
$$f(w)=_{w\to 0}1+i\kappa w+\ldots,\quad f(w)=_{w\to\infty}A(-iw)^{-\kappa}+\ldots \tag{43}$$

The first of relations (43), for $\Phi(w)$ known, determines the coefficients $b_{1,2}$ in Eq. (42):

$$b_1=\frac{1}{B'}=\frac{1}{\Phi(0)},\quad b_2=\kappa\frac{B'}{C'}=i\kappa\frac{\Phi(0)}{\Phi'(0)}. \tag{44}$$

The amplitude $A=A'(B')^{-1}(\kappa B'/C')^{-\kappa}$ in relation (43), in contrast to A', B', C' in relations (41), is universal: it is the same for all materials from an experimental standpoint, and from a theoretical standpoint it does not depend on the choice of renormalization scheme in the calculations.

If we use some approximate expression for the scaling function $f(w)$ over a wide range of frequencies, it is important that this approximate function faithfully reproduce the asymptotic limit (43) at least qualitatively. We cannot simply use for $f(w)$ the initial terms in its ε expansion

$$f(w)=1+\varepsilon f_1(w)+\varepsilon^2 f_2(w)+\ldots, \tag{45}$$

because this fails to accurately reproduce the asymptotic limit as $w\to\infty$: the fractional power $w^{-\kappa}$ with $\kappa=O(\varepsilon)$ is replaced by its ε expansion $w^{-\kappa}=1-\kappa\ln w+\ldots$. This problem can be avoided if, for example, we factor out of $f(w)$ some function with qualitatively correct asymptotic limits and regular behavior in the intermediate range. Following Kroll and Ruhland,[16] we use the simple representation

$$f(w)=(1-iw)^{-\kappa}h(w),\quad \kappa\equiv\alpha/z\nu. \tag{46}$$

For the asymptotic limits of the function $h(w)$, from relations (46) and (43) we have

$$h(w)=_{w\to 0}1+0\cdot iw+\ldots,\quad h(w)=_{w\to\infty}A+\ldots. \tag{47}$$

The notation $0\cdot iw$ in relation (47) emphasizes the lack of a linear contribution in w to $h(w)$ as $w\to 0$.

A series expansion of $h(w)$ in ε does not share this failing. It can be obtained by successively finding the ε expansion of the renormalized response function R_R, then on the basis of definitions (35), (42), and (43) the coefficients $f_{1,2}$ in (45), and finally the ε expansion of the function $h(w)$ itself in (46). To obtain the latter it is necessary to know the

ε expansion of the exponent κ in (46); for α/ν it is given by (36), $z=4(1-9\varepsilon/38)+\ldots$ (Ref. 3); hence $\kappa=\varepsilon(1-479\varepsilon/1026)/12+\ldots$. The two-loop ε expansions of the functions Φ, f, and h are given in the Appendix. As can be seen from the asymptotic limits of the functions $f_{1,2}(w)$ as $w\to\infty$ derived in the Appendix, they contain powers of $\ln w$, which cancel in the asymptotic limit of $h(w)$ according to relation (47). This enables us to find the ε expansion of the universal amplitude A in (43):

$$A=1+\frac{\varepsilon}{6}+\frac{\varepsilon^2}{456}\left(\frac{89\pi^2}{4}-\frac{7267}{27}\right)+O(\varepsilon^3), \tag{48}$$

or numerically $A=1+0.1667\varepsilon-0.1086\varepsilon^2+\ldots$

6. DISCUSSION OF RESULTS; COMPARISON WITH EXPERIMENT

In the comparison with experiment the nonuniversal additive constant can be dropped by subtracting the quantity $c_\infty^{-2}=c^{-2}(\infty)$ from the first of equalities (5). Thus, from relations (5), (7), and (42) we obtain

$$F_1(\omega,\tau)\equiv[c^{-2}(\omega,\tau)-c^{-2}(\infty)]\tau^\alpha=\mathrm{Re}\,\bar f(\bar\omega),$$

$$F_2(\omega,\tau)\equiv\alpha_\lambda(\omega,\tau)c^{-2}(\omega,\tau)\tau^\alpha=\pi\,\mathrm{Im}\,\bar f(\bar\omega),$$

$$\bar f(\bar\omega)\equiv C_1 f(C_2\bar\omega),\quad \bar\omega\equiv\omega\tau^{-z\nu},\quad \tau\equiv(T-T_c)/T_c, \tag{49}$$

with the universal function $f(w)$ from Eq. (42) and nonuniversal factors $C_{1,2}$. It follows from the representation (49) that the experimentally measured quantities $F_{1,2}(\omega,\tau)$ depend in fact not on the two variables ω and τ, but only on their combination $\bar\omega=\omega\tau^{-z\nu}$ to within IR-negligible corrections. This is the fundamental assertion of dynamic scaling theory in regard to the problem under consideration, and it is confirmed by the experimental data.[13] In experiments with different ω and τ for one material the constants $C_{1,2}$ in Eqs. (49) are fixed, but they can change (in contrast to the function $f(w)$) if one goes to another material, and in this consists their nonuniversality.

In comparing Eqs. (49) with experiment we used the experimental data of Ref. 13, and for $f(w)$ we used expression (46) with experimental values of the exponent κ, including the first term (in the one-loop approximation) or the first two terms (in the two-loop approximation) of the ε expansion of the function $h(w)$. The experimental data are conveniently represented on a log–log scale, plotting $\ln F$ against $\ln\bar\omega$. The uncertainty in the coefficients $C_{1,2}$ in Eqs. (49) thus reduces to the possibility of translating the theoretical curves along the axes, with this translation being the same for F_1 and F_2. The ''shape'' of these curves for both F_1 and F_2, which is not altered by parallel translation along the axes, depends only on the scaling function $f(w)$. In constructing the curves in Figs. 1–3, the ''theoretical curves'' (49) were plotted as y on a logarithmic scale with $C_1=C_2=1$, and the arbitrariness in $C_{1,2}$ was used to advantage for the corresponding ''parallel translation'' of the experimental curves.

Figure 1 compares the experimental data with the dependence of the ratio $F_2/F_1=\pi\mathrm{Im}f(C_2\bar\omega)/\mathrm{Re}f(C_2\bar\omega)$ on the reduced frequency. The ratio F_2/F_1 depends on just one nonuniversal constant, C_2, and the uncertainty in C_2 corresponds to the possibility of a horizontal shift of the theoretical or experimental curve. It follows from (43) that $\pi\mathrm{Im}f(C_2\bar\omega)/\mathrm{Re}f(C_2\bar\omega)\to\pi\tan(\kappa\pi/2)$ as $\bar\omega\to\infty$ (note that this limit is a universal constant; numerically $\pi\tan(\kappa\pi/2)\simeq0.28$ for $\kappa\simeq0.057$). From the graph in Fig. 1 it can be seen that the experimental data agree with this prediction of the theory (in fact, based solely on the dynamic scaling hypothesis and the general properties (41) of the scaling function), and this asymptotic limit of F_2/F_1 is actually reached. Taking advantage of the possibility of horizontal translation, the theoretical curve for F_2/F_1 can also be aligned with the experimental curve in the low-frequency region; then the difference between experiment and theory shows up only at intermediate frequencies. The points in Fig. 1 plot the experimental data[13] for assorted values of ω and τ (note that for ^{3}He, which was the material used in the experiment in Ref. 13, it turns out that $c_\infty^{-2}\simeq0$; for a detailed discussion see Ref. 21). Curve *1* was constructed with the function $f(w)$, calculated in the one-loop approximation, curve *2* was constructed with the added two-loop contribution, and the dotted curve was calculated in Ref. 16. It is clear from the figure that for the one-loop approximation there is appreciable divergence from experiment at intermediate frequencies, but taking the two-loop approximation into account improves the situation significantly. The difference between the theoretical curve calculated in Ref. 16 in the one-loop approximation and curve *1* can be explained by the fact that in their calculation of the scaling function using a relation similar to (46), Kroll and Ruhland[16] replaced the coefficient ε in the one-loop expression $h=1+\varepsilon h_1$ in (A11) with the quantity $12\alpha(z\nu-2\alpha)^{-1}=\varepsilon+\ldots$ and substituted the experimental value 0.77 for $\varepsilon=1$ in the three-dimensional problem. As can be seen from the figure, this substitution yields qualitatively the same effect as the two-loop correction, improving the agreement with experiment. But from a theoretical standpoint, such a substitution is invalid within the framework of the ε expansion, as it ''oversteps the ε-accuracy'' of the one-loop scaling function. (Helpful remark: ''numerical experiments'' show that the two-loop approximation (A11) for h with $\varepsilon=1$

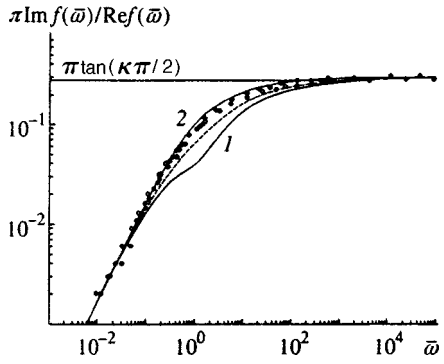


FIG. 1. The dependence of the ratio $\pi\mathrm{Im}f(\bar\omega)/\mathrm{Re}f(\bar\omega)$ on the reduced frequency $\bar\omega$ is compared with the experimental data of Roe and Meyer (D. B. Roe and H. Meyer, J. Low Temp. Phys. **30**, 91 (1978)) on the absorption in one wavelength $\alpha_\lambda(\omega)$ for ^{3}He. Curve *1* corresponds to the first order of the ε-expansion for the function $h(w)$, curve *2* corresponds to the second order. The dotted curve was constructed using the Kroll–Ruhland function.

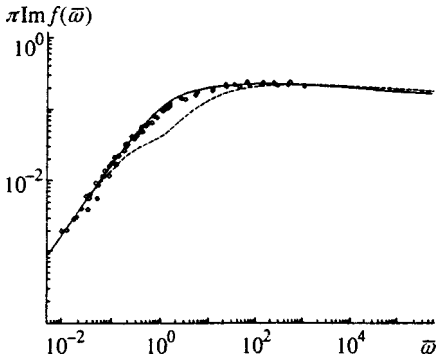


FIG. 2. Dependence of the reduced absorption $\alpha_\lambda(\omega)c^{-2}(\omega)\tau^\alpha$ on $\bar{\omega}$ according to the data of Roe and Meyer for ^{3}He. The dashed curve corresponds to calculation to order ε, and the solid curve to order ε^2.

is almost exactly reproduced by the one-loop expression $h=1+\varepsilon h_1$ if the coefficient $\varepsilon=1$ is replaced by 0.42. This can be used to construct graphs of the two-loop approximation.)

The foregoing discussion of overstepping the accuracy of the problem does not pertain to the exponents α and $z\nu$ in Eqs. (49), since calculating these exponents and the scaling function involves two independently solvable problems. Therefore, for the exponents we are not prohibited from using the experimental data for the three-dimensional problem: $\alpha=0.11$, $\nu=0.63$, $z=3.07$, $z\nu=1.93$, $\kappa=\alpha/z\nu=0.057$. In fact, just these data were used (as in Ref. 16) to construct the theoretical curves in Figs. 1–3.

Independently comparing the velocity dispersion and absorption coefficient of sound in (49) with experiment, the subtraction is usually carried out at zero frequency, because the constant $c(0)$ can more easily be extracted from the experimental data than the constant $c(\infty)$. This yields

$$[c^{-2}(0,\tau)-c^{-2}(\omega,\tau)]\tau^\alpha = C_1 \mathrm{Re}[1-f(C_2\bar{\omega})]. \qquad (50)$$

The experimental data[13] are compared with curves of the dispersion (50) and the absorption (the second of Eqs. (49)) in Figs. 2 and 3. The arbitrariness in $C_{1,2}$ was used to achieve the most accurate fit possible between the asymptotic limits of the experimental data and the theoretical curves constructed for $C_1=C_2=1$. For the sound absorption (Fig. 2) it

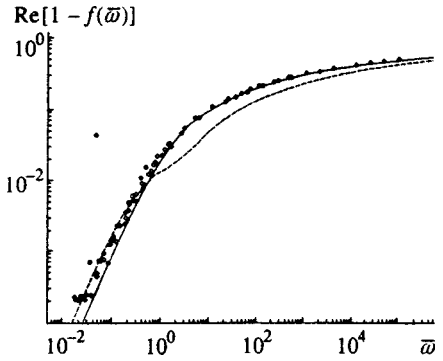


FIG. 3. Dependence of the reduced dispersion of the speed of sound $[c^{-2}(0)-c^{-2}(\omega)]\tau^\alpha$ on $\bar{\omega}$ according to the data of Roe and Meyer for ^{3}He. The dashed curve corresponds to calculation to order ε, and the solid curve to order ε^2.

was possible in this way to align both asymptotic limits, but for the velocity dispersion (Fig. 3) an appreciable discrepancy with experiment remains in the low-frequency region (for the reason see below). It can be seen from the graphs in Figs. 2 and 3 that in all cases taking the two-loop correction into account substantially improves the agreement with experiment.

We now discuss in more detail the low-frequency behavior of the velocity dispersion (50). As can be seen from the graph in Fig. 3, there is a considerable discrepancy here between the theoretical curve of the two-loop approximation and the experimental data, which shows up in particular as a difference in slope. This is not a random result. The point is that in general the first two terms (of order unity and $i\omega$) of the low-frequency asymptotic limit $f(w)$ (43) do not contribute to expression (50) due to the subtraction at zero frequency made there. The low-frequency asymptotic limit (50) is thus not determined by the correction terms shown in (43). As can be seen from Eqs. (A14), the ε expansion yields corrections of the form

$$w^2[(a_1\varepsilon+a_2\varepsilon^2)\ln w+(b_2\varepsilon^2+\ldots)\ln^2 w+\ldots]. \qquad (51)$$

It is clear that higher powers of $\ln w$ will appear in higher orders of the ε expansion, so that for a correct determination of the low-frequency asymptotic limit of the dispersion (50), it is necessary to sum the ''leading logarithms'' $\sim(\varepsilon\ln w)^n$ in all orders of the ε expansion.

The solution of this problem does not pertain just to the competency of the RG method. Analogous problems arise in the theory of critical behavior, for example, in the study of the asymptotic limit $\tau\to 0$ of the static correlator in the simple ψ^4 model, and can be solved with the help of the Wilson operator expansion,[11] where logarithms of the type (51) are summed, as a rule, into some fractional exponents. By analogy, it is natural to assume that the result of summing over the leading logarithms in expression (51) will be some fractional power $w^{2+a(\varepsilon)}$ with $a(\varepsilon)=O(\varepsilon)$. An analysis of the diagrams confirms this assumption. The asymptotic limit of the one-loop graph Γ_1^R (A3), calculated beyond the scope of the ε expansion (with finite $0<\varepsilon<2$), yields $a=-\varepsilon/2$ for the exponent a. It is important that in the given case, as an analysis of the two-loop diagrams $\Gamma_{2,3,4}^R$ (A3) shows, higher orders of perturbation theory lead to the same power-law asymptotic limit $\sim w^{2-\varepsilon/2}$. It therefore stands to reason that an improved low-frequency asymptotic limit (43) of the function $f(w)$ has the form

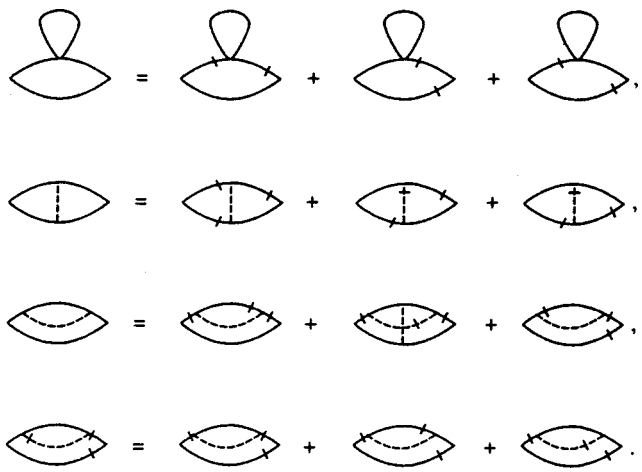
$$f(w)=1+i\kappa w+D(-iw)^{2-\varepsilon/2}+O(w^2) \qquad (52)$$

with a real coefficient $D=D(\varepsilon)$. It is not hard to show that expanding the D term in Eq. (52) in a series in ε leads to an asymptotic limit that agrees with that obtained in Eq. (A13) via the ε expansion. The initial slope 3/2 of the velocity dispersion (50) predicted by (52) (for $\varepsilon=1$) is in good agreement with the experimental data. A rigorous proof of this relation can be obtained with the help of ''infrared perturbation theory.''[22,23]

This work was carried out with the support of the Russian Fund for Fundamental Research (Grant No. 96-02-17-033) and the Goskomvuz Concourse Center for Basic Science (Grant No. 97-0-14.1-30).

APPENDIX A: RESULTS OF CALCULATIONS

We expand the terms in diagrams $\Gamma_{3,4}$ of Eqs. (37):



(A1)

Only diagrams contributing in the H_0 model are shown.

The renormalized analogs Γ_i^R of the quantities Γ_i (38), defining the renormalized response function

$$R_R(\omega,\tau)=\sum_{i=1}^{4}\Gamma_i^R(\omega,\tau), \tag{A2}$$

are given by the expressions (we use the notation of Eqs. (39) throughout)

$$\Gamma_1^R=\int \mathscr{D}\mathbf{k}\ \frac{\lambda k^2}{\alpha_k(2\varepsilon_k-i\omega)}-\frac{\mu^{-\varepsilon}}{16\pi^2\varepsilon},\quad \Gamma_2^R=-g_1\mu^{\varepsilon}(\Gamma_1^R)^2,$$

$$\Gamma_3^R=g_1\mathscr{F}^R\partial_\tau\Gamma_1^R,\quad \mathscr{F}^R=-\frac{1}{2}\tau\left(\int \mathscr{D}\mathbf{q}\ \frac{\mu^{\varepsilon}}{q^2\alpha_q}-\frac{1}{8\pi^2\varepsilon}\right),$$

$$\Gamma_4^R=-\lambda g_2\mu^{\varepsilon}i\omega\left\{\int \mathscr{D}\mathbf{k}\int \mathscr{D}\mathbf{q}\right.$$

$$\times\frac{k^2q^2-(\mathbf{kq})^2}{|\mathbf{k}+\mathbf{q}|^4\alpha_k\alpha_q(2\varepsilon_k-i\omega)^2(2\varepsilon_q-i\omega)^2}$$

$$-\int \mathscr{D}\mathbf{k}\ \frac{\tau}{\alpha_k(2\varepsilon_k-i\omega)^2}\int \mathscr{D}\mathbf{q}\ \frac{k^2q^2-(\mathbf{kq})^2}{q^2|\mathbf{k}+\mathbf{q}|^4\alpha_q}$$

$$\left.-\int \mathscr{D}\mathbf{k}\ \frac{k^2}{\alpha_k(2\varepsilon_k-i\omega)^2}\left[\int \mathscr{D}\mathbf{q}\ \frac{k^2q^2-(\mathbf{kq})^2}{k^2q^2|\mathbf{k}+\mathbf{q}|^4}-\frac{3\mu^{-\varepsilon}}{32\pi^2\varepsilon}\right]\right\}. \tag{A3}$$

The quantity Γ_4^R is given in a form suitable for calculation—each term in braces in its definition remains finite as $\varepsilon\to 0$.

The functions

$$\overline{\Gamma}_i(w)\equiv 32\pi^2\mu^{\varepsilon}\Gamma_i^R(\tau=\mu^2,\omega=2\lambda w\mu^4,g=g_*) \tag{A4}$$

calculated according to formulas (A3) with two terms of the ε expansion for $\overline{\Gamma}_1$ and one term for $\overline{\Gamma}_{2,3,4}$ taken into account, have the following form:

$$\overline{\Gamma}_1=\beta+\varepsilon(\beta^2+\pi^2/6)/4+iw\{q_1[1+(\beta+1)\varepsilon/2]-\varepsilon q_2/2\},$$

$$\overline{\Gamma}_2=-\varepsilon(\beta+iwq_1)^2/6,\quad \overline{\Gamma}_3=(\beta+1)\varepsilon(1+2iwq_3)/6,$$

$$\overline{\Gamma}_4=9\varepsilon iw[q_4+2q_5-(\beta+4/3)q_3]/19. \tag{A5}$$

Here $\beta\equiv\ln 4\pi-C$, C is the Euler constant, $q_i\equiv q_i(w)$ are functions defined by the relations

$$q_1=\int_0^\infty dx\,\frac{x}{(x+1)^2(\tilde{x}+z)},\quad \tilde{x}\equiv x(x+1),\quad z\equiv -iw,$$

$$q_2=\int_0^\infty dx\ \frac{x\ln x}{(x+1)^2(\tilde{x}+z)},\quad q_3=\partial_z(zq_1),$$

$$q_4=\int_0^\infty dx\ \frac{x+(x^2-1)\ln(x+1)}{(x+1)(\tilde{x}+z)^2},$$

$$q_5=\int_0^\infty dx\int_0^x dy\ \frac{y^2}{(x+1)(\tilde{x}+z)(y+1)(\tilde{y}+z)}$$

$$=q_6\ln z+\frac{(1-3z)}{2}q_6^2+\int_0^\infty dx\,\frac{3\ln(x+1)-\ln(\tilde{x}+z)}{(x+1)(\tilde{x}+z)},$$

$$q_6=\int_0^\infty\frac{dx}{(x+1)(\tilde{x}+z)}. \tag{A6}$$

The integrals $q_{1,2,3,6}$ can be expressed, if desired, in terms of elementary functions: in particular,

$$q_1=-\frac{1}{z}-\frac{1-z}{2z^2}\ln z-\frac{1-3z}{2z^2\Delta}\ln\left[\frac{1+\Delta}{1-\Delta}\right], \tag{A7}$$

where $\Delta\equiv(1-4z)^{1/2}$. The logarithms of complex arguments are defined in the above relations in the usual way: $\ln z=\ln|z|+i\arg z$, $-\pi<\arg z<\pi$ with $\arg z=0$ for $z>0$. Invoking relations (35), (36), (A2), and (A4) we find the scaling function $\Phi(w)$ out to terms $\sim\varepsilon^2$ inclusive:

$$\Phi(w)=1+\frac{\varepsilon}{6}\left(1-\frac{19}{27}\varepsilon\right)\sum_{i=1}^{4}\overline{\Gamma}_i(w). \tag{A8}$$

Substituting the above expressions for $\overline{\Gamma}_i$ (A5) into Eq. (A.8) yields

$$\Phi(w)=1+\frac{\varepsilon}{6}(\beta+iwq_1)+\frac{\varepsilon^2}{6}\left\{\frac{\beta^2}{16}+\frac{\pi^2}{24}-\frac{29\beta}{54}+\frac{1}{6}\right.$$

$$+iw\left[\left(\frac{\beta}{6}-\frac{11}{54}\right)q_1-\frac{iwq_1^2}{6}-\frac{q_2}{2}-\frac{q_3(8\beta+17)}{57}\right.$$

$$\left.\left.+\frac{9(q_4+2q_5)}{19}\right]\right\}. \tag{A9}$$

From (42), (44), and (A9) we obtain the ε expansion for the normalized function $f(w)$:

$$f(w)=1+\varepsilon f_1(w)+\varepsilon^2 f_2(w)+\ \ldots\ , \quad f_1=iwq_1/6,$$

$$f_2=\frac{iw}{12}\Big\{-\frac{11}{27}q_1-\frac{1}{3}iwq_1^2-q_2+\frac{1}{19}[(241-30\pi^2)q_3 + 18(q_4+2q_5)]\Big\}. \tag{A10}$$

From the definition (46) of the function $h(w)$ in terms of the known ε expansion of the function $f(w)$ and of the index $\kappa=\alpha/z\nu$ in (46), we find the ε expansion of $h(w)$:

$$h(w)=1+\varepsilon h_1(w)+\varepsilon^2 h_2(w)+\ \ldots,$$

$$h_1=f_1+\frac{1}{12}\ln(1-iw),$$

$$h_2=f_2+\frac{1}{12}f_1\ln(1-iw)-\frac{479}{12312} \times\ln(1-iw)+\frac{1}{288}\ln^2(1-iw) \tag{A11}$$

with $f_{1,2}$ from (A10).

We now give the asymptotic limits of the resulting functions. The functions q_i given by Eqs. (A6), which enter into expressions (A9) and (A10), have the following asymptotic limits (everywhere $z\equiv -i\omega$): for $w\to 0$

$$q_1=\frac{1}{2}+z\ \ln\ z+O(z), \quad q_2=-\frac{1}{2}+\frac{1}{2}z\ \ln^2\ z+O(z),$$

$$q_3=\frac{1}{2}+2z\ \ln\ z+O(z), \quad q_4=\frac{\pi^2-9}{2}+z\ \ln\ z+O(z),$$

$$q_5=\frac{\pi^2-9}{6}+O(z), \tag{A12}$$

and for $w\to\infty$

$$q_1=\frac{\ln\ z}{2z}-\frac{1}{z}+o\left(\frac{1}{z}\right), \quad q_2=\frac{\ln^2 z}{8z}-\frac{\pi^2}{8z}+o\left(\frac{1}{z}\right),$$

$$q_3=\frac{1}{2z}+o\left(\frac{1}{z}\right), \quad q_4=\frac{\ln\ z}{4z}+o\left(\frac{1}{z}\right), \quad q_5=\frac{\pi^2}{24z}+o\left(\frac{1}{z}\right). \tag{A13}$$

From expressions (A12) and (A13) we obtain the following asymptotic limits for the functions $f_{1,2}$ (A.10): for $w\to 0$

$$f_1=-\frac{1}{12}z-\frac{1}{6}z^2\ln\ z+O(z^2),$$

$$f_2=\frac{1}{12}\Big\{\frac{479}{1026}z+\frac{1}{2}z^2\ln^2 z+\Big[\frac{20}{19}(3\pi^2-25)+\frac{11}{27}\Big]z^2\ln\ z\Big\} + O(z^2), \tag{A14}$$

and for $w\to\infty$

$$f_1=-\frac{1}{12}\ln\ z+\frac{1}{6}+O\left(\frac{1}{z}\right),$$

$$f_2=\frac{1}{288}\ln^2\ z+\frac{77}{3078}\ln\ z+\frac{1}{8}\left(\frac{89\pi^2}{228}-\frac{7267}{1539}\right)+O\left(\frac{1}{z}\right). \tag{A15}$$

Substituting the asymptotic limits (A14) into formulas (A11) yields

$$h=1-\frac{1}{6}\varepsilon z^2\ln\ z+\frac{1}{12}\varepsilon^2 z^2 \times\Big\{\frac{1}{2}\ln^2\ z+\Big[\frac{20}{19}(3\pi^2-25)+\frac{11}{27}\Big]\ln\ z\Big\}+O(z^2), \tag{A16}$$

which agrees with the first of relations (47). Substituting (A15) into (A11) leads to the second of relations (47) with the constant A given in Eq. (48). All occurrences of $\ln z$ in (A15) cancel out in the analogous asymptotic limit (A11).

*)E-mail: adjemyan@snoopy.niif.spb.su
†)E-mail: Alexander.Vasiljev@pobox.spbu.ru
‡)E-mail: saha@snoopy.niif.spb.su

[1] P. Hohenberg and B. Halperin, Rev. Mod. Phys. **49**, 435 (1977).
[2] A. Z. Patashinskiĭ and V. L. Pokrovskiĭ, *Fluctuation Theory of Phase Transitions*, Pergamon Press, Oxford (1979).
[3] E. Siggia, B. Halperin, and P. Hohenberg, Phys. Rev. B **13**, 2110 (1976).
[4] R. Dengler and F. Schwable, Europhys. Lett. **4**, 1233 (1987).
[5] E. V. Gurovich, E. I. Kats, V. V. Lebedev, and A. R. Muratov, JETP Lett. **55**, 57 (1992).
[6] E. V. Gurovich, E. I. Kats, and V. V. Lebedev, Zh. Éksp. Teor. Fiz. **100**, 855 (1991) [Sov. Phys. JETP **73**, 473 (1991)].
[7] M. Fixman, J. Chem. Phys. **36**, 1961 (1962).
[8] W. Botch and M. Fixman, J. Chem. Phys. **42**, 199 (1965).
[9] R. A. Ferrel and J. K. Bhattacharjee, Phys. Lett. A **86**, 109 (1981).
[10] R. A. Ferrel and J. K. Bhattacharjee, Phys. Rev. A **31**, 1788 (1985).
[11] J. Zinn-Justin, *Quantum Field Theory and Critical Phenomena*, Oxford Univ. Press, Oxford (1989).
[12] H. C. Burstyn, J. V. Sengers, J. K. Bhattacharjee, and R. A. Ferrel, Phys. Rev. A **28**, 1567 (1983).
[13] D. B. Roe and H. Meyer, J. Low Temp. Phys. **30**, 91 (1978).
[14] D. B. Roe, B. A. Wallace, and H. Meyer, J. Low Temp. Phys. **16**, 51 (1974).
[15] D. Sarid and D. S. Cannel, Phys. Rev. A **15**, 735 (1977).
[16] D. M. Kroll and J. M. Ruhland, Phys. Lett. A **80**, 45 (1980).
[17] R. Folk and G. Moser, Phys. Rev. Lett. **75**, 2706 (1995).
[18] C. de Dominicis and L. Peliti, Phys. Rev. B **18**, 353 (1978).
[19] R. Baush, H. K. Janssen, and H. Z. Wagner, Z. Phys. B **24**, 113 (1976).
[20] L. D. Landau and E. M. Lifshitz, *Statistical Physics*, Pt. 1, 3rd ed., Pergamon Press, Oxford (1980).
[21] J. K. Bhattacharjee and R. A. Ferrel, Phys. Lett. A **88**, 77 (1982).
[22] L. Ts. Adzhemyan, A. N. Vasil'ev, and Yu. M. Pis'mak, Teor. Mat. Fiz. **74**, 360 (1988).
[23] M. Yu. Nalimov, Teor. Mat. Fiz. **80**, 212 (1989).

Translated by Paul F. Schippnick

Lattice dynamics of K_2NaAlF_6, K_3AlF_6, and Na_3AlF_6 crystals with the elpasolite structure

V. I. Zinenko*) and N. G. Zamkova

L. V. Kirenskiĭ Physics Institute Siberian Section, Russian Academy of Sciences, 660036 Krasnoyarsk, Russia

S. N. Sofronova

Krasnoyarsk State University, 660041 Krasnoyarsk, Russia

(Submitted 11 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1742–1756 (November 1998)

This paper presents the results of a nonempirical calculation of the static and dynamic properties of K_2NaAlF_6, K_3AlF_6, and Na_3AlF_6 crystals with the elpasolite structure. The calculation is based on a microscopic model of an ionic crystal that allows for the deformability and polarizability of the ions. The deformability parameters of the ions are determined by minimizing the total energy of the crystal. The total energy is regarded as a functional of the electron density, using the local Thomas–Fermi approximation and taking into account exchange (correlation) effects. The results of the calculations of the equilibrium lattice parameters and of the permittivities are in good agreement with the experimental data. Unstable vibrational modes are found in the spectrum of the lattice vibrations, with these modes occupying the phase space in the entire Brillouin zone. © *1998 American Institute of Physics.* [S1063-7761(98)01311-0]

1. INTRODUCTION

The family of crystals with the elpasolite structure $A_2BB^{3+}X_6$ can be classified as perovskite-like compounds, a typical structural feature of which is the presence of octahedral groups. Most crystals of this family, like the representatives of the perovskite family, experience diverse structural phase transitions associated with instability of the crystal lattice against various vibrational lattice modes.

Crystals of the elpasolite family in the high-symmetry phase belong to the cubic space group O_h^5, with a face-centered lattice. The unit cell contains one molecule. Depending on the chemical composition, various distorted low-symmetry phases are observed, with sequences of structural phase transitions being detected in many crystals of this family.

Compounds with the elpasolite structure have been intensively studied by various methods, and by now there is much experimental information for many crystals of the given family concerning the structures, the physical properties, and their changes during phase transformations. In particular, Raman scattering and inelastic neutron scattering in certain crystals have been used to determine the soft vibrational modes of the crystal lattice.[1] The experimental data on the structures of the low-symmetry phases and the soft modes of the lattice vibrations are evidence that, in most of the compounds of the elpasolite family that have been studied, the phase transitions are associated with small rotations of the $B^{3+}X_6$ octahedra. However, it is also experimentally known that the structures of the distorted phases in certain elpasolites correspond not only to rotations of the octahedra, but also to substantial displacements of the A and B ions from the equilibrium positions of the cubic phase. There have been virtually no calculations of the frequency spectrum of the lattice vibrations in crystals of the elpasolite family. Such a calculation of the incomplete vibrational spectrum of the $Cs_2NaTmBr_6$ crystal in the rigid-ion model is given in Ref. 1. Since the unit cell of elpasolite contains ten atoms, a large number of unknown parameters are needed in the rigid-ion model to take into account short-range forces (Ref. 1 used nine parameters). For this reason, it is difficult to use the rigid-ion model to study the crystal lattice's instability against one vibrational mode or another as a function of the chemical composition of the compounds.

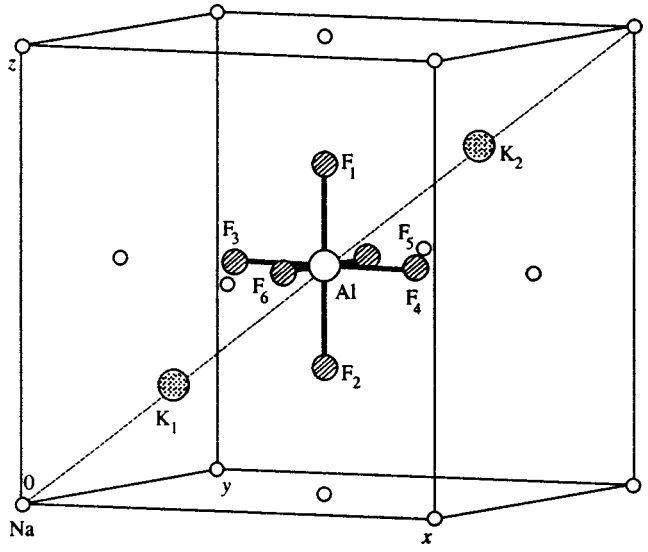


FIG. 1. Structure of the elpasolite K_2NaAlF_6. One molecule and the face-centered Na lattice are shown.

The goal of this paper is to calculate from first principles the equilibrium volume, the total spectrum of the lattice vibrations, and the radio-frequency (rf) permittivity in K_2NaAlF_6, K_3AlF_6, and Na_3AlF_6 crystals in terms of the generalized Gordon–Kim model proposed by Ivanov and Maksimov.[2]

Section 2 presents the results of a group-theoretical analysis of the normal modes of the lattice vibrations of the elpasolite structure for all symmetry points and directions of the Brillouin zone. The results of such an analysis appear in the literature for only two symmetry points (Γ and X) of the Brillouin zone.[1,3] The model and the method of computing the frequencies of the lattice normal modes and the rf permittivity are presented in Sec. 3. The results of the calculations and a discussion of the results are presented in Sec. 4.

2. SYMMETRY ANALYSIS OF THE NORMAL MODES

The crystal structure of elpasolite in the high-symmetry phase is cubic with space group $O_h^5(Fm3m)$. The ions occupy ten interpenetrating fcc lattices, as shown in Fig. 1.

The characters of the various symmetry elements in the vibrational representation are as follows (only the symmetry elements associated with the z direction are shown below, $\tau = a_0/2$, and a_0 is the lattice parameter):

the identity element

$$\chi(E) = 30,$$

rotation about a fourfold axis

$$\chi(C_{4z}) = 1 + 3\exp(i\mathbf{q}\cdot(2\tau,0,0)),$$

rotation about a twofold axis

$$\chi(C_{2z}) = (-1)[1 + 3\exp(i\mathbf{q}\cdot(2\tau,2\tau,0)) + \exp(i\mathbf{q}\cdot(\tau,\tau,0)) + \exp(i\mathbf{q}\cdot(3\tau,3\tau,0))],$$

rotation about a twofold axis along a face diagonal

$$\chi(C_{2xy}) = (-1)[1 + \exp(i\mathbf{q}\cdot(2\tau,2\tau,2\tau))],$$

reflection in a plane perpendicular to a twofold axis

$$\chi(C_{\sigma_z}) = 1 + 5\exp(i\mathbf{q}\cdot(0,0,2\tau)),$$

reflection in a plane perpendicular to a face diagonal

$$\chi(C_{\sigma_{xy}}) = 1 + 3\exp(i\mathbf{q}\cdot(2\tau,2\tau,0)) + \exp(i\mathbf{q}\cdot(\tau,\tau,0)) + \exp(i\mathbf{q}\cdot(3\tau,3\tau,0)),$$

TABLE I. Displacements of ions of an elpasolite in the normal modes of the center and of boundary point X of the Brillouin zone.

Irreducible representation	Normal mode	Number of modes
	Zone center	
A_{1g}	$-F_{1z}=F_{2z}=-F_{3y}=F_{4y}=F_{5x}=-F_{6x}$	1
E_g	$-F_{1z}=-F_{2z}=F_{3y}=F_{4y}=-F_{5x}=F_{6x}$ $-F_{1z}=F_{2z}=F_{3x}=-F_{4x}=F_{5y}=F_{6y}$	1
T_{1g}	$-F_{1y}=F_{2y}=F_{5z}=-F_{6z}$ $-F_{1x}=F_{2x}=-F_{3z}=F_{4z}$ $-F_{3y}=F_{4y}=-F_{5x}=F_{6x}$	1
T_{2g}	$K_{1x}=K_{1y}=K_{1z}=-K_{2x}=-K_{2y}=-K_{2z}$; $-F_{1x}=-F_{1y}=F_{2x}=F_{2y}=F_{3y}=F_{3z}=-F_{4y}=-F_{4z}$ $=-F_{5x}=-F_{5z}=F_{6x}=F_{6z}$ $K_{1x}=-K_{2x}$; $-K_{1y}=K_{2y}$; $K_{1z}=K_{2z}$; $F_{1x}=-F_{2x}=-F_{3z}=F_{4z}$; $-F_{1y}=F_{2y}=F_{5z}=F_{6z}$; $F_{3y}=-F_{4y}=-F_{5x}=F_{6x}$ $K_{1x}=-K_{2x}$; $K_{1y}=-K_{2y}$; $-K_{1z}=K_{2z}$; $-F_{1y}=F_{2y}-F_{5z}=F_{6z}$; $-F_{1y}=F_{2y}=-F_{5z}=F_{6z}$; $-F_{3y}=F_{4y}=F_{5x}=-F_{6x}$	2
T_{2u}	$F_{1y}=F_{2y}=-F_{5y}=-F_{6y}$ $F_{1x}=F_{2x}=-F_{3x}=-F_{4x}$ $F_{3z}=F_{4z}=-F_{5z}=-F_{6z}$	1
T_{1u}	all ions are displaced	4
	point X	
τ_3	$F_{3y}=-F_{4y}=F_{5x}=-F_{6x}$	1
τ_5	$F_{3x}=-F_{4x}=F_{5y}=-F_{6y}$	1
τ_7	$F_{3y}=-F_{4y}=-F_{5x}=F_{6x}$	1
τ_8	$F_{3z}=F_{4z}=-F_{5z}=-F_{6z}$	1
τ_6	$K_{1z}=-K_{2z}$	1
τ_1	all ions are displaced	3
τ_9	all ions are displaced	3
τ_{10}	all ions are displaced	6

inversion

$$\chi(J)=(-3)[1+\exp(i\mathbf{q}\cdot(2\tau,2\tau,2\tau))],$$

inverted rotation by 60°

$$\chi(S_6)=0,$$

inverted rotation by 90°

$$\chi(S_{4z})=(-1)[1+\exp(i\mathbf{q}\cdot(0,2\tau,2\tau))+\exp(i\mathbf{q}\cdot(0,\tau,\tau))+\exp(i\mathbf{q}\cdot(0,3\tau,3\tau))].$$

The expansion of the modal representation T into irreducible representations can be found by the standard procedure:[4]

$$C_\lambda=n^{-1}\sum_g \chi(\mathbf{q},g)\chi^\lambda(\mathbf{q},g), \tag{1}$$

where n is the order of the group of wave vector $\mathbf{q}$, and $\chi^\lambda(\mathbf{q},g)$ is the character of the small representation of the group of vector $\mathbf{q}$. This decomposition has the following form for the symmetry points and directions of the Brillouin zones of the fcc lattice (the symbols for the wave vectors and irreducible representations are from Kovalev's tables;[5] for the zone center, the standard symbols for the representations of the space groups are shown in parentheses):

(a) Center of the Brillouin zone, $\mathbf{q}=(0,0,0)$

$$T=\tau_1(A_{1g})+\tau_3(E_g)+\tau_5(T_{1g})+2\tau_4(T_{2g})+\tau_9(T_{2u})+5\tau_{10}(T_{1u}).$$

Here the splitting of the longitudinal and transverse optical frequencies of symmetry T_{1u} by the macroscopic electric field is neglected.

(b) $\mathbf{q}=(0,0,2\mu\pi/\tau)$

$$T=7\tau_1+\tau_2+2\tau_3+2\tau_4+9\tau_5.$$

The mode with symmetry τ_5 is doubly degenerate. At the zone boundary (point X),

$$T=3\tau_1+\tau_3+4\tau_4+\tau_5+\tau_6+\tau_7+\tau_8+3\tau_9+6\tau_{10},$$

τ_9 and τ_{10} correspond to doubly degenerate modes.

(c) $\mathbf{q}=(2\mu\pi/\tau,\ 2\mu\pi/\tau,0)$ (point K corresponds to $\mathbf{q}=(3\pi/4\tau,3\pi/4\tau,0)$)

$$T=10\tau_2+4\tau_2+8\tau_3+8\tau_4.$$

(d) $\mathbf{q}=(\mu\pi/\tau,\mu\pi/\tau,\mu\pi/\tau)$

$$T=8\tau_1+2\tau_2+10\tau_3,$$

τ_3 corresponds to doubly degenerate modes. At the zone boundary (point L),

$$T=4\tau_1+\tau_2+\tau_3+4\tau_4+5\tau_5+5\tau_6,$$

τ_5 and τ_6 correspond to doubly degenerate modes.

(e) $\mathbf{q}=(0,\pi/\tau,\pi/2\tau)$ (point W)

$$T=5\tau_1+2\tau_2+5\tau_3+2\tau_4+8\tau_5,$$

τ_5 corresponds to doubly degenerate modes.

The displacements of the ions in certain normal modes are given in Table I.

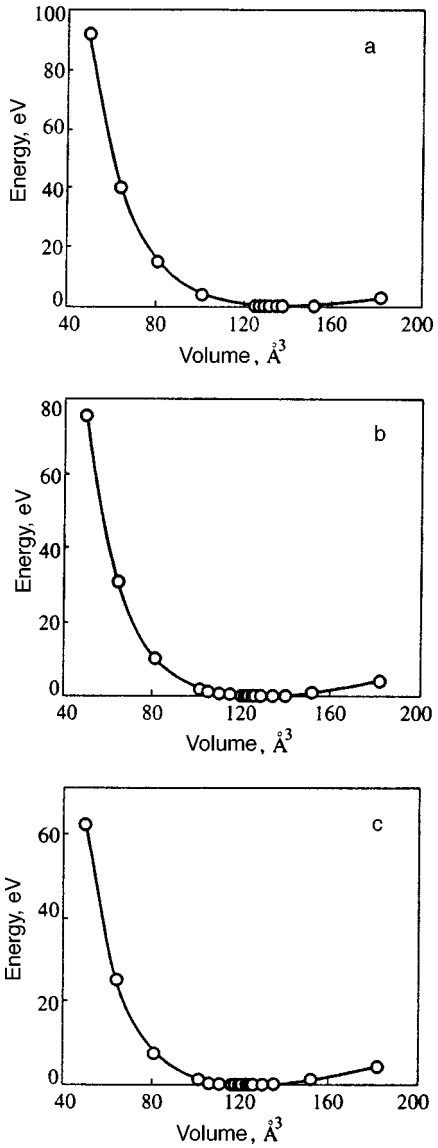


FIG. 2. Dependence of the total energy of the crystal on the volume: (a) K_3AlF_6; (b) K_2NaAlF_6; (c) Na_3AlF_6. The origin of the energy readings in (a), (b), and (c) corresponds to 72 784 eV, 60 751 eV, and 36 683 eV.

3. MODEL. METHOD OF CALCULATION

The model of the ionic crystal proposed by Ivanov and Maximov,[2] which takes into account the polarizability of the ions, is used to compute the frequency spectrum of the lattice vibrations of crystals of the elpasolite family. In this model, the ionic crystal is represented as consisting of individual intersecting spherically symmetric ions. The total electron density of the crystal in this case can be written

$$\rho(\mathbf{r})=\sum_i \rho_i(\mathbf{r}-\mathbf{R}_i),$$

where symmetrization is carried out over all the ions of the crystal.

The total energy of the crystal in terms of the density-functional method, taking into account only pairwise interaction, has the form

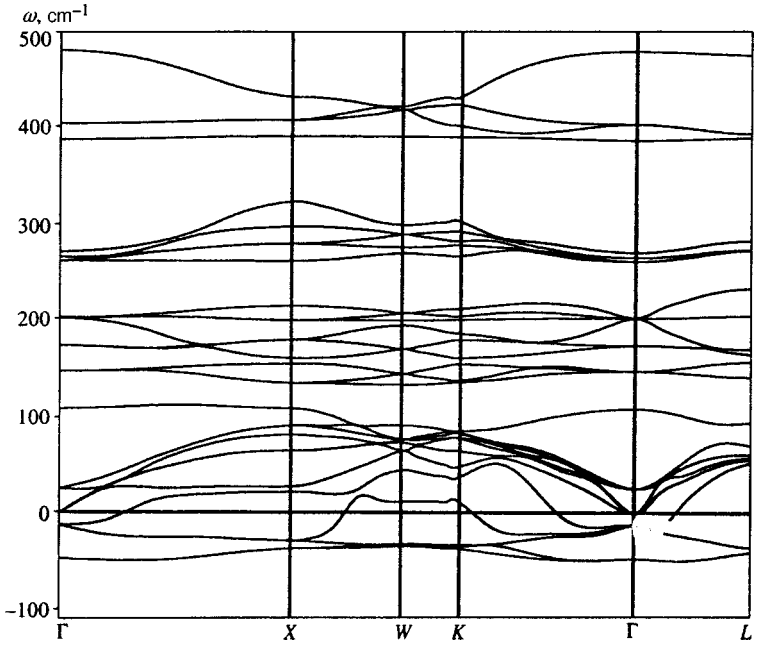


FIG. 3. Calculated dispersion curves for cubic K_2NaAlF_6. (Imaginary frequencies are indicated by negative values.)

$$E_{\rm cr}=\frac{1}{2}\sum_{i\neq j}\frac{Z_iZ_j}{|\mathbf{R}_i-\mathbf{R}_j|}+\sum_i E_i^{\rm self}(R_w^i)$$
$$+\frac{1}{2}\sum_{i\neq j}\Phi_{ij}(R_w^i,R_w^j,|\mathbf{R}_i-\mathbf{R}_j|), \qquad (2)$$

where Z_i is the charge of the ith ion,

$$\Phi_{ij}(R_w^i,R_w^j,|\mathbf{R}_i-\mathbf{R}_j|)=\mathbf{E}\{\rho_i(\mathbf{r}-\mathbf{R}_i)+\rho_j(\mathbf{r}-\mathbf{R}_j)\}$$
$$-\mathbf{E}\{\rho(\mathbf{r}-\mathbf{R}_i)\}-\mathbf{E}\{\rho(\mathbf{r}-\mathbf{R}_j)\}\ , \qquad (3)$$

energy $\mathbf{E}\{\rho\}$ is calculated by the density-functional method,[2] using the local approximation for the kinetic and exchange (correlation) energies, and $E_i^{\rm self}(R_w^i)$ is the self-energy of the ion. The electron density of an individual ion and its self-energy are calculated taking into account the crystal potential, approximated by a charged sphere (the Watson Sphere)

$$v(r)=\begin{cases} Z_i^{\rm ion}/R_w, & r<R_w \\ Z_i^{\rm ion}/r, & r>R_w \end{cases},$$

where R_w is the radius of the Watson sphere. The radii R_w^i of the spheres at individual ions are found by minimizing the total energy of the crystal.

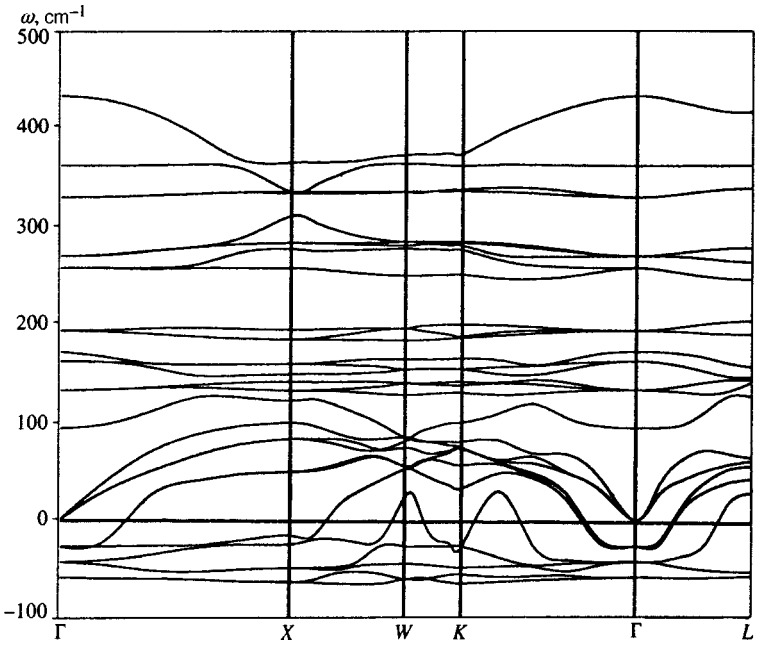


FIG. 4. Calculated dispersion curves for cubic K_3AlF_6. (Imaginary frequencies are indicated by negative values.)

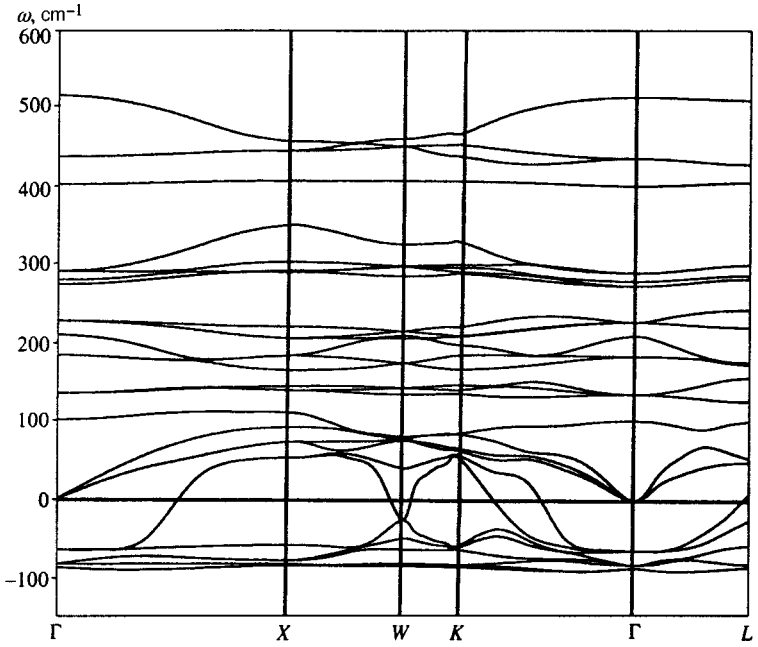


FIG. 5. Calculated dispersion curves for cubic Na_3AlF_6. (Imaginary frequencies are indicated by negative values.)

To calculate the lattice dynamics in the expression for the energy of the crystal, Eq. (2), it is necessary to add terms that describe the energy changes caused by displacing the ions from their equilibrium positions. When the frequencies of the lattice vibrations of the ionic crystals were calculated, the electronic polarizability of the ions and the ''breathing'' of the ion in the crystal environment were taken into account both in terms of the phenomenological models of Ref. 6 and the microscopic approach of Ref. 7. In the model considered here, the expression for the dynamic matrix has the form

$$
\begin{aligned}
D_{jj'}^{\alpha\beta}=\frac{\exp(-i\mathbf{q}(\mathbf{x}_j-\mathbf{x}_{j'}))}{\sqrt{M_jM_{j'}}}\Bigg\{&\frac{1}{2}Z_j^{\text{ion}}Q^{\alpha\beta}(\mathbf{q};jj')Z_{j'}^{\text{ion}}\\
&+\Phi_{RR}^{\alpha\beta}(\mathbf{q};jj')+\sum_{k=1}^{N}\left[\Phi_{Rv}^{\alpha}(\mathbf{q};jk)V^{\beta}(\mathbf{q};kj')\right.\\
&\left.+V^{*\alpha}(\mathbf{q};kj)\Phi_{vR}^{\beta}(\mathbf{q};kj')\right]\\
&+\sum_{k,k'=1}^{N}V^{*\alpha}(\mathbf{q};kj)\Phi_{vv}(\mathbf{q};kk')V^{\beta}(\mathbf{q};k'j')\\
&+\sum_{\gamma=1}^{3}\sum_{k=1}^{N}\left[W^{*\gamma\alpha}(\mathbf{q};kj)\Phi_{wR}^{\gamma\beta}(\mathbf{q};jk)\right.\\
&\left.+\Phi_{Rw}^{\alpha\gamma}(\mathbf{q};jk)W^{\gamma\beta}(\mathbf{q};kj')\right]+\sum_{\gamma,\gamma'=1}^{3}\sum_{k,k'=1}^{N}W^{*\gamma\alpha}\\
&\times(\mathbf{q};kj)\Phi_{ww}^{\gamma,\gamma'}(\mathbf{q};kk')W^{\gamma'\beta}(\mathbf{q};k'j')\\
&+\sum_{\gamma=1}^{3}\sum_{k,k'=1}^{N}\left[W^{*\gamma\alpha}(\mathbf{q};kj)\Phi_{wv}^{\gamma}(\mathbf{q};kk')V^{\beta}(\mathbf{q};k'j')\right.\\
&\left.+V^{*\alpha}(\mathbf{q};kj)\Phi_{wv}^{\gamma}(\mathbf{q};kk')W^{\gamma\beta}(\mathbf{q};k'j')\right]\Bigg\},
\end{aligned}
\tag{4}
$$

where N is the number of atoms per unit cell, $\mathbf{x}_j$ are the coordinates of atom j inside the unit cell, and $Q^{\alpha\beta}(\mathbf{q};jj')$ is the contribution to the dynamic matrix from long-range Coulomb interactions. The matrices entering into Eq. (4) have the form

$$\hat{V}=-\hat{\Phi}_{vv}^{-1}\hat{P}_{vR},\quad \hat{W}=\hat{R}_{ww}^{-1}\hat{\Phi}_{ww}^{-1}\hat{S}_{wR},$$

$$\hat{R}_{ww}=1-\hat{\Phi}_{ww}^{-1}\hat{\Phi}_{wv}\hat{\Phi}_{vv}^{-1},\quad \hat{S}_{wR}=\hat{\Phi}_{wv}\hat{\Phi}_{vv}^{-1}\hat{\Phi}_{vR}-\hat{\Phi}_{wR},$$

$$\hat{P}_{vR}=\hat{\Phi}_{vR}+\hat{\Phi}_{vw}\hat{R}_{ww}^{-1}\hat{\Phi}_{ww}^{-1}\hat{S}_{wR}.$$

The matrix $\hat{\Phi}$ is defined as

$$\Phi_{RR}^{\alpha\beta}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial^2\Phi\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial R_{\alpha}\begin{pmatrix}\mathbf{l}\\ j\end{pmatrix}\partial R_{\beta}\begin{pmatrix}0\\ j'\end{pmatrix}}\exp(-i\mathbf{ql}),$$

$$\Phi_{vv}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial^2\Phi\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial v_j\partial v_{j'}}\exp(-i\mathbf{ql}),$$

$$\Phi_{ww}^{\alpha\beta}(\mathbf{q};jj')=Q^{\alpha\beta}(\mathbf{q};jj')+\Gamma^{\alpha\beta}(\mathbf{q};jj')+\frac{\delta_{jj'}\delta_{\alpha\beta}}{\alpha_j},\tag{5}$$

$\Gamma^{\alpha\beta}(\mathbf{q})$ is the matrix of the short-range part of the dipole–dipole interaction, α_j is the polarizability of the jth ion,

TABLE II. Equilibrium values of the lattice parameters, the polarizabilities of the ions, and the rf permittivities.

Crystal	Model	a_0, Å		Polarizability, Å^3					ε_∞	
		Calc.	Exp.	$\alpha_{K'}$	α_K	α_{Na}	α_{Al}	α_F	Calc.	Exp.
K_2NaAlF_6	I	8.12	8.11(Ref. 9)							
	II	8.12			0.696	0.122	0.034	1.123	2.27	1.79(Ref. 9)
	III	7.94								
	IV	7.94			0.836	0.122	0.034	0.720	1.80	
K_3AlF_6	I	8.20	8.38(Ref. 9)							
	II	8.20		0.696	0.696		0.034	1.123	2.23	1.80(Ref. 9)
	III	8.12								
	IV	8.12		0.726	0.836		0.034	0.749	1.86	
Na_3AlF_6	I	8.09	7.95(Ref. 10)							
	II	8.09				0.122	0.034	1.123	2.05	1.78(Ref. 9)
	III	7.86								
	IV	7.86				0.122	0.034	0.720	1.61	

$$\Phi^{\alpha}_{Rv}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial^2\Phi\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial R_\alpha\begin{pmatrix}\mathbf{l}\\ j\end{pmatrix}\partial v_{j'}}\exp(-i\mathbf{ql}),$$

$$\Phi^{\beta}_{vR}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial^2\Phi\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial v_j\partial R_\beta\begin{pmatrix}0\\ j'\end{pmatrix}}\exp(-i\mathbf{ql}),$$

$$\Phi^{\alpha\beta}_{wR}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial E^{\mathrm{sh}}_{\alpha}\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial R_\beta\begin{pmatrix}0\\ j'\end{pmatrix}}\exp(-i\mathbf{ql}),$$

$$\Phi^{\alpha}_{wv}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial E^{\mathrm{sh}}_{\alpha}\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial v_{j'}}\exp(-i\mathbf{ql}),$$

$$\Phi^{\beta}_{vw}(\mathbf{q};jj')=\sum_{\mathbf{l}}\frac{\partial E^{\mathrm{sh}}_{\beta}\begin{pmatrix}\mathbf{l}\,0\\ jj'\end{pmatrix}}{\partial v_{j}}\exp(-i\mathbf{ql}),$$

$$\hat{\Phi}_{Rw}=\hat{\Phi}^{+}_{wR}, \tag{6}$$

and $\mathbf{E}^{\mathrm{sh}}$ is the short-range crystal field created at the jth ion. The expression for the rf permittivity ε_∞ can be written

$$\varepsilon^{\alpha\beta}_{\infty}=\delta_{\alpha\beta}+\frac{4\pi q_\alpha}{q^2}\sum_{\gamma=1}^{3}\sum_{k,k'=1}^{N}q_\gamma[\Phi^{-1}_{ww}]^{\gamma\beta}(0;kk').$$

The Coulomb contribution to the dynamic matrix $Q^{\alpha\beta}(\mathbf{q};jj')$ was calculated by the Ewald method. The calculation for the ion was carried out according to Liberman's program,[8] and the energy of the pairwise interaction from Eq. (3) and the polarizability of the ion were calculated according to Ivanov and Maksimov's program,[2] using the Thomas–Fermi approximation for the kinetic energy and the Hedin–Lundquist approximation for the exchange energy. The technique of approximating the dependences of the energy on the distance $\mathbf{R}$ and the potentials v of the Watson sphere was used to compute the partial derivatives in Eqs. (5) and (6), entering into the dynamic matrix of Eq. (4). Chebyshev polynomials were used for the approximations.[2]

4. RESULTS AND DISCUSSION

This section presents the results of calculations of the total energy, the equilibrium volume, and the lattice vibration spectra for three crystals and four models. The calculations in Model I use the electron density of free spherically symmetrical ions (the rigid-ion model). Model II takes into account the polarizability of the ions. In Model III, the effect of the crystal environment is taken into account by using the potential of the Watson sphere when calculating the electron density of the ions. For simplicity, we used the Watson-sphere potential for only two types of ions in the crystals under discussion: for the K^+ ion and the F^- ion. As shown by our estimates for Al^{3+} and Na^+ ions, the electron density of the free ions is virtually the same as the electron density of these ions in the Watson sphere. Finally, Model IV takes into account the deformability and polarizability of the ions.

TABLE III. Limiting frequencies of the ($q=0$) vibrations of K_2NaAlF_6.

ω_i (cm^{-1})	Degeneracy	Type of vibration	Models			
			I	II	III	IV
ω_1^L	1	T_{1u}	537.9	435.7	558.5	478.6
ω_2^T	2	T_{1u}	399.3	380.2	427.5	403.2
ω_3	1	A_{1g}	394.9	268.9	456.4	386.6
ω_4^L	1	T_{1u}	299.5	194.9	359.4	270.5
ω_5^T	2	T_{1u}	279.1	194.8	356.6	265.5
ω_6	2	E_g	227.5	226.8	268.0	261.4
ω_7^L	1	T_{1u}	197.9	176.0	213.5	202.5
ω_8	3	T_{2g}	264.7	148.0	308.6	202.1
ω_9^T	2	T_{1u}	150.3	146.0	178.8	173.6
ω_{10}	3	T_{2u}	117.8	96.3	166.0	146.8
ω_{11}^L	1	T_{1u}	124.0	102.9	120.8	108.3
ω_{12}	3	T_{1g}	37.7i	37.7i	30.2	25.1
ω_{13}	3	T_{1u}	0.0	0.0	0.0	0.0
ω_{14}	3	T_{2g}	88.8	73.5i	70.7	12.1i
ω_{15}^T	2	T_{1u}	87.8	46.7i	61.1	47.0i

TABLE IV. Limiting frequencies of the $(q=0)$ vibrations of K_3AlF_6.

ω_i (cm^{-1})	Degeneracy	Type of vibration	Models I	II	III	IV
ω_1^L	1	T_{1u}	489.9	385.0	517.8	432.8
ω_2^T	2	T_{1u}	297.8	381.3	329.4	349.5
ω_3	1	A_{1g}	377.5	268.6	432.9	361.4
ω_4^L	1	T_{1u}	314.8	247.9	352.2	269.8
ω_5^T	2	T_{1u}	305.4	241.5	350.1	269.7
ω_6	2	E_g	229.4	229.1	260.6	257.7
ω_7^L	1	T_{1u}	175.5	143.8	188.9	172.3
ω_8	3	T_{2g}	249.0	150.6	287.6	193.8
ω_9^T	2	T_{1u}	166.6	143.8	171.3	162.4
ω_{10}	3	T_{2u}	109.9	93.1	148.4	132.8
ω_{11}^L	1	T_{1u}	117.2	74.8	112.5	94.2
ω_{12}	3	T_{1g}	31.6i	31.6i	24.7i	26.6i
ω_{13}	3	T_{1u}	0.0	0.0	0.0	0.0
ω_{14}	3	T_{2g}	66.2	85.1i	53.5	42.3i
ω_{15}^T	2	T_{1u}	58.8	69.6i	38.8	58.2i

TABLE V. Limiting frequencies of the $(q=0)$ vibrations of Na_3AlF_6.

ω_i (cm^{-1})	Degeneracy	Type of vibration	Models I	II	III	IV
ω_1^L	1	T_{1u}	487.9	380.1	579.3	513.5
ω_2^T	2	T_{1u}	349.5	305.3	454.1	435.9
ω_3	1	A_{1g}	335.6	215.6	480.6	400.8
ω_4^L	1	T_{1u}	257.7	191.9	372.9	280.2
ω_5^T	2	T_{1u}	248.4	190.0	366.0	274.0
ω_6	2	E_g	151.0	149.0	294.1	291.0
ω_7^L	1	T_{1u}	171.5	154.5	221.6	210.4
ω_8	3	T_{2g}	258.3	170.3	314.5	227.9
ω_9^T	2	T_{1u}	123.2	114.4	188.7	184.3
ω_{10}	3	T_{2u}	98.8	85.2	147.4	135.6
ω_{11}^L	1	T_{1u}	70.3	34.0	116.7	101.9
ω_{12}	3	T_{1g}	57.6i	58.5i	81.9i	82.1i
ω_{13}	3	T_{1u}	0.0	0.0	0.0	0.0
ω_{14}	3	T_{2g}	81.6i	95.3i	80.9i	64.4i
ω_{15}^T	2	T_{1u}	90.7i	96.4i	58.1i	86.4i

The results of the calculations are shown in Figs. 2–5 and in Tables II–V. The equilibrium values of the lattice parameters were determined by minimizing the total energy of the crystal as a function of volume (Fig. 2). The lattice parameters are shown in Table II along with the experimental values. For all three materials, the calculated lattice parameters agree with the experimental data to within 2%. The radii of the Watson spheres for the K^+ and F^- ions, found by minimizing the total energy, are 2.0 au and 2.2–2.3 au, respectively.

Table II shows the calculated polarizabilities of the ions and the rf permittivities of the materials under consideration. This table also shows the experimental values of ε_∞. As can be seen from the table, the calculated polarizabilities of the fluorine ions are substantially different in the free-ion approximation (and taking into account the crystal environment within the Watson sphere), and this in turn results in a difference in the calculated rf permittivities for all three materials.

The calculated dispersion curves of the frequencies of the lattice vibrations for the three compounds are shown in Figs. 3–5. In order not to clutter the figures, we show in them the calculated results only for Model IV, since the $\omega(\mathbf{q})$ dependences are qualitatively the same for all four models, while the quantitative differences in the frequencies of the lattice vibrations calculated in Models I–IV are shown in Tables III–V, which display the limiting frequencies of the $(q=0)$ vibrations. As can be seen from Figs. 3–5 and Tables II–IV, there are imaginary frequencies of the lattice vibrations in all compounds under discussion; this is evidence of structural instability of the cubic phase in these materials. It should be emphasized that the unstable modes occupy all the phase space in the Brillouin zone. In the K_3AlF_6 and Na_3AlF_6 crystals, there is instability of the structure in all four models. In the K_2NaAlF_6 crystal, the cubic phase is stable at zero temperature only in the model of the deformed ion that neglects polarizability. As can be seen from Tables III–V, taking the polarizability of the ions into account in all the compounds under consideration reduces almost all the frequencies of the lattice vibrations, increases the number of unstable modes, and appreciably reduces the splitting of the longitudinal and transverse vibrational frequencies of the polar modes.

It can be seen from Figs. 3–5 and Tables III–V that the cubic phase in the compounds under consideration is most unstable in the Na_3AlF_6 crystal and most stable in K_2NaAlF_6. This conclusion qualitatively agrees with the results of experimental studies of structural phase transitions in these crystals.[9] It has been established that the phase-transition temperature in Na_3AlF_6 significantly exceeds the transition temperature in K_3AlF_6, while no phase transitions are detected in the K_2NaAlF_6 crystal up to liquid-nitrogen temperatures.

There are three types of instability of the cubic structure at the center of the Brillouin zone. One is the ferroelectric instability associated with transverse vibrations of the polar mode T_{1u}. In this mode, all the atoms in a unit cell are displaced from the equilibrium positions of the cubic phase. Ferroelectric phase transitions, as far as we know, have not been experimentally observed in halide crystals with the elpasolite structure. Another instability is associated with the triply degenerate T_{1g} mode.

Only the four fluorine atoms are displaced from the equilibrium positions in this mode, and these displacements cause the AlF_6 octahedron to rotate as a whole (see Table I). Finally, a third type of instability is associated with the triply degenerate T_{2g} mode. In one of the eigenvectors of this mode, the displacements of the atoms cause the AlF_6 octahedron to rotate about the body diagonal while the potassium (sodium) atoms located on that diagonal are simultaneously displaced toward each other. Note that there is another, stable mode with the same T_{2g} symmetry in the vibrational spectrum of the crystals under consideration (see Tables III–V).

5. CONCLUSION

The static and dynamic properties of three crystals with the elpasolite structure have thus been calculated in this paper in terms of a simple nonempirical model of an ionic crystal. The calculated equilibrium values of the lattice parameters and the permittivity are in good agreement with the experimental data. Unfortunately, we cannot compare the calculated frequencies of the lattice vibrations with measured results, since such measurements have apparently not been made for the crystals considered here. Our results concerning the instability of the cubic structure and the presence of unstable modes in a large phase space of the Brillouin zone are apparently common to crystals with the given structure.

The authors are grateful to the Russian Fund for Fundamental Research (Projects 97-02-16277 and 96-15-96700) for financial support of this work. We are grateful to O. V. Ivanov and E. G. Maksimov for allowing us to use their programs to calculate the total energy and polarizability of the ions.

*)E-mail: zinenko@ph.krasnoyarsk.su

[1] W. Bührer and H. U. Güdel, J. Phys. C **20**, 3809 (1987).
[2] O. V. Ivanov and E. G. Maksimov, Zh. Éksp. Teor. Fiz. **108**, 1841 (1995) [JETP **81**, 1008 (1995)] .
[3] M. Couzi, S. Khairoun, and A. Tressand, Phys. Status Solidi A **98**, 423 (1986).
[4] A. A. Maradudin and V. Vosko, Rev. Mod. Phys. **40**, 1 (1968).
[5] O. V. Kovalev, *Irreducible Representations of the Space Groups*, Gordon & Breach, New York (1965).
[6] V. Nüsslein and U. Schröder, Phys. Status Solidi **21**, 309 (1967).
[7] A. Chizmeshya, F. M. Zimmermann, R. A. LaViolette, and G. H. Wolf, Phys. Rev. **50**, 15 559 (1994).
[8] D. A. Liberman, D. T. Cromer, and J. J. Waber, Comput. Phys. Commun. **2**, 107 (1971).
[9] *Minerals. A Collection*, vol. 2, Akad. Nauk SSSR, Moscow (1963).
[10] H. Bode and E. Voss, Z. Anorg. Allg. Chem. **200**, 1 (1957).

Translated by W. J. Manthey

Low-temperature thermal conductivity of highly enriched and natural germanium

A. P. Zhernov*) and D. A. Zhernov

Kurchatov Institute, 123182 Moscow, Russia
(Submitted 10 February 1998)
Zh. Éksp. Teor. Fiz. **114**, 1757–1764 (November 1998)

Experimental data on the thermal conductivity $K(T)$ of crystals of natural and highly enriched germanium (99.99%) ^{70}Ge with lapped and polished surfaces are analyzed in the temperature range $\sim 1.5-8$ K. In all the samples in the temperature range $\sim 1.5-4$ K the standard boundary mechanism of scattering dominates. As the temperature is raised, an isotopic scattering mechanism is observed in the natural samples. In the highly enriched samples the theoretical values of $K(T)$ turn out to be much smaller than the experimental ones. It is conjectured that a Poiseuille viscous flow regime of the phonon gas emerges in this case. © *1998 American Institute of Physics.* [S1063-7761(98)01411-5]

1. INTRODUCTION

Chemically pure, perfect crystals of the ^{70}Ge isotope of germanium with 99.99% enrichment have recently been synthesized by the group led by V. I. Ozhogin. A study of various of their properties has commenced. An experimental study of their thermal conductivity $K(T)$ over a wide temperature range has already been carried out on samples of varied composition, including both highly enriched ^{70}Ge and natural germanium.[1,2] The corresponding data were obtained for crystals with a finely polished surface and with a surface processed relatively more crudely by lapping (for details, see Ref. 2). As a result, we possess a unique body of experimental data for more detailed examination of kinetic processes in regular systems, and of their role in the kinetics of isotopic disorder.

In the present paper we analyze experimental data[1,2] obtained for a range of helium temperatures from 1.5 to 8 K. First we consider the effect of the degree of surface processing and the possible role of strong dispersion of the acoustic phonon spectrum of germanium in the natural samples. Second, for the case of highly enriched perfect crystals, we discuss the possible emergence of an effect predicted by Gurzhi.[3] In essence, the migration of quasiparticles taking place in the phonon flux can under certain conditions be considered a random walk, analogous to Poiseuille viscous flow of a liquid.[3,4] At very low temperatures, the relaxation rate for the acoustic phonon mode with frequency ω and polarization j is then exponentially small, due to anharmonic Umklapp (U) processes in which short-wavelength phonons with large momenta (comparable to the reciprocal lattice vector) invariably participate; specifically,

$$\frac{1}{\tau_U^{(j)}} \propto \omega^2 T \exp\left(-\frac{B_j}{T}\right).$$

At the same time, the relaxation rates for the longitudinal (l) and transverse (t) modes due to normal (N) processes in which any phonons can participate, including only long-wavelength ones, are described by power-law temperature dependences of the form

$$\frac{1}{\tau_N^{(t)}} \propto \omega T^4, \quad \frac{1}{\tau_N^{(l)}} \propto \omega^2 T^3$$

(see, e.g., Refs. 4 and 5). Therefore, for the corresponding effective mean free paths a temperature range can exist in which $l_N \ll d \ll l_U$, where d is the characteristic linear dimension of the sample. In that range, for extremely high-quality single-isotope crystals, the effect of static defects and isotopic disorder on the structure of the nonequilibrium phonon distribution function is obscured by N processes. Only scattering processes involving the walls of the sample are then resistive. However, due to the more frequent nonresistive N processes the transport mean free path grows as $\sim d^2/l_N$ (Ref. 3, see also Refs. 4–6).

Mezhov-Deglin[7,8] investigated the thermal conductivity of very perfect, single-isotope samples of solid ^{4}He and explicitly observed the features predicted by Gurzhi[3] in the temperature behavior of the thermal conductivity $K(T)$ in the region to the left of the maximum (see also Ref. 9). Moreover, manifestations of the hydrodynamic regime have been analyzed in NaF and Bi crystals, which lacks the isotopic scattering mechanism. The manifestation of the cited regime in NaF and Bi is less pronounced in comparison with solid helium. The corresponding results are discussed, for example, in Ref. 5.

2. BASIC RELATIONS

We restrict the discussion to extremely low temperatures. In this situation the relaxation time τ_j of the phonon mode with polarization index j and group velocity v_j is due mainly to boundary scattering. For diffuse boundary scattering and an infinitely long sample, $\tau_j^{(b,c)} = l_C / v_j$, where l_C is the mean free path of the phonon mode (or the Casimir length[10]). For samples with rectangular cross section S the Casimir length is $l_C = 1.12\sqrt{S}$.

We consider the effect of fine surface polishing, and also of lapping (somewhat cruder processing) on the temperature dependence of the thermal conductivity in the context of the

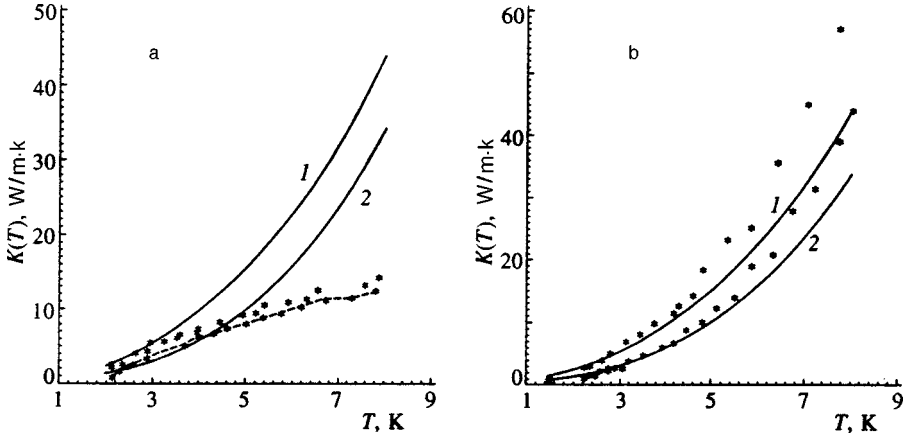


FIG. 1. Temperature dependence of the thermal conductivity for natural (a) and highly enriched samples (b). Theoretical curves *1* and *2* correspond to samples with polished and lapped surfaces, respectively. The specularity parameter $\eta=36$ (curve *1*), 65 (curve *2*). The experimental points are taken from Ref. 2 (for the lapped surface they are joined by a dotted line in Fig. 1a).

Ziman–Soffer theory,[11,12] which takes account of both diffuse and specular phonon scattering at the sample boundaries. The relaxation time

$$\tau_j^{(b)}=\frac{l_C}{v_j}\frac{1+P(k_j,\phi)}{1-P(k_j,\phi)},\quad P(k_j,\phi)=\exp[-(2k_j\eta\cos\phi)^2] \tag{1}$$

figures in this theory. Here P is the specularity factor, which for the jth mode depends on the magnitude of the phonon wave vector k_j and its orientation, i.e., the angle ϕ. The roughness factor η characterizes the degree of surface polish. Note that $k_j=\omega/v_j$ and $\phi=\pi/2-\theta$, where θ is the angle formed by the phonon wave vector and the unit vector in the direction of the temperature gradient.

It can be shown that the lattice thermal conductivity K can be written (see Ref. 13)

$$K(T)=\sum_{j=1}^{3}\frac{k_Bl_C}{2\pi^2v_j^2}\left(\frac{k_BT}{\hbar}\right)^3\times\int_0^{z_{Dj}}dz\,\frac{z^4e^z}{(e^z-1)^2}\int_0^1 dy\,y^2\coth Z_j(y). \tag{2}$$

Here $z_{Dj}=\Theta_j/T$ is the reduced Debye temperature. The factor Z_j is

$$Z_j(y)=\begin{cases}\dfrac{4k_B^2}{2v_j^2\hbar^2}\eta^2z^2T^2(1-y^2), & \coth Z_j(y)\leqslant r,\\ ry^2, & \coth Z_j(y)>r.\end{cases} \tag{3}$$

Here $y=\cos\theta$, and the parameter $r=l_{\max}/l_C$, where $l_{\max}$ is the linear extent of the sample (its length). In what follows, the Boltzmann constant k_B and Planck constant $\hbar$ are set to unity.

Note that the value of Z_j is sensitive not only to the degree of surface polish, i.e., to η, but also depends on the temperature T. Note also that if $y\approx1$, i.e., the phonons ''migrate'' parallel to the axis of the sample, then the integral over y diverges. In actuality, the length of the sample is finite. This is taken into account by the Ziman–Soffer theory. When $F(y)=\coth Z_j(y)$ exceeds r, $F(y)$ can be replaced by ry^2.

3. RESULTS AND DISCUSSION

In the present work we use (1)–(3) to analyze low-temperature experimental data in the temperature range $T\approx1.5-8$ K on the thermal conductivity of natural and isotopically enriched samples (with enrichment of 99.99%). As noted above, these were prepared both with a finely polished surface and with a more crudely lapped surface.[1,2] Thus, there are four sets of experimental curves.

The sole fitting parameter of the theory is η. The isotopic disorder parameter

$$g=\frac{\sum c_iM_i^2-\left(\sum c_iM_i\right)^2}{\left(\sum c_iM_i\right)^2},$$

where c_i and M_i are the concentration and mass of the isotope of type i, is 5.87×10^{-4} and 8.18×10^{-8}, respectively, for the natural and highly enriched crystals. For the group velocities of the transverse and longitudinal acoustic phonon modes we adopted standard values, $v_t=3.16\times10^5$ cm/s and $v_l=5.21\times10^5$ cm/s. To determine the Casimir length l_C and the parameter r, we used the actual geometrical dimensions of the samples (see Refs. 1 and 2). Note that their cross sections were rectangular. The corresponding data are listed in Table I.

Calculated theoretical curves for two values of the parameter η, which characterizes the degree of surface polish, are plotted in Fig. 1 along with experimental data in the form of points for all four samples.

Some comments on the figure are in order. It can be seen at once that the factor η decreases substantially as the degree of surface polish rises. In fine polishing, both the natural and the highly enriched samples correspond to $\eta\approx36$. For samples with lapped surfaces, the values of η are also similar, but $\eta\approx65-75$ (see Table I).

Note that for the natural samples there is reasonable agreement between theory and experiment in the temperature range $\sim1.5-4$ K. For $T\geqslant4$ K the theoretical curves $K(T)$ lie substantially higher than the experimental points (also see Refs. 13–15). For the highly enriched samples the agreement

TABLE I. Geometrical dimensions l_x, l_y, l_z and values of the parameter η of the Ziman–Soffer theory.

Sample	l_x, l_y, l_z, mm	η, Å
$n(s)$	2.33×2.3×40.7	65
$n(p)$	2.4×2.35×40.8	36
$h(s)$	2.2×2.5×40.4	65–75
$h(p)$	2.44×2.13×40.4	36

*Remark.*The isotopic disorder parameter g for natural (n) and highly enriched (h) samples is 5.87×10^{-4} and 8.18×10^{-8}, respectively. The notation s and p denote lapped and polished samples, respectively.

with theory is also reasonable in the range $T\approx1.5-4$ K. But for $T\geqslant4$ K the calculated values of $K(T)$ are considerably lower than the experimental values.

We now comment on these results. In the simplest approximation

$$K(T)=\frac{1}{3}C_L(T)v_a^2\tau_s, \quad \tau_s^{-1}=\sum_i (\tau^{(i)})^{-1}. \tag{4}$$

Here C_L is the phonon thermal conductivity and v_a is the mean phonon velocity. The symbol τ_s denotes the total relaxation time, whose reciprocal is the sum of the reciprocal relaxation times of the various processes.

According to (4), for the natural samples

$$K(T)\approx\frac{1}{3}C_L(T)v_a^2\tau^{(b,c)}\left(1-\frac{\tau^{(b,c)}}{\tau^{(is)}}\right),$$

where $\tau^{(is)}$ is the relaxation time due to isotopic disorder. Numerical estimates show that the ''negative'' deviation of $K(T)$ for $T\geqslant4$ K is due to neglect of isotopic phonon scattering under conditions in which soft transverse modes are present (see also Refs. 5 and 6).

For the highly enriched samples $\tau_s=\tau^{(b,c)}$. At the same time, for germanium the effective Debye temperature $\Theta(T)$ (due to soft t modes) has a pronounced minimum at 25 K (see, e.g., Refs. 16 and 17). Since

$$C_L(T)\propto\Theta^{-3}(T), \tag{5}$$

by taking relations (4) and (5) into account it is possible to explain qualitatively the specific behavior of $K(T)$—the ''positive'' deviation for $T>4$ K as a result of a manifestation of soft t modes. However, specific estimates using the results of Refs. 16 and 17 do not afford even a crude description of the observed deviation.

As noted in the Introduction, the hydrodynamic viscous flow regime for phonons is treated in the literature. If this regime is in fact realized, it leads to substantial growth of the thermal conductivity on the left side of the temperature maximum. At present, evidence for this regime has been explicitly observed in crystals of solid helium (see Refs. 7 and 8, and also Ref. 9). Specifically, at the lowest temperatures $T<0.6$ K the phonon mean free path in ^{4}He due to anharmonic collisions turns out to be less than the diameter of the sample. In the temperature range 0.6–1 K we have $l_N\ll d$, $l_Rl_N\gg d^2$, where l_N and l_R are the mean free paths corresponding to normal and resistive processes. This then yields $l_{\rm eff}=0.1\,d^2/l_N$. In this situation, according to experiment and simple theoretical estimates, $K(T)\propto C_Ll_N\propto T^8$ to the left of the maximum.

Note also that Ref. 18 established criteria under which the motion of a phonon gas subject to an applied temperature gradient can be considered Poiseuille viscous flow:

$$l_R/l_N\gg10^3, \quad d/l_N\gg30. \tag{6}$$

Thus, it should be possible qualitatively to explain the distinctive ''positive'' deviation of the experimental values of $K(T)$ from the theoretical values for $T>4$ K in ^{70}Ge as a manifestation of viscous Poiseuille flow in perfect, highly enriched samples. Note that the degree of surface polish has a significant influence on the thermal conductivity. Numerical estimates require a knowledge of the mean free paths due to the anharmonic processes. Moreover, it is necessary to take into account effects associated with phonon focusing.[18]

To find the mean free paths, we adduce results of Callaway's theory of the thermal conductivity of a crystal lattice (see, e.g., Ref. 5). This theory takes account of the contribution of N and U processes to the thermal conductivity. Since the phonon spectrum of germanium is substantially anisotropic, we consider contributions to the thermal conductivity in the form of a sum of contributions from t and l modes. We have chosen τ_N and τ_U in the standard form

$$\frac{1}{\tau_N^{(t)}}=A_N^{(t)}\omega T^4, \quad \frac{1}{\tau_N^{(l)}}=A_N^{(l)}\omega^2T^3,$$

$$\frac{1}{\tau_U^{(t,l)}}=A_U^{(t,l)}\,\omega^2\,T\exp\left(-\frac{B^{(t,l)}}{T}\right).$$

As a result of estimates and a comparison of theoretical and experimental results, the authors of Ref. 2 find that

$$A_N^{(t)}=2\times10^{-13}\,\mathrm{K}^{-4}, \quad A_N^{(l)}=2\times10^{-21}\,\mathrm{s\cdot K}^{-3},$$

$$A_U^{(t)}=1\times10^{-19}\,\mathrm{K}^{-4}, \quad A_U^{(l)}=5\times10^{-19}\,\mathrm{s\cdot K}^{-3},$$

$$B^{(t)}=55\,\mathrm{K}, \quad B^{(l)}=180\,\mathrm{K}. \tag{7}$$

It is noteworthy that in the model under consideration, the thermal conductivity of germanium is a superposition of two peaks corresponding to t and l modes. Interestingly enough, for a highly enriched sample (99.99%) the main peak in $K(T)$, right at the maximum at $T_m\approx16.5$ K, turns out to be associated with l modes. For a sample with 96% enrichment, on the other hand, the dominant contribution to $K(T)$ at the maximum is due to t modes. The influence of l modes is partly veiled due to isotopic scattering. This explains the shift of the maximum by a few degrees (~4 K) toward lower temperatures for the sample with 96% enrichment in comparison with the curves for the highly enriched (99.99%) sample. For the natural sample, under conditions of strong isotopic scattering, the role of l modes is somewhat enhanced relative to t modes, and the maximum is raised by a few fractions of a degree.

We now determine the mean free path l_i corresponding to the phonon mode of i type ($i=l,t$). We have

$$\frac{1}{l_i} = \int d\mathbf{q} \frac{C_\mathbf{q}}{v_\mathbf{q} l_\mathbf{q}^{(i)}} \int d\mathbf{q} \frac{C_\mathbf{q}}{v_\mathbf{q}}. \tag{8}$$

Here $C_\mathbf{q}$ is the partial contribution to the lattice specific heat from the $\mathbf{q}$ modes, and $v_\mathbf{q}$ is their group velocity.

In the temperature range under consideration with the parameters (7) and using (8), it turns out not to be the case that $l_N \ll d \ll l_U$ for transverse modes. For longitudinal modes, the situation is the following. It can be shown that $l_N^{(l)} \ll l_U^{(l)}$, whereupon the mean free paths differ substantially. At the same time we have

$$\frac{l_C}{l_N} \approx \frac{5 \cdot 10^8 T^5}{T_{Dl}^5},$$

where $T_{Dl} = 330$ K is the Debye temperature for longitudinal modes. Thus, at $T > 4$ K, l_C is greater than l_N, but of the same order. Generally speaking, condition (6) is not satisfied.

Thus, in the present case, there is no developed hydrodynamic regime, but effects associated with it should be observed to a certain extent in longitudinal modes.

We emphasize that when the surface of the samples is polished, l_C grows. In the simplest case the resultant change in the relaxation time can be written[10]

$$\frac{1}{\tau_b} = \frac{v_a}{l_C} \frac{1-P}{1+P}.$$

P, which characterizes the fraction of quasiparticles specularly reflected from the boundaries, ranges from 0 to 1. According to Ref. 2, as one goes from ''lapped'' to ''polished'' Ge samples, P rises from 0.2 to ~ 0.4.

We emphasize that as the specularity of the surface (characterized by the parameter P) and consequently the difference between l_N and l_C decreases, the deviation of the experimental values of $K(T)$ from the theoretically calculated values increases (see Fig. 1b).

To summarize, in the temperature range 1.5–4 K, the change in the behavior of the thermal conductivity of germanium as a function of the degree of specularity can be described qualitatively by the Ziman–Soffer theory. In perfect, highly enriched samples at $T \geqslant 4$ K, when most collisions are with the walls, the phonon flux is observed to deviate from Knudsen flow. In natural samples of Ge, isotopic scattering emerges at $T \geqslant 4$ K in a substantial way.

Note that under conditions of Poiseuille flow, the existence of second sound becomes possible, that is to say, oscillations of the density of thermal excitations. A similar effect is observed in helium,[9] and probably in NaF and Bi as well.[5] It would be interesting to investigate second sound in perfect germanium crystals.

This work was suggested by V. I. Ozhogin and carried out with the support of N. A. Chernoplekov.

*)E-mail: zhernov@kurm.polyn.kiae.su

[1] V. I. Ozhogin, A. V. Inyushkin, A. N. Toldenkov, G. É. Popov, E. E. Haller, and K. M. Ito, JETP Lett. **63**, 490 (1996).
[2] M. Asen-Palmer, K. Bartkowsky, E. Gmelin, M. Cardona, A. P. Zhernov, A. V. Inyushkin, A. V. Taldenkov, V. I. Ozhogin, K. M. Itoh, and E. E. Haller, Phys. Rev. B **56**, 9431 (1997).
[3] R. N. Gurzhi, Zh. Éksp. Teor. Fiz. **46**, 719 (1965) [Sov. Phys. JETP **19**, 490 (1964)]; Usp. Fiz. Nauk **94**, 689 (1968) [Sov. Phys. Usp. **11**, 255 (1968)].
[4] E. M. Lifshitz and L. P. Pitaevskiĭ, *Physical Kinetics*, Pergamon Press, Oxford (1981).
[5] R. Berman, *Thermal Conduction in Solids*, Clarendon Press, Oxford (1976).
[6] V. L. Gurevich, *Kinetics of Phonon Systems* [in Russian], Nauka, Moscow (1980).
[7] L. P. Mezhov-Deglin, Zh. Éksp. Teor. Fiz. **49**, 66 (1965) [Sov. Phys. JETP **22**, 47 (1966)]; Zh. Éksp. Teor. Fiz. **71**, 1453 (1976) [Sov. Phys. JETP **44**, 761 (1976)].
[8] L. P. Mezhov-Deglin, Fiz. Tverd. Tela **22**, 1748 (1980) [Sov. Phys. Solid State **22**, 1018 (1980)].
[9] B. N. Esel'son, V. N. Grigor'ev, V. G. Ivantsov, É. Ya. Rudakovskiĭ, D. G. Sanikidze, and I. A. Serbin, *Solutions of Quantum Liquids* [in Russian] Nauka, Moscow (1973).
[10] H. B. Casimir, Physica (Amsterdam) **5**, 495 (1938).
[11] J. M. Ziman, *Electrons and Phonons*, Clarendon Press, Oxford (1960), p. 456.
[12] S. B. Soffer, J. Appl. Phys. **38**, 1710 (1967).
[13] D. P. Singh and Y. P. Foshi, Phys. Rev. B **19**, 3133 (1979).
[14] D. R. Frankl and G. J. Campisi, in *Proceedings of the International Conference on Phonon Scattering in Solids*, Paris, 1972, H. J. Albany (ed.), La Documentation Française, Paris (1972), p. 88.
[15] W. S. Hust and D. R. Frankl, Phys. Rev. **186**, 801 (1969).
[16] A. D. Zdetsis and C. S. Wang, Phys. Rev. B **19**, 2999 (1979).
[17] Resul Eryiğit and Irving P. Herman, Phys. Rev. B **53**, 7775 (1996).
[18] A. K. McCurdy, H. J. Maris, and C. Elbaum, Phys. Rev. B **2**, 4077 (1970).

Translated by Paul F. Schippnick

Spin fluctuations and the superconducting state in doped insulators

A. I. Agafonov and É. A. Manykin*)

Institute of Superconductivity and Solid State Physics, Kurchatov Institute, 123182 Moscow, Russia
(Submitted 23 July 1998)
Zh. Éksp. Teor. Fiz. **114**, 1765–1784 (November 1998)

We propose a model of electron pairing via spin fluctuations in doped insulators. The bare states for the superconducting condensate correspond to impurity bands in the original band gap of the undoped material. We obtain a complete set of equations for the superconducting state. We show that fermion pairing in impurity bands of extended states is possible, and thus so is superconductivity, if localized spin-0 bosons are produced. The latter are necessarily accompanied by localized spin-1 bosons, which are responsible for the relationship between singlet and triplet pairing channels of quasiparticles. © *1998 American Institute of Physics.* [S1063-7761(98)01511-X]

1. INTRODUCTION

Ever since the discovery of high-temperature superconductivity, it has been the subject of an enormous amount of experimental and theoretical research. Nevertheless, neither the pairing mechanism in the superconducting state nor the origin of charge carriers in the normal state of doped cuprates have been elucidated. It has also remained unclear whether the mechanism of superconductivity in 3D bismuth-based oxides ($Ba_{1-x}K_xBiO_{3-y}$ and others) is identical to that in layered copper-based oxides (p-doped materials like $La_{2-x}Sr_xCuO_4$, $YBa_2Cu_3O_{7-x}$, etc., and n-doped $Nd_{2-x}Ce_xO_{4-x}$).

The parent cuprates usually have an antiferromagnetic insulating state. The magnetic phase disappears[1,2] at low doping levels. When the doping level is increased, the material goes to a relatively low-conductivity metallic phase with a high T_c. Previous research shows that HTSC materials belong to the class of doped insulators, and impurities exert a dominant influence on phase transitions.[2,3]

The cuprates have unusual anisotropic properties in the insulating, metallic, and superconducting phases.[1–4] Their experimental behavior stimulated the development of various theoretical approaches to describing the properties of these materials. These approaches fall into two groups.

After it was shown that carrier mobilities in the cuprates are comparable to the Mott–Ioffe–Regel limit ($ea^2/\hbar$), it became clear that localization effects are important in these systems.[5] The first group of approaches is based on the fact that, as follows from experimental data, the Fermi level is located among the electron states produced by doping in the charge transfer gap of the material, for instance in the gap due to charge transfer between O2p- and Cu 3d-bands in cuprates.[1,4,6–11] These impurity states, known as *in-gap states*, form in all doped compounds.[1] Typical spectral features have been observed both in the optical conductivity[12,13] and in electron spectroscopy.[8–11,14] Measurements obtained by angle-resolved photoelectron spectroscopy (ARPES) have established that both localized and extended states coexist at the Fermi level, and can contribute to the formation of a superconducting gap.[15]

Starting with the notion that the Fermi level is pinned in impurity states of doped cuprates, the superconducting gap should open inside the insulating gap. In this case, the zero-bias tunneling conductance peak (ZBCP)[16] and absence of a superconducting gap in very low enregy optical spectra[17] (in contrast to BCS superconductors) must be considered a manifestation of these states. Assuming that the nature of these states is determined by polarons,[18,19] impurity complexes,[20] or bistable impurity trapping centers,[21] theories interpreting properties of high-temperature superconductors in terms of the Fermi-liquid theory, BCS theory (not necessarily with phonon-mediated pairing between quasiparticles), and quantum percolation theory have been developed.

On the other hand, many researchers, doubting that these concepts apply to properties of the cuprates, associate high-temperature superconductivity with strong electron correlations in these materials.[1,3,22–27] This second approach is based on the most commonly adopted assumption that high-temperature superconductivity derives from processes occurring in CuO_2 planes. This assumption may turn out to be true, as long as the Fermi level in the doped cuprates is located that part of the electron spectrum whose density of states is governed by a CuO_2 plane.[1] Since the available experimental data indicate the importance of spin fluctuations in doped systems,[1,23,28] we consider it important to direct the reader's attention to an investigation[29,30] of the pairing mechanism due to spin–spin interactions between charge carriers in CuO_2 planes.

Note that, although electronic states can appear in the initial insulating gap under doping, the motion of the Fermi level across the insulating gap as p-doping transforms to n-doping is inherent in models of strongly correlated electrons (see Ref. 1 and references therein). At the same time, experimental data indicate that Fermi levels in p- and n-doped compounds are close, notwithstanding the wide insulating gap (1.5–2 eV).[4]

We are confident that apart from detailed investigations of the mechanism of superconductivity in CuO_2 planes, the

feasibility of different ways of producing superconducting states must be studied. At this point, we focus primarily on the search for superconductivity in semiconductors[31] and investigations of electronic mechanisms of superconductivity in alloys.[32] A long time ago, Pines[33] did not rule out the existence of superconductivity in degenerate doped semiconductors. Presently, it seems worthwhile to analyze this possibility, and in particular to study superconductivity in impurity bands in the absence of translational symmetry. This study is also important if we are to search for new superconductors whose conductivity is not necessarily determined by planes of their lattices.

In this paper, we present a theory of superconductivity in doped insulators. We assume that the Fermi level can be pinned to the impurity bands that are deep inside the insulator gap of the undoped material. Then the superconducting gap should open within the initial gap. States that feed the creation of a condensate in the superconducting state should correspond to impurity bands. This approach, probably, can be also applied to HTSC materials.

Previously we considered formation of impurity bands of both extended and localized states in the model of doped insulators with electron correlations on bare impurity orbitals and hybridization between the orbitals and initial insulator band states.[34,35] These narrow impurity bands of extended states are generated in the initial gap due to hybridization, which leads to virtual single-electron transitions over the impurity ensemble: initial impurity site → band state → another site → band state, and so on. We proved using the self-consistent Hartree–Fock approximation that an insulator–metal quantum transition occurs in such a system as the impurity concentration increases. This approximation, however, does not take account of spin fluctuations in a system of correlated electrons.

We suppose that a superconducting state can be due to spin fluctuations in a doped material. In this study we utilize a model Hamiltonian for the normal and superconducting states. We derive a complete set of equations for normal and anomalous Green's functions and show that the system has several order parameters that determine the state of the doped material. Two different types of these parameters correspond to spin fluctuations. One, associated with normal Green's functions for localized states, initially shows up in properties of the insulating and metallic states of the doped material. At the same time, the order parameters of the other type associated with anomalous Green's functions, which are nondiagonal in the spin variables and diagonal in the impurity sites, can lead to a new nonphonon channel of quasiparticle pairing deep inside the initial insulator gap of an undoped material.

2. MODEL

The formation of impurity states due to substitutional impurities is quite common in doped insulators. For example, when a substitutional impurity is introduced into $Ba_{1-x}K_xBiO_3$, Ba^{2+} ions in the host lattice are randomly replaced by K^{1+} ions. The valence electron of K goes to satisfy the bonding requirements, so a singly occupied acceptor level arises.

Taking into account the Coulomb correlations in the original impurity orbitals, we represent the Hamiltonian of the doped system in the form

$$H=\sum_{\mathbf{k}\sigma}\varepsilon_{\mathbf{k}}a^{+}_{\mathbf{k}\sigma}a_{\mathbf{k}\sigma}+\sum_{j\sigma}\varepsilon_0 a^{+}_{j\sigma}a_{j\sigma}+\frac{1}{2}\sum_{j\sigma}Un_{j\sigma}n_{j,-\sigma}+\sum_{j,\mathbf{k}\sigma}\{V_{\mathbf{k}j}a^{+}_{\mathbf{k}\sigma}a_{j\sigma}+\text{H.c.}\},\quad(1)$$

where $a_{\mathbf{k}\sigma}$ and $a_{j\sigma}$ are the usual annihilation operators of original band states and impurity states labeled by j, respectively; $\sigma=\pm(1/2)(\uparrow,\downarrow)$ is the spin index; $\mathbf{k}$ is the three-dimensional electron wave vector in the band state with energy $\varepsilon_{\mathbf{k}};\varepsilon_0$ is the impurity state energy; $V_{\mathbf{k}j}$ is the hybridization matrix element; and U is the on-site electron correlation for impurity levels.

Hamiltonian (1) can be applied to doped cuprates for the following reason. When substitutional impurities are introduced, for example, into $La_{2-x}Sr_xCuO_4$, La^{3+} ions in LaO layers are replaced randomly by Sr^{2+} ions. Both valence electrons of Sr are involved in valence bonds, so formation of a singly occupied acceptor impurity orbital might be expected. The CuO_2 layer of the cuprate is between two intermediate $La_{1-x/2}M_{x/2}O$ layers. Since the centers of impurity orbitals are located off the CuO_2 plane, the interaction between impurity orbitals and band states of the CuO_2 plane leads to their hybridization. Our analysis is limited for simplicity to the single-band approximation for the density of states of the CuO_2 plane. When the strong anisotropy of cuprates is taken into account, the Hamiltonian for a single lattice cell not coupled to other cells by charge transfer along the $c(z)$ axis can be expressed in the form of Eq. (1) in a two-dimensional space with corresponding wave vectors.

The Nd_2CuO_4 structure is similar to that of La_2CuO_4, the only difference being the location of oxygen atoms off the CuO_2 planes. Under substitutional doping in $Nd_{2-x}Ce_xCuO_4$, Nd^{3+} is randomly replaced by Ce^{4+} and a singly occupied donor level can arise in the original gap. In La_2CuO_{4+y} additional oxygen atoms can occupy interstitial positions near CuO_2 planes. Then all valence electrons of oxygen atoms can take part in formation of impurity bands.

Note that the impurity centers that preserve the Fermi level and are located between cuprate planes (not in these planes) in doped cuprates were investigated by Phillips.[21]

Solution of the problem with Hamiltonian (1) in the Hartree–Fock approximation yields a rather complex pattern of impurity bands with localized and delocalized (extended) states in the insulator gap. Changes in this pattern of impurity bands with the impurity concentrations near the Fermi level determine the insulator–metal phase transition. In insulators with narrower bands of the Bloch states, this transition occurs at lower impurity concentrations. In what follows, we give a solution of the problem with Hamiltonian (1) corresponding to the superconducting state of the system. We will

show that a superconducting phase in the doped material can emerge owing to nothing but spin fluctuations. The calculations will be performed at zero temperature.

3. EQUATIONS FOR THE GREEN'S FUNCTIONS

Let us introduce the normal Green's functions

$$G^{\sigma\sigma_1}_{\eta\eta_1}(t,t_1)=-i\langle N_{\sigma_1-\sigma}|T\tilde{a}_{\eta\sigma}(t)\tilde{a}^+_{\eta_1\sigma_1}(t_1)|N_0\rangle$$

and anomalous Green's functions

$$F^{(+)\sigma\sigma_1}_{\eta\eta_1}(t,t_1)=\exp(-2i\mu t)\langle(N+2)_{\sigma+\sigma_1}\times|T\tilde{a}^+_{\eta\sigma}(t)\tilde{a}^+_{\eta_1\sigma_1}(t_1)|N_0\rangle,$$

$$F^{\sigma\sigma_1}_{\eta\eta_1}(t,t_1)=\exp(2i\mu t)\langle N_0|T\tilde{a}_{\eta,\sigma}(t)\tilde{a}_{\eta_1\sigma_1}(t_1)|\times(N+2)_{\sigma+\sigma_1}\rangle,$$

where we have used operators in the Heisenberg picture with Hamiltonian (1); η denotes $\mathbf{k}$ or j; μ is the chemical potential in the system. Averaging is performed over the system ground states with total particle numbers N and $N+2$.[36] The subscripts at the ground states $|N_{\sigma_1-\sigma}\rangle$ and $|(N+2)_{\sigma+\sigma_1}\rangle$ indicate the existence of additional spins for a given total number of particles, as compared to the spin states $|N_0\rangle$ and $|(N+2)_0\rangle$.

At this point, note the following circumstance. Among the Green's functions introduced previously, there is, for example, the anomalous Green's function $F^{(+)\uparrow\uparrow}_{jj}$. Given Hamiltonian (1), the emergence of this function does not conflict with the Pauli principle. The point is that we are considering a system of interacting fermions, so operator $\tilde{a}^+_{j\uparrow}(t)$ acting on $|N\rangle$ does not change the population of a specific localized (site) state by 1. The average occupation number of any localized (site) state can be assumed to be much smaller than 1. This is the fundamental difference between systems of interacting and noninteracting particles, in which occupation numbers can be either 0 or 1.

Using the equation of motion for Heisenberg operators based on Hamiltonian (1), we obtain the equations for the Green's functions introduced previously:

$$\left(i\frac{\partial}{\partial t}-\varepsilon_{\mathbf{k}}\right)G^{\sigma\sigma_1}_{\mathbf{k}\eta}(t,t_1)=\delta_{\mathbf{k}\eta}\delta_{\sigma\sigma_1}\delta(t-t_1)+\sum_j V_{\mathbf{k}j}G^{\sigma\sigma_1}_{j\eta}(t-t_1), \tag{2}$$

$$\left(i\frac{\partial}{\partial t}+\varepsilon_{\mathbf{k}}-2\mu\right)F^{(+)\sigma\sigma_1}_{\mathbf{k}\eta}(t-t_1)=-\sum_j V_{j\mathbf{k}}F^{(+)\sigma\sigma_1}_{j\eta}(t-t_1), \tag{3}$$

$$\left(i\frac{\partial}{\partial t}-\varepsilon_0\right)G^{\sigma\sigma_1}_{j\eta}(t-t_1)=\delta_{j\eta}\delta_{\sigma\sigma_1}\delta(t-t_1)+\sum_{\mathbf{k}} V_{j\mathbf{k}}G^{\sigma\sigma_1}_{\mathbf{k}\eta}\times(t-t_1)+K^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1), \tag{4}$$

$$\left(i\frac{\partial}{\partial t}+\varepsilon_0-2\mu\right)F^{(+)\sigma\sigma_1}_{j\eta}(t-t_1)=-\sum_{\mathbf{k}} V_{\mathbf{k}j}F^{(+)\sigma\sigma_1}_{\mathbf{k}\eta}(t-t_1)-H^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1). \tag{5}$$

Here we use the following notation for two-particle Green's functions emerging in our calculations:

$$K^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=-iU\langle N_{\sigma_1-\sigma}|T\tilde{a}_{j\sigma}(t)\tilde{a}^+_{j,-\sigma}\times(t)\tilde{a}_{j,-\sigma}(t)\tilde{a}^+_{\eta\sigma_1}(t_1)|N_0\rangle, \tag{6}$$

$$H^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=U\exp(-2i\mu t)\langle(N+2)_{\sigma+\sigma_1}\times|T\tilde{a}^+_{j\sigma}(t)\tilde{a}^+_{j,-\sigma}(t)\tilde{a}_{j,-\sigma}(t)\tilde{a}^+_{\eta\sigma_1}\times(t_1)|N_0\rangle. \tag{7}$$

We now transform to the interaction picture on the right-hand side of Eqs. (6) and (7), with the interaction operator

$$\mathbf{H}_{\rm int}(\tau)=\mathbf{H}^c_{\rm int}(\tau)+\mathbf{H}^g_{\rm int}(\tau),$$

where $\mathbf{H}^c_{\rm int}(\tau)$ and $\mathbf{H}^g_{\rm int}(\tau)$ are determined by the last two terms of Eq. (1), respectively, with operators $a_{\eta\sigma}$ and $a^+_{\eta\sigma}$ in the Schrödinger picture replaced by $a_{\eta\sigma}(t)$ and $a^+_{\eta\sigma}(t)$ in the interaction picture. The unperturbed Hamiltonian is expressed by the first two terms on the right of Eq. (1). As a result, we have

$$K^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=-iU\langle N_{\sigma_1-\sigma}|T[S(\infty)a_{j\sigma}(t)a^+_{j,-\sigma}\times(t)a_{j,-\sigma}(t)a^+_{\eta\sigma_1}(t_1)]|N_0\rangle_c, \tag{8}$$

and

$$H^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=U\exp(-2i\mu t)\langle(N+2)_{\sigma+\sigma_1}\times|T[S(\infty)a^+_{j\sigma}(t)a^+_{j,-\sigma}(t)a_{j,-\sigma}\times(t)a^+_{\eta\sigma_1}(t_1)]|N_0\rangle_c, \tag{9}$$

where only connected diagrams are taken into account, which is indicated by subscript c.

All connected diagrams in Eqs. (8) and (9) can be divided into two groups. One of them contains diagrams corresponding to all possible pairing among the four operators in the Heisenberg picture on the right of Eqs. (6) and (7). Going to the interaction picture, the right-hand side of Eq. (6) becomes

$$\begin{aligned}&-iU\langle N_{\sigma_1-\sigma}|T[S(\infty)a_{j\sigma}(t)a^+_{j,-\sigma}(t)]|N_{\sigma_1+\sigma}\rangle_c\langle N_{\sigma_1+\sigma}\\&\quad\times|T[S(\infty)a_{j,-\sigma}(t)a^+_{\eta\sigma_1}(t_1)]|N_0\rangle_c,\\&\quad+iU\langle N_{\sigma_1-\sigma}|T[S(\infty)a_{j\sigma}(t)a_{j,-\sigma}(t)]|(N+2)_{\sigma_1-\sigma}\rangle_c\\&\quad\times\langle(N+2)_{\sigma_1-\sigma}|T[S(\infty)a^+_{j,-\sigma}(t)a^+_{\eta\sigma_1}(t_1)]|N_0\rangle_c,\\&\quad-iU\langle N_{\sigma_1-\sigma}|T[S(\infty)a_{j\sigma}(t)a^+_{\eta\sigma_1}(t_1)]|N_0\rangle_c\langle N_0\\&\quad\times|T[S(\infty)a^+_{j,-\sigma}(t)a_{j,-\sigma}(t)]|N_0\rangle_c,\end{aligned}$$

which is, by definition, equivalent to

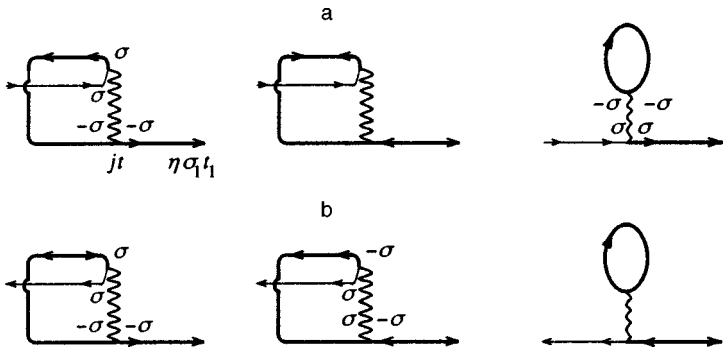


FIG. 1.

$$iUG_{jj}^{\sigma,-\sigma}(0^+)G_{j\eta}^{-\sigma,\sigma_1}(t,t_1)+iUF_{jj}^{\sigma,-\sigma}(0^+)F_{j\eta}^{(+)-\sigma,\sigma_1}\times(t,t_1)-iUG_{jj}^{-\sigma,-\sigma}(-0^+)G_{j\eta}^{\sigma,\sigma_1}(t,t_1). \quad (10)$$

Similarly, we transform the right-hand side of Eq. (7) to

$$iUF_{jj}^{(+)\sigma,-\sigma}(0^+)G_{j\eta}^{-\sigma,\sigma_1}(t,t_1)+iUG_{jj}^{-\sigma,\sigma}\times(0^+)F_{j\eta}^{(+)-\sigma,\sigma_1}(t,t_1)-iUG_{jj}^{-\sigma,-\sigma}\times(-0^+)F_{j\eta}^{(+)\sigma,\sigma_1}(t,t_1). \quad (11)$$

The diagrams for Eqs. (10) and (11) are shown in Figs. 1a and 1b, respectively. Here thin lines correspond to $G_{jj}^{(0)\sigma\sigma}$, wavy lines to iU. The first two diagrams for Eqs. (10) and (11) are related to the on-site interaction with spin-flip processes, whereas the last loop diagrams, which are the Hartree–Fock contribution, result in the impurity level renormalization.

Another group contains all diagrams in which, after the expansion of the S-matrix in terms of $\mathbf{H}_{\rm int}(\tau)$ and application of Wick's theorem, any of the four operators to the right of the S-matrix in Eqs. (8) and (9) is coupled only to the operators of this S-matrix expansion. Since the operator of resonant scattering (hybridization) $\mathbf{H}_{\rm int}^{g}(\tau)$ contains only a pair of operators, the diagrams corresponding to the first-order terms in the S-matrix expansion in terms of $\mathbf{H}_{\rm int}^{g}(\tau)$ have been already taken into account in Eqs. (10) and (11). In the first order in $\mathbf{H}_{\rm int}^{c}(\tau)$ we have for Eqs. (8) and (9) twenty-four simple diagrams of the second group. After partial summation of more complex diagrams, the bare Green's functions are replaced by full functions. As a result, we obtain diagrams shown in Fig. 2 for $K_{jjj\eta}^{\sigma,-\sigma,-\sigma,\sigma_1}(t,t_1)$. Each wavy line corresponds to iU. The summation is performed over j_1 and σ_2, and integration over τ. In front of the resulting expression, the factor $(-1)^{\sigma-\sigma_2}$ must be added. Here we do not show the corresponding diagrams for $H_{jjj\eta}^{\sigma,-\sigma,-\sigma,\sigma_1}(t,t_1)$. They are similar to those of the diagrams in Fig. 2, but the thin line entering the j-site, $G_{jj}^{(0)\sigma\sigma}(t)$, must be replaced by the outgoing line, $G_{jj}^{(0)\sigma\sigma}(-t)$, and the solid lines $G_{jj_1}(t,\tau)$ and $F_{jj_1}(t,\tau)$ associated with $G^{(0)\sigma\sigma}(-t)$ must be replaced

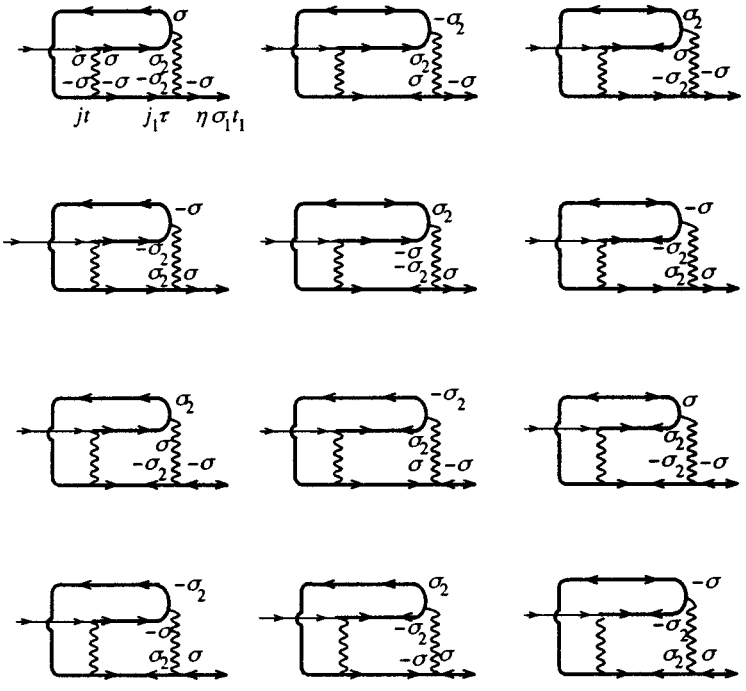


FIG. 2.

by the full Green's functions $G_{j_1 j}(\tau,t)$ and $F^{(+)}_{jj_1}(t,\tau)$. The calculation rules for these diagrams are the same.

In the second order of the *S*-matrix expansion in terms of $\mathbf{H}_{\rm int}(\tau)$, we obtain for Eqs. (8) and (9) *T*-averages of operators containing terms of the second order in $\mathbf{H}^g_{\rm int}(\tau)$, the product of $\mathbf{H}^g_{\rm int}(\tau_1)$ and $\mathbf{H}^c_{\rm int}(\tau_2)$, and a term of the second order in $\mathbf{H}^c_{\rm int}(\tau)$. Note that the *T*-average containing the terms of the second order in hybridization yields diagrams that have already been taken into account in Eqs. (10) and (11). Further, the *T*-average containing the product of $\mathbf{H}^g_{\rm int}(\tau_1)$ and $\mathbf{H}^c_{\rm int}(\tau_2)$ yields diagrams some of which are taken into account in Eqs. (10) and (11) or are among the diagrams of Fig. 1; others are taken into account in the partial summation for the diagrams shown in Fig. 2. The situation is similar in higher orders of perturbation theory. Only the *T*-average, containing terms of the second order in $\mathbf{H}^c_{\rm int}(\tau)$ generates new connected and topologically distinguishable diagrams with three wavy lines. After partial summation of more complex diagrams, one can replace the bare Green's functions with full Green's functions. The diagrams are not shown here.

It is clear that in Eq. (4), which is written, for example, for $G^{\uparrow\uparrow}_{j\mathbf{k}}$, the two-particle Green's function $K^{\uparrow\downarrow\uparrow\downarrow}_{jjj\mathbf{k}}(t,t_1)$ corresponds to a set of diagrams starting with a thin line $G^{(0)\uparrow\uparrow}_{jj}(t)$ and ending with full Green's functions $G^{\uparrow\uparrow}_{j_1\mathbf{k}}(\tau,t_1)$, $G^{\downarrow\uparrow}_{j_1\mathbf{k}}(\tau,t_1)$, $F^{(+)\uparrow\uparrow}_{j_1\mathbf{k}}(\tau,t_1)$, and $F^{(+)\downarrow\uparrow}_{j_1\mathbf{k}}(\tau,t_1)$ (integration is performed over τ and summation over j_1). This allows one to introduce vertex parts to express two-particle Green's functions in the form

$$K^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=\sum_{j_1,\sigma_2}\int d\tau A^{\sigma,\sigma_2}_{jj_1}(t,\tau)G^{\sigma_2,\sigma_1}_{j_1\eta}(\tau,t_1)+\sum_{j_1,\sigma_2}\int d\tau B^{\sigma,\sigma_2}_{jj_1}(t,\tau)F^{(+)\sigma_2,\sigma_1}_{j_1\eta}(\tau,t_1) \tag{12}$$

and

$$H^{\sigma,-\sigma,-\sigma,\sigma_1}_{jjj\eta}(t,t_1)=\sum_{j_1,\sigma_2}\int d\tau C^{\sigma,\sigma_2}_{jj_1}(t,\tau)G^{\sigma_2,\sigma_1}_{j_1\eta}(\tau,t_1)+\sum_{j_1,\sigma_2}\int d\tau D^{\sigma,\sigma_2}_{jj_1}(t,\tau)F^{(+)\sigma_2,\sigma_1}_{j_1\eta}(\tau,t_1). \tag{13}$$

The functions A, B, C, and D can be calculated by summing diagrams. Examples of such diagrams are given in Figs. 1 and 2, where hybridization effects are taken into account by full Green's functions, which directly follows from Eqs. (2)–(5).

Equations (2)–(5), (10)–(13) must be supplemented by the equation for the chemical potential μ of the system:

$$N_t+N_{\rm im}=\frac{1}{\pi}\int_{-\infty}^{\mu}d\omega\,{\rm Im\,Tr}\,G(\omega), \tag{14}$$

where N_t is the total number of electrons in the original energy band of the insulator (for definiteness, in the valence band of the insulator or CuO_2 plane in the case of *p*-doping); $N_{\rm im}$ is the concentration of substitutional impurities. Here we have assumed that the impurity level is singly populated.

Now let us discuss order parameters in the studied model. The loop diagrams in Fig. 1 represent renormalization of the bare impurity level ε_0 due to on-site Coulomb correlations in the Hartree–Fock approximation. This renormalization is determined by the average population numbers of the localized orbitals, which can be expressed in the form

$$G^{\sigma\sigma}_{jj}(-0^+)=\lim_{t\to 0^+}\int\frac{d\omega}{2\pi}G^{\sigma\sigma}_{jj}(\omega)\exp(i\omega t)=i\frac{1}{\pi}\int_{-\infty}^{\mu}d\omega\,{\rm Im}\,G^{\sigma\sigma}_{jj}(\omega)=iA_{j,\sigma}. \tag{15}$$

Since impurity atoms randomly occupy equivalent sites of the crystal lattice, there are only the following options:

1) $A_{j,\sigma}$ is independent of j and σ. As was shown,[14] this case always corresponds to a paramagnetic metallic state in the Hartree–Fock approximation.

2) $A_{j,\sigma}$ is independent of j and $A_{j,\sigma}\neq A_{j,-\sigma}$. This situation corresponds to either a magnetic insulating state or a magnetic metallic state in the Hartree–Fock approximation.[14]

3) At a specific j_0, we have $A_{j_0,\sigma}=A_1$, $A_{j_0,-\sigma}=A_2$, and $A_1\neq A_2$. But for an ensemble of impurities, $A_{j,\sigma}$ as a function of j at a given σ randomly assumes the values A_1 and A_2. For example, for $j_1\neq j_0 A_{j_1,\sigma}=A_2$ and $A_{j_1,-\sigma}=A_1$. This case always corresponds to a paramagnetic state.

Note that it is of interest to investigate the phonon-mediated pairing mechanism in narrow impurity bands in the paramagnetic state 3). In a magnetic metal, we did so previously, assuming triplet pairing.[35] In the paramagnetic phase, the impurity bands are degenerate in the spin variable. Therefore the singlet pairing channel, which is important for cuprates, is also active in this phase.

Moreover, the system has order parameters of two different types associated with spin fluctuations. One is determined by the σ-off-diagonal matrix element of the normal Green's functions for the localized states:

$$G^{\sigma-\sigma}_{jj}(0^+)=\lim_{t\to 0^+}\int\frac{d\omega}{2\pi}G^{\sigma-\sigma}_{jj}(\omega)\exp(-i\omega t)=i\lambda_j, \tag{16}$$

hence

$$G^{-\sigma\sigma}_{jj}(0^+)=-(G^{\sigma-\sigma}_{jj}(0^+))^*=i\lambda^*_j.$$

It is clear that for the normal state, spin fluctuations in the system are described just by λ.

The feasibility of a superconducting state in a system is determined by anomalous Green's functions. An important factor is the order parameter, which is determined by the σ-off-diagonal Green's functions $F^{(+)\sigma,-\sigma}_{jj}(\omega)$ and $F^{\sigma,-\sigma}_{jj}(\omega)$. These functions describe correlated creation or annihilation of a pair of quasiparticles localized at one site, which can be considered a localized zero-spin boson. We introduce the notation

$$F_{jj}^{\sigma,-\sigma}(0^+)=\lim_{t\to 0^+}\int\frac{d\omega}{2\pi}F_{jj}^{\sigma-\sigma}(\omega)\exp(-i\omega t)=i\beta_j. \tag{17}$$

Then we have

$$F_{jj}^{-\sigma,\sigma}(0^+)=-i\beta_j,\quad F_{jj}^{(+)\sigma,-\sigma}(0^+)=i\beta_j^*,$$
$$F_{jj}^{(+)-\sigma,\sigma}(0^+)=-i\beta_j^*.$$

The integrodifferential equations (2)–(5), (10)–(17), which take the diagrams for the vertex parts into account, is complete from a mathematical standpoint. This is a complicated set of equations. In the final analysis, it is important to show that solutions describing superconductivity in impurity bands are feasible. A specific precursor of this is formation of localized bosons, and accordingly the existence of a nonzero solution of Eq. (17). The on-site interaction between bare particles, however, is repulsive. Therefore it is important to analyze the renormalization of the on-site interaction due to hybridization and spin fluctuations. This renormalization is determined by vertex parts A, B, C, and D introduced previously. This analysis will be performed below in an approximation in which for the two-particle Green's function one takes into account only the diagrams in Fig. 1 and in case 1), when $A_{j,\sigma}$ is independent of both j and σ.

4. SOLUTION OF THE EQUATIONS IN THE PARAMAGNETIC PHASE

Note that either the parameter λ [Eq. (16)] or β [Eq. (17)] can be chosen to be real because Eq. (1) does not change after the following transformation in the two spin subspaces:

$$a_{\eta,\uparrow}\to a_{\eta,\uparrow}\exp(i\phi),\quad a_{\eta,\downarrow}\to a_{\eta,\downarrow}\exp(-i\phi),$$

where the phase ϕ is constant. For definiteness, we assume λ to be real.

Using Eqs. (16) and (17), and $F_{jj}^{(+)\downarrow\uparrow}(0^+)=-i\beta^*$, we transform equations (2)–(9) for extended states to

$$G_{\mathbf{k}\mathbf{k}_1}^{\uparrow\uparrow}(\omega)=G_{\mathbf{k}\mathbf{k}}^{(0)}(\omega)\left(\delta_{\mathbf{k}\mathbf{k}_1}+\sum_j V_{\mathbf{k}j}G_{j\mathbf{k}_1}^{\uparrow\uparrow}(\omega)\right), \tag{18}$$

$$G_{j\mathbf{k}}^{\uparrow\uparrow}=\tilde{G}_{ll}\left(\sum_{\mathbf{k}_1}V_{j\mathbf{k}_1}G_{\mathbf{k}_1\mathbf{k}}^{\uparrow\uparrow}-\lambda UG_{j\mathbf{k}}^{\downarrow\uparrow}-\beta UF_{j\mathbf{k}}^{(+)\downarrow\uparrow}\right), \tag{19}$$

$$G_{\mathbf{k}\mathbf{k}_1}^{\downarrow\uparrow}(\omega)=G_{\mathbf{k}\mathbf{k}}^{(0)}(\omega)\sum_j V_{\mathbf{k}j}G_{j\mathbf{k}_1}^{\downarrow\uparrow}(\omega), \tag{20}$$

$$G_{j\mathbf{k}}^{\downarrow\uparrow}=\tilde{G}_{ll}\left(\sum_{\mathbf{k}_1}V_{j\mathbf{k}_1}G_{\mathbf{k}_1\mathbf{k}}^{\downarrow\uparrow}-\lambda UG_{j\mathbf{k}}^{\uparrow\uparrow}+\beta UF_{j\mathbf{k}}^{(+)\uparrow\uparrow}\right), \tag{21}$$

$$F_{\mathbf{k}\mathbf{k}_1}^{(+)\downarrow\uparrow}=-F_{\mathbf{k}\mathbf{k}}^{(0)}\sum_j V_{j\mathbf{k}}F_{j\mathbf{k}_1}^{(+)\downarrow\uparrow}, \tag{22}$$

$$F_{j\mathbf{k}}^{(+)\downarrow\uparrow}=-\tilde{F}_{ll}\left(\sum_{\mathbf{k}_1}V_{\mathbf{k}_1 j}F_{\mathbf{k}_1\mathbf{k}}^{(+)\downarrow\uparrow}+\beta^*UG_{j\mathbf{k}}^{\uparrow\uparrow}-\lambda UF_{j\mathbf{k}}^{(+)\uparrow\uparrow}\right), \tag{23}$$

$$F_{\mathbf{k}\mathbf{k}_1}^{(+)\uparrow\uparrow}=-F_{\mathbf{k}\mathbf{k}}^{(0)}\sum_j V_{j\mathbf{k}}F_{j\mathbf{k}_1}^{(+)\uparrow\uparrow}, \tag{24}$$

$$F_{j\mathbf{k}}^{(+)\uparrow\uparrow}=-\tilde{F}_{ll}\left(\sum_{\mathbf{k}_1}V_{\mathbf{k}_1 j}F_{\mathbf{k}_1\mathbf{k}}^{(+)\uparrow\uparrow}-\beta^*UG_{j\mathbf{k}}^{\downarrow\uparrow}-\lambda UF_{j\mathbf{k}}^{(+)\downarrow\uparrow}\right). \tag{25}$$

Similarly, we obtain for localized states

$$G_{jj_1}^{\uparrow\uparrow}=\tilde{G}_{ll}\left(\delta_{jj_1}+\sum_{\mathbf{k}}V_{j\mathbf{k}}G_{\mathbf{k}j_1}^{\uparrow\uparrow}-\lambda UG_{jj_1}^{\downarrow\uparrow}-\beta UF_{jj_1}^{(+)\downarrow\uparrow}\right), \tag{26}$$

$$G_{\mathbf{k}j}^{\uparrow\uparrow}=G_{\mathbf{k}\mathbf{k}}^{(0)}\sum_{j_1}V_{\mathbf{k}j_1}G_{j_1 j}^{\uparrow\uparrow}, \tag{27}$$

$$G_{jj_1}^{\downarrow\uparrow}=\tilde{G}_{ll}\left(\sum_{\mathbf{k}}V_{j\mathbf{k}}G_{\mathbf{k}j_1}^{\downarrow\uparrow}-\lambda UG_{jj_1}^{\uparrow\uparrow}+\beta UF_{jj_1}^{(+)\uparrow\uparrow}\right), \tag{28}$$

$$G_{\mathbf{k}j}^{\downarrow\uparrow}=G_{\mathbf{k}\mathbf{k}}^{(0)}\sum_{j_1}V_{\mathbf{k}j_1}G_{j_1 j}^{\downarrow\uparrow}, \tag{29}$$

$$F_{jj_1}^{(+)\downarrow\uparrow}=-\tilde{F}_{ll}\left(\sum_{\mathbf{k}}V_{\mathbf{k}j}F_{\mathbf{k}j_1}^{(+)\downarrow\uparrow}+\beta^*UG_{jj_1}^{\uparrow\uparrow}-\lambda UF_{jj_1}^{(+)\uparrow\uparrow}\right), \tag{30}$$

$$F_{\mathbf{k}j}^{(+)\downarrow\uparrow}=-F_{\mathbf{k}\mathbf{k}}^{(0)}\sum_{j_1}V_{j_1\mathbf{k}}F_{j_1 j}^{(+)\downarrow\uparrow}, \tag{31}$$

$$F_{jj_1}^{(+)\uparrow\uparrow}=-\tilde{F}_{ll}\left(\sum_{\mathbf{k}}V_{\mathbf{k}j}F_{\mathbf{k}j_1}^{(+)\uparrow\uparrow}-\beta^*UG_{jj_1}^{\downarrow\uparrow}-\lambda UF_{jj_1}^{(+)\downarrow\uparrow}\right), \tag{32}$$

$$F_{\mathbf{k}j}^{(+)\uparrow\uparrow}=-F_{\mathbf{k}\mathbf{k}}^{(0)}\sum_{j_1}V_{j_1\mathbf{k}}F_{j_1 j}^{(+)\uparrow\uparrow}. \tag{33}$$

Here we have used the notation

$$G_{\mathbf{k}\mathbf{k}}^{(0)}(\omega)=(\omega-\varepsilon_{\mathbf{k}})^{-1},$$
$$F_{\mathbf{k}\mathbf{k}}^{(0)}(\omega)=(\omega+\varepsilon_{\mathbf{k}}-2\mu)^{-1},$$
$$\tilde{G}_{ll}(\omega)=(\omega-\varepsilon_0+iUG_{jj}^{-\sigma-\sigma}(-0^+))^{-1},$$
$$\tilde{F}_{ll}(\omega)=(\omega+\varepsilon_0-2\mu-iUG_{jj}^{-\sigma-\sigma}(-0^+))^{-1}.$$

Since $A_{j,\sigma}$ is independent of j and σ, functions $\tilde{G}_{jj}^{\sigma\sigma}$ and $\tilde{F}_{jj}^{\sigma\sigma}$ are independent of j, which can be replaced by l, and of σ, which can be omitted. We next solve Eqs. (18)–(33).

4.1. Localized states

Consider Eqs. (26) and (27). Substituting the off-diagonal matrix element $G_{j_1 j}$ given by Eq. (26) into Eq. (27), we obtain

$$G_{\mathbf{k}j}^{\uparrow\uparrow}(G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}-N_{\mathrm{im}}|V_{\mathbf{k}l}|^2\tilde{G}_{ll})=V_{\mathbf{k}j}G_{jj}^{\uparrow\uparrow}-\lambda U\tilde{G}_{ll}\sum_{j_1\neq j}V_{\mathbf{k}j_1}G_{j_1 j}^{\downarrow\uparrow}$$
$$-\beta U\tilde{G}_{ll}\sum_{j_1\neq j}V_{\mathbf{k}j_1}F_{j_1 j}^{(+)\downarrow\uparrow}$$
$$+\tilde{G}_{ll}\sum_{\mathbf{k}_1\neq\mathbf{k},j_1\neq j}V_{\mathbf{k}j_1}V_{j_1\mathbf{k}_1}G_{\mathbf{k}_1 j}^{\uparrow\uparrow}. \tag{34}$$

For the impurity ensemble, the hybridization matrix element $V_{\mathbf{k}j}$ has the form

$$V_{\mathbf{k}j}=\frac{1}{\sqrt{\Omega}}V_{\mathbf{k}l}\exp(i\mathbf{k}\mathbf{R}_j),$$

where $\mathbf{R}_j$ is the three-dimensional radius vector of the jth impurity atom and Ω is the sample volume. In cuprates one must replace Ω by S, the area of the CuO_2 plane, and take into account that $\mathbf{R}_j$ is the impurity two-dimensional radius vector in an intermediate plane, for example, in $La_{1-x/2}M_{x/2}O$ planes in $La_{2-x}Sr_xCuO_4$.

Summing over j_1 in the last term on the right-hand side of Eq. (34), one must average the result over random distributions of impurity atoms. In calculating these configuration averages, we have utilized the technique developed by Yonezawa and Matsubara.[37] As a result, we have found that the last term on the right of Eq. (34) is proportional to

$$\left\langle \sum_{\mathbf{k}_1\neq\mathbf{k},j_1\neq j}\exp(i(\mathbf{k}-\mathbf{k}_1)\mathbf{R}_{j_1})\right\rangle_{\text{av}}=N_{\text{im}}\sum_{\mathbf{k}_1\neq\mathbf{k}}\delta(\mathbf{k}_1-\mathbf{k})=0,$$

where $\langle\ldots\rangle_{\text{av}}$ denotes averaging over all possible configurations of the impurity ensemble. Similar expressions are obtained in dealing with the pairs of equations (28) and (29), (30) and (31), and (32) and (33). In view of this result, in what follows we neglect terms with such double sums.

Substituting

$$\sum_{j_1\neq j}V_{\mathbf{k}j_1}G^{\downarrow\uparrow}_{j_1j}=G^{\downarrow\uparrow}_{\mathbf{k}j}G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}-V_{\mathbf{k}j}G^{\downarrow\uparrow}_{jj}$$

according to Eq. (29) and

$$\sum_{j_1\neq j}V_{\mathbf{k}j_1}F^{(+)\downarrow\uparrow}_{j_1j}=-F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}-V_{\mathbf{k}j}F^{(+)\downarrow\uparrow}_{jj},$$

according to (31) into (34), we obtain an expression relating the off-diagonal Green's functions, $G^{\uparrow\uparrow}_{\mathbf{k}j}$, $G^{\downarrow\uparrow}_{\mathbf{k}j}$, and $F^{(+)\downarrow\uparrow}_{\mathbf{k}j}$, to the j-diagonal Green's functions, $G^{\uparrow\uparrow}_{jj}$, $G^{\downarrow\uparrow}_{jj}$, and $F^{(+)\downarrow\uparrow}_{jj}$.

Applying this approach to the paired equations (28) and (29), (30) and (31), and (32) and (33), we obtain

$$\begin{aligned}&(G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{G}_{ll})G^{\uparrow\uparrow}_{\mathbf{k}j}+\lambda U\tilde{G}_{ll}G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}G^{\downarrow\uparrow}_{\mathbf{k}j}\\&\quad-\beta U\tilde{G}_{ll}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}=V_{\mathbf{k}j}(G^{\uparrow\uparrow}_{jj}+\lambda U\tilde{G}_{ll}G^{\downarrow\uparrow}_{jj}\\&\quad+\beta U\tilde{G}_{ll}F^{(+)\downarrow\uparrow}_{jj}),\end{aligned}\tag{35}$$

$$\begin{aligned}&\lambda U\tilde{G}_{ll}G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}G^{\uparrow\uparrow}_{\mathbf{k}j}+(G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{G}_{ll})G^{\downarrow\uparrow}_{\mathbf{k}j}\\&\quad+\beta U\tilde{G}_{ll}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}F^{(+)\uparrow\uparrow}_{-\mathbf{k}j}=V_{\mathbf{k}j}(G^{\downarrow\uparrow}_{jj}+\lambda U\tilde{G}_{ll}G^{\uparrow\uparrow}_{jj}\\&\quad-\beta U\tilde{G}_{ll}F^{(+)\uparrow\uparrow}_{jj}),\end{aligned}\tag{36}$$

$$\begin{aligned}&-\beta^*U\tilde{F}_{ll}G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}G^{\uparrow\uparrow}_{\mathbf{k}j}+(F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{F}_{ll})F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}\\&\quad-\lambda U\tilde{F}_{ll}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}F^{(+)\uparrow\uparrow}_{-\mathbf{k}j}=-V_{\mathbf{k}j}(F^{(+)\downarrow\uparrow}_{jj}\\&\quad+\beta^*U\tilde{F}_{ll}G^{\uparrow\uparrow}_{jj}-\lambda U\tilde{F}_{ll}F^{(+)\uparrow\uparrow}_{jj}),\end{aligned}\tag{37}$$

$$\begin{aligned}&\beta^*U\tilde{F}_{ll}G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}G^{\downarrow\uparrow}_{\mathbf{k}j}+(F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{F}_{ll})F^{(+)\uparrow\uparrow}_{-\mathbf{k}j}\\&\quad-\lambda U\tilde{F}_{ll}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}=-V_{\mathbf{k}j}(F^{(+)\uparrow\uparrow}_{jj}\\&\quad-\beta^*U\tilde{F}_{ll}G^{\downarrow\uparrow}_{jj}-\lambda U\tilde{F}_{ll}F^{(+)\downarrow\uparrow}_{jj}).\end{aligned}\tag{38}$$

The solution of Eqs. (35)–(38) yields expressions for the off-diagonal Green's functions expressed in terms of the diagonal Green's functions. Substitution of these expressions into Eqs. (26), (28), (30), and (32) at $j=j_1$ makes possible a transformation to a system of four linear algebraic equations for the j-diagonal Green's functions. The latter functions determine the three desired parameters and spectra of localized states.

There is, however, a more convenient option: calculate the sum and difference of the paired equations (35), (36) and (37), (38). We have

$$\begin{aligned}&a_1(\lambda)(G^{\uparrow\uparrow}_{\mathbf{k}j}+G^{\downarrow\uparrow}_{\mathbf{k}j})-c_3F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}(F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}-F^{(+)\uparrow\uparrow}_{-\mathbf{k}j})\\&\quad=V_{\mathbf{k}j}(c_1(\lambda)(G^{\uparrow\uparrow}_{jj}+G^{\downarrow\uparrow}_{jj})+c_3(F^{(+)\downarrow\uparrow}_{jj}-F^{(+)\uparrow\uparrow}_{jj})),\end{aligned}\tag{39}$$

$$\begin{aligned}&a_1(-\lambda)(G^{\uparrow\uparrow}_{\mathbf{k}j}-G^{\downarrow\uparrow}_{\mathbf{k}j})-c_3F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}(F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}+F^{(+)\uparrow\uparrow}_{-\mathbf{k}j})\\&\quad=V_{\mathbf{k}j}(c_1(-\lambda)(G^{\uparrow\uparrow}_{jj}-G^{\downarrow\uparrow}_{jj})+c_3(F^{(+)\downarrow\uparrow}_{jj}+F^{(+)\uparrow\uparrow}_{jj})),\end{aligned}\tag{40}$$

$$\begin{aligned}&a_2(-\lambda)(F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}+F^{(+)\uparrow\uparrow}_{-\mathbf{k}j})-c_4^*G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}(G^{\uparrow\uparrow}_{\mathbf{k}j}-G^{\downarrow\uparrow}_{\mathbf{k}j})\\&\quad=-V_{\mathbf{k}j}(c_2(-\lambda)(F^{(+)\downarrow\uparrow}_{jj}+F^{(+)\uparrow\uparrow}_{jj})+c_4^*(G^{\uparrow\uparrow}_{jj}-G^{\downarrow\uparrow}_{jj})),\end{aligned}\tag{41}$$

$$\begin{aligned}&a_2(\lambda)(F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}-F^{(+)\uparrow\uparrow}_{-\mathbf{k}j})-c_4^*G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}(G^{\uparrow\uparrow}_{\mathbf{k}j}+G^{\downarrow\uparrow}_{\mathbf{k}j})\\&\quad=-V_{\mathbf{k}j}(c_2(\lambda)(F^{(+)\downarrow\uparrow}_{jj}-F^{(+)\uparrow\uparrow}_{jj})+c_4^*(G^{\uparrow\uparrow}_{jj}+G^{\downarrow\uparrow}_{jj})).\end{aligned}\tag{42}$$

Here we use the notation $c_1(-\lambda)=1-\lambda U\tilde{G}_{ll}$, $c_2(\lambda)=1+\lambda U\tilde{F}_{ll}$, $a_1(\lambda)=c_1(\lambda)G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{G}_{ll}$, $a_2(-\lambda)=c_2(-\lambda)F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}-N_{\text{im}}V^2_{\mathbf{k}l}\tilde{F}_{ll}$, $c_3=\beta U\tilde{G}_{ll}$, $c_4^*=\beta^*U\tilde{F}_{ll}$.

Now, using Eqs. (39) and (42), we can derive, for example, expressions for $G^{\uparrow\uparrow}_{\mathbf{k}j}+G^{\downarrow\uparrow}_{\mathbf{k}j}$ and $F^{(+)\downarrow\uparrow}_{-\mathbf{k}j}-F^{(+)\uparrow\uparrow}_{-\mathbf{k}j}$ in terms of the j-diagonal Green's functions. Substitute the former expression into the sum of Eqs. (26) and (28) for diagonal functions ($j=j_1$), and the latter into the difference between Eqs. (30) and (32). As a result, we obtain

$$(G^{\uparrow\uparrow}_{jj}+G^{\downarrow\uparrow}_{jj})(\omega-\varepsilon_0-S_1)+\beta UZ(F^{(+)\downarrow\uparrow}_{jj}-F^{(+)\uparrow\uparrow}_{jj})=1\tag{43}$$

and

$$\begin{aligned}&(F^{(+)\downarrow\uparrow}_{jj}-F^{(+)\uparrow\uparrow}_{jj})(\omega+\varepsilon_0-2\mu-S_2)\\&\quad+\beta^*UZ(G^{\uparrow\uparrow}_{jj}+G^{\downarrow\uparrow}_{jj})=0.\end{aligned}\tag{44}$$

Here we have introduced the self-energies, which are functions of A, μ, λ, and β:

$$S_1(\omega;\lambda)=U(A-\lambda)+\sum_{\mathbf{k}}\frac{V^2_{\mathbf{k}l}[(c_1(\lambda)a_2(\lambda)-c_3c_4^*F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}]}{a_1(\lambda)a_2(\lambda)-c_3c_4^*G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}},\tag{45}$$

$$Z(\omega;\lambda)=1+N_{\text{im}}\tilde{G}_{ll}\tilde{F}_{ll}\times\sum_{\mathbf{k}}\frac{V^4_{\mathbf{k}l}}{a_1(\lambda)a_2(\lambda)-c_3c_4^*G^{(0)^{-1}}_{\mathbf{k}\mathbf{k}}F^{(0)^{-1}}_{-\mathbf{k}-\mathbf{k}}},\tag{46}$$

$$S_2(\omega;\lambda)=-U(A+\lambda)+\sum_{\mathbf{k}}\frac{V_{\mathbf{k}l}^2[(c_2(\lambda)a_1(\lambda)-c_3c_4^*G_{\mathbf{kk}}^{(0)^{-1}}]}{a_1(\lambda)a_2(\lambda)-c_3c_4^*G_{\mathbf{kk}}^{(0)^{-1}}F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}}. \quad (47)$$

The solution of (43) and (44) has the form

$$G_{jj}^{\uparrow\uparrow}+G_{jj}^{\downarrow\uparrow}=G_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta), \quad (48)$$

where

$$G_{\mathrm{loc}}=\frac{\omega+\varepsilon_0-2\mu-S_2(\lambda)}{(\omega+\varepsilon_0-2\mu-S_2(\lambda))(\omega-\varepsilon_0-S_1(\lambda))-|\beta|^2U^2Z(\lambda)^2}, \quad (49)$$

and

$$F_{jj}^{(+)\downarrow\uparrow}-F_{jj}^{(+)\uparrow\uparrow}=-F_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta), \quad (50)$$

where

$$F_{\mathrm{loc}}=\frac{\beta^*UZ(\omega;A,\mu,\lambda;\beta)}{(\omega+\varepsilon_0-2\mu-S_2(\lambda))(\omega-\varepsilon_0-S_1(\lambda))-|\beta|^2U^2Z^2(\lambda).} \quad (51)$$

One can easily verify that the procedure applied to Eqs. (40) and (41), which yields definitions of $G_{jj}^{\uparrow\uparrow}-G_{jj}^{\downarrow\uparrow}$ and $F_{jj}^{(+)\downarrow\uparrow}+F_{jj}^{(+)\uparrow\uparrow}$, results in expressions for these quantities in a form similar to Eqs. (48), (49) and (50), (51) when $\lambda\to-\lambda$. As a result, we have

$$G_{jj}^{\uparrow\uparrow}-G_{jj}^{\downarrow\uparrow}=G_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta) \quad (52)$$

and

$$F_{jj}^{(+)\downarrow\uparrow}+F_{jj}^{(+)\uparrow\uparrow}=-F_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta). \quad (53)$$

Finally, we derive from Eqs. (48)–(53) the j-diagonal Green's functions

$$G_{jj}^{\uparrow\uparrow}=\frac{1}{2}G_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta)+\frac{1}{2}G_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta), \quad (54)$$

$$G_{jj}^{\downarrow\uparrow}=\frac{1}{2}G_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta)-\frac{1}{2}G_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta), \quad (55)$$

$$F_{jj}^{(+)\downarrow\uparrow}=-\frac{1}{2}F_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta)-\frac{1}{2}F_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta), \quad (56)$$

$$F_{jj}^{(+)\uparrow\uparrow}=\frac{1}{2}F_{\mathrm{loc}}(\omega;\mu,A,\lambda,\beta)-\frac{1}{2}F_{\mathrm{loc}}(\omega;\mu,A,-\lambda,\beta). \quad (57)$$

4.2. Extended states

Consider, for example, Eqs. (18) and (19). Substituting the off-diagonal $\mathbf{k}$ matrix element $G_{\mathbf{kk}_1}$ given by Eq. (18) into Eq. (19), we have

$$G_{j\mathbf{k}}^{\uparrow\uparrow}(\omega)\left(\tilde{G}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq\mathbf{k}}\left|V_{\mathbf{k}_1l}\right|^2G_{\mathbf{k}_1\mathbf{k}_1}^{(0)}\right)+\lambda UG_{j\mathbf{k}}^{\downarrow\uparrow}+\beta UF_{j\mathbf{k}}^{(+)\downarrow\uparrow}=V_{j\mathbf{k}}G_{\mathbf{kk}}^{\uparrow\uparrow}+\sum_{\mathbf{k}_1\neq\mathbf{k},j_1\neq j}V_{j\mathbf{k}_1}G_{\mathbf{k}_1\mathbf{k}_1}^{(0)}V_{\mathbf{k}_1j_1}G_{j_1\mathbf{k}}^{\uparrow\uparrow}. \quad (58)$$

In accordance with Eqs. (35)–(38), the Green's function $G_{j\mathbf{k}}^{\uparrow\uparrow}\propto V_{j\mathbf{k}}$. Taking into account the above reasoning, we neglect terms like the last on the right-hand side of Eq. (58). As a result, we derive from Eqs. (18)–(25)

$$G_{j\mathbf{k}}^{\uparrow\uparrow}(\omega)\left(\tilde{G}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq\mathbf{k}}V_{\mathbf{k}_1l}^2G_{\mathbf{k}_1\mathbf{k}_1}^{(0)}\right)+\lambda UG_{j\mathbf{k}}^{\downarrow\uparrow}+\beta UF_{j\mathbf{k}}^{(+)\downarrow\uparrow}=V_{j\mathbf{k}}G_{\mathbf{kk}}^{\uparrow\uparrow}, \quad (59)$$

$$G_{j\mathbf{k}}^{\downarrow\uparrow}(\omega)\left(\tilde{G}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq\mathbf{k}}V_{\mathbf{k}_1l}^2G_{\mathbf{k}_1\mathbf{k}_1}^{(0)}\right)+\lambda UG_{j\mathbf{k}}^{\uparrow\uparrow}-\beta UF_{j\mathbf{k}}^{(+)\uparrow\uparrow}=V_{j\mathbf{k}}G_{\mathbf{kk}}^{\downarrow\uparrow}, \quad (60)$$

$$F_{j\mathbf{k}}^{(+)\downarrow\uparrow}(\omega)\left(\tilde{F}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq-\mathbf{k}}V_{\mathbf{k}_1l}^2F_{\mathbf{k}_1\mathbf{k}_1}^{(0)}\right)+\beta^*UG_{j\mathbf{k}}^{\uparrow\uparrow}-\lambda UF_{j\mathbf{k}}^{(+)\uparrow\uparrow}=-V_{j\mathbf{k}}F_{-\mathbf{kk}}^{(+)\downarrow\uparrow}, \quad (61)$$

$$F_{j\mathbf{k}}^{(+)\uparrow\uparrow}(\omega)\left(\tilde{F}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq-\mathbf{k}}V_{\mathbf{k}_1l}^2F_{\mathbf{k}_1\mathbf{k}_1}^{(0)}\right)-\beta^*UG_{j\mathbf{k}}^{\downarrow\uparrow}-\lambda UF_{j\mathbf{k}}^{(+)\downarrow\uparrow}=-V_{j\mathbf{k}}F_{-\mathbf{kk}}^{(+)\uparrow\uparrow}. \quad (62)$$

Equations (59)–(62) make it possible to obtain expressions for functions that are off-diagonal in the lower indices in terms of diagonal functions. We take the paired equations (59), (60) and (61), (62) and calculate their sums and differences. As a result, we obtain

$$b_1(\lambda)(G_{j\mathbf{k}}^{\uparrow\uparrow}+G_{j\mathbf{k}}^{\downarrow\uparrow})+\beta U(F_{j\mathbf{k}}^{(+)\downarrow\uparrow}-F_{j\mathbf{k}}^{(+)\uparrow\uparrow})=V_{j\mathbf{k}}(G_{\mathbf{kk}}^{\uparrow\uparrow}+G_{\mathbf{kk}}^{\downarrow\uparrow}), \quad (63)$$

$$b_2(\lambda)(F_{j\mathbf{k}}^{(+)\downarrow\uparrow}-F_{j\mathbf{k}}^{(+)\uparrow\uparrow})+\beta^*U(G_{j\mathbf{k}}^{\uparrow\uparrow}+G_{j\mathbf{k}}^{\downarrow\uparrow})$$
$$=-V_{j\mathbf{k}}(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}-F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow}), \tag{64}$$

$$b_1(-\lambda)(G_{j\mathbf{k}}^{\uparrow\uparrow}-G_{j\mathbf{k}}^{\downarrow\uparrow})+\beta U(F_{j\mathbf{k}}^{(+)\downarrow\uparrow}+F_{j\mathbf{k}}^{(+)\uparrow\uparrow})$$
$$=V_{j\mathbf{k}}(G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}-G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow}), \tag{65}$$

$$b_2(-\lambda)(F_{j\mathbf{k}}^{(+)\downarrow\uparrow}+F_{j\mathbf{k}}^{(+)\uparrow\uparrow})+\beta^*U(G_{j\mathbf{k}}^{\uparrow\uparrow}-G_{j\mathbf{k}}^{\downarrow\uparrow})$$
$$=-V_{j\mathbf{k}}(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}+F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow}). \tag{66}$$

Here we have introduced the notation

$$b_1(\lambda)=\tilde{G}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq\mathbf{k}}V_{\mathbf{k}_1 l}^2G_{\mathbf{k}_1\mathbf{k}_1}^{(0)}+\lambda U,$$

$$b_2(-\lambda)=\tilde{F}_{ll}^{-1}-\sum_{\mathbf{k}_1\neq-\mathbf{k}}V_{\mathbf{k}_1 l}^2F_{\mathbf{k}_1\mathbf{k}_1}^{(0)}-\lambda U.$$

Using Eqs. (63)–(66), (18), (20), (22), and (24), we obtain the first pair of equations:

$$(G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}+G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow})\left(G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}-\frac{N_{\mathrm{im}}V_{\mathbf{k}_1 l}^2}{b_1(\lambda)}\right)$$
$$-\frac{\beta U}{b_1(\lambda)}F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}-F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow})=1, \tag{67}$$

$$(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}-F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow})\left(F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}-\frac{N_{\mathrm{im}}V_{\mathbf{k}_1 l}^2}{b_2(\lambda)}\right)$$
$$-\frac{\beta^*U}{b_2(\lambda)}G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}(G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}+G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow})=-\frac{\beta^*U}{b_2(\lambda)}, \tag{68}$$

and the second pair of equations

$$(G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}-G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow})\left(G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}-\frac{N_{\mathrm{im}}V_{\mathbf{k}_1 l}^2}{b_1(-\lambda)}\right)$$
$$-\frac{\beta U}{b_1(-\lambda)}F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}+F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow})=1, \tag{69}$$

$$(F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}+F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow})\left(F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}-\frac{N_{\mathrm{im}}V_{\mathbf{k}_1 l}^2}{b_2(-\lambda)}\right)$$
$$-\frac{\beta^*U}{b_2(-\lambda)}G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}(G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}-G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow})=-\frac{\beta^*U}{b_2(-\lambda)}. \tag{70}$$

The solution of Eqs. (67)–(70) has the form

$$G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}+G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow}=G_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta), \tag{71}$$

where

$$G_{\mathrm{ext}}=\frac{F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}-b_1(\lambda)W(\lambda)}{(G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}-b_2(\lambda)W(\lambda))(F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}-b_1(\lambda)W(\lambda))-|\beta|^2U^2W^2(\lambda)}, \tag{72}$$

and

$$F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}-F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow}=F_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta), \tag{73}$$

where

$$F_{\mathrm{ext}}=\frac{\beta^*UW(\lambda)}{(G_{\mathbf{k}\mathbf{k}}^{(0)^{-1}}-b_2(\lambda)W(\lambda))(F_{-\mathbf{k}-\mathbf{k}}^{(0)^{-1}}-b_1(\lambda)W(\lambda))-|\beta|^2U^2W^2(\lambda)}. \tag{74}$$

Here we have used the notation

$$W(\lambda)=\frac{N_{\mathrm{im}}V_{\mathbf{k}l}^2}{b_1(\lambda)b_2(\lambda)-|\beta|^2U^2}. \tag{75}$$

The expressions for $G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}-G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow}$ and $F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}+F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow}$ equal the right-hand sides of Eqs. (71) and (73) after substituting $\lambda\to-\lambda$ in Eqs. (72) and (74).

Finally we have the **k**-diagonal Green's functions

$$G_{\mathbf{k}\mathbf{k}}^{\uparrow\uparrow}=\frac{1}{2}G_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta)+\frac{1}{2}G_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,-\lambda,\beta), \tag{76}$$

$$G_{\mathbf{k}\mathbf{k}}^{\downarrow\uparrow}=\frac{1}{2}G_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta)-\frac{1}{2}G_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,-\lambda,\beta), \tag{77}$$

$$F_{-\mathbf{k}\mathbf{k}}^{(+)\downarrow\uparrow}=\frac{1}{2}F_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta)+\frac{1}{2}F_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,-\lambda,\beta), \tag{78}$$

$$F_{-\mathbf{k}\mathbf{k}}^{(+)\uparrow\uparrow}=-\frac{1}{2}F_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,\lambda,\beta)+\frac{1}{2}F_{\mathrm{ext}}(\mathbf{k},\omega;\mu,A,-\lambda,\beta). \tag{79}$$

5. DISCUSSION AND CONCLUSIONS

We have derived the solution of Eq. (1) corresponding to a superconducting state of a doped dielectric due to spin fluctuations. In the approximation applied to vertex parts (10) and (11), we have taken into account spin-flip electron scattering. This process has generated the last two terms on the right-hand side of Eqs. (19), (21), (23), (25), (26), (28), (30) and (32). The amplitudes of these scattering processes are determined by two spin-fluctuation order parameters.

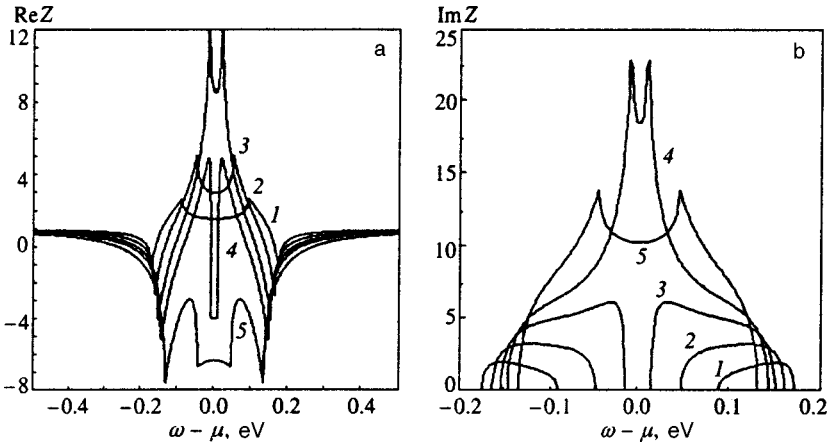


FIG. 3. Curves of (a) $\mathrm{Re}\,Z(\omega)$ and (b) $\mathrm{Im}\,Z(\omega)$ calculated by Eqs. (80)–(83) at $N_t=0.1$ Å^{-d} ($d=2, 3$), $D_b=1.5$ eV, $N_{\mathrm{im}}=0.1N_t$, $\varepsilon_0=D_b+0.5$ eV, $U=0.5$ eV, $\mu=2.5$ eV, and $\mu-\varepsilon_0-UA=0.2$ eV: *1*) $V_h=1.0$ eV; *2*) 1.2 eV; *3*) 1.35 eV; *4*) 1.45 eV; *5*) 1.6 eV.

One of these parameters, λ, is derived from Eq. (16) and expressed in terms of the normal Green's function $G_{jj}^{-\sigma\sigma}$ [Eq. (55)], off-diagonal in the spin index. It is clear that nonzero solutions for λ can be obtained in both normal (metallic or insulating) and superconducting states of the system.

On the other hand, a nonzero solution for the second parameter β, which can be expressed using Eq. (17) in terms of the anomalous Green's function $F_{jj}^{(+)-\sigma\sigma}$ [Eq. (56)], signifies the creation of localized bosons with spin 0 in the system. Besides these, the system should contain localized bosons with spin 1, as follows from Eqs. (51) and (57). It is the parameter β that determines function (74), which describes a superconducting condensate in both singlet [Eq. (78)] and triplet [Eq. (79)] quasiparticle pairing channels.

The possibility of a nontrivial solution of Eq. (17) is dictated by the contribution of poles or their neighborhoods in the anomalous Green's function (56), which are located near the Fermi level on the semiaxis $\mathrm{Re}\,\omega\geqslant\mu$ in the lower half of the complex plane ($\mathrm{Im}\,\omega<0$). This Green's function is a function of Z, which is the renormalized energy of electron correlations due to spin fluctuations and hybridization (46). If the second term on the right-hand side of Eq. (46) is less than the 1 due to the bare correlation energy, one can easily find that Eq. (17) has only the trivial solution $\beta=0$, hence the system is in a normal state (either metallic or insulating). A solution $\beta\neq0$ exists if there is at least a region near the Fermi level where $\mathrm{Re}\,Z<0$. This would mean that the correlation energy UZ in this spectral range has changed sign from plus in the bare state to minus, and corresponds to attractive interaction between quasiparticles.

The function Z depends on μ, A, λ, and β, which are self-consistent parameters and determined by Eqs. (14)–(17). In order to prove the possibility of $\mathrm{Re}\,Z<0$, let us estimate Z in the Hartree–Fock approximation ($\lambda=0$ and $\beta=0$). Assuming that $V_{\mathbf{k}l}$ is independent of $\mathbf{k}$ and considering for definiteness the valence band of the undoped insulator, we derive from Eq. (46)

$$Z=1+\frac{0.5N_{\mathrm{im}}V_{\mathbf{k}l}^4}{\omega^2-(\mu-\varepsilon_0-UA)^2-N_{\mathrm{im}}V_{\mathbf{k}l}^2}R(\omega), \tag{80}$$

where ω is measured with respect to μ,

$$R(\omega)=\frac{1}{\omega}\left[\sum_{\mathbf{k}}\frac{1}{\xi(\omega)-\varepsilon_{\mathbf{k}}-i0^+}-\sum_{\mathbf{k}}\frac{1}{\xi(-\omega)-\varepsilon_{\mathbf{k}}-i0^+}\right], \tag{81}$$

and

$$\xi(\omega)=\omega+\mu-\frac{N_{\mathrm{im}}V_{\mathbf{k}l}^2}{\omega-\varepsilon_0-UA+\mu}. \tag{82}$$

It follows from Eqs. (80)–(82) that the effective correlation energy UZ is a complex, even function of ω. In our calculations, we approximate the density of states in the valence band by

$$\rho^{(0)}(\varepsilon)=\begin{cases}\dfrac{N_t}{\pi D_b^2}[D_b^2-\varepsilon^2]^{1/2}, & |\varepsilon|\leqslant D_b,\\ 0, & |\varepsilon|>D_b.\end{cases} \tag{83}$$

Figure 3 shows $\mathrm{Re}\,Z$ and $\mathrm{Im}\,Z$ as functions of ω for various hybridization parameters $V_h=V_{\mathbf{k}l}N_t^{1/2}$. The function $Z(\omega)$ is very sensitive to the relation between μ and ε_0+UA, but in any case, $\mu\geqslant\varepsilon_0+UA$, as follows from Eq. (15). For the given $\mu-\varepsilon_0-UA=0.2$ eV and at relatively small V_h, there are only narrow regions far from the Fermi level where $\mathrm{Re}\,Z<0$ (curve *1* in Fig. 3a). The function $\mathrm{Im}\,Z$ is also relatively small and nonzero at a considerable distance from the Fermi level (curve *1* in Fig. 3b). As V_h increases, the regions where $\mathrm{Re}\,Z<0$ widen and overlap at the Fermi level, as shown by curves *2*, *3*, *4*, and *5* in Fig. 3a. A similar behavior is demonstrated by $\mathrm{Im}\,Z$ (curves *2*, *3*, *4*, and *5* in Fig. 3b). At the hybridization parameter corresponding to curves *5* in Fig. 3 but a larger detuning parameter $\mu-\varepsilon_0-UA=0.3$ eV, the curves of $Z(\omega)$ are similar to curves *1* in Fig. 3. At a higher hybridization parameter of $V_h=1.9$ eV, however, the curves of $\mathrm{Re}\,Z(\omega)$ and $\mathrm{Im}\,Z(\omega)$ are again similar to curves *5* in Fig. 3.

In conclusion note the following circumstance. We have shown that in the Hartree–Fock approximation, the correlation energy renormalized by the hybridization can be negative near the Fermi level. But in this approximation only single-particle states are generated, hence the system state is a normal one. In order to obtain a superconducting state in a

doped insulator due to the negative correlation energy, one must take into consideration spin fluctuations in the system. In this case, we obtain an equation system including anomalous Green's functions for localized bosons and a superconducting condensate, so the effective correlation energy also becomes a function of spin-fluctuation order parameters.

In this theory of the spin-fluctuation-induced superconductivity in impurity bands of doped insulators, the superconducting gap cannot be described by a BCS-like equation. It is replaced by the key equation (17), which describes formation of localized zero-spin bosons. The mechanism of formation of a superconducting condensate near the Fermi level in the insulator gap is determined by virtual two-quasiparticle transitions over the impurity ensemble: a boson localized at an impurity site → a pair of quasiparticles in the $(-\mathbf{kk})$ extended states of the impurity band → a localized boson at another impurity site, etc. In this case, the singlet and triplet quasiparticle pairing channels are mutually related.

*)E-mail: aai@isssph.kiae.su

[1] E. Dagotto, Rev. Mod. Phys. **66**, 673 (1994).
[2] Yu. A. Izyumov, N. M. Plakida, and Yu. N. Skryabin, Sov. Phys. Usp. **159**, 621 (1989).
[3] B. Batlogg *et al.*, Physica C **135–140** 130 (1994); C. Y. Chen *et al.*, Phys. Rev. B **51**, 3671 (1995).
[4] G. M. Éliashberg, in *Physical Properties of High-Temperature Superconductors*, D. M. Ginzberg (ed.), World Scientific, Singapore (1989).
[5] M. V. Sadovskii, Sverkhprovodimost': Fizika, Khimiya, Tekhnologiya **3** 337 (1995).
[6] K. Tamasaku and S. Uchida, Physica C **235-24**, 1321 (1994); K. Kitazawa, Physica C **235-24** xxiii (1994).
[7] S. Uchida, T. Ido, H. Takagi, T. Arima, Y. Tokura, and S. Tajima, Phys. Rev. B **43** 7942 (1991).
[8] Z.-X. Shen, J. W. Allen, J. J. Yeh, J.-S. Kang, W. Ellis, W. Spicer, I. Lindau, M. B. Maple, Y. D. Dalichaouch, M. S. Torikachvili, J. Z. Sun, and T. H. Geballe, Phys. Rev. B **36**, 8414 (1987).
[9] H. Matsuyama, T. Takahashi, H. Katayama-Yoshida, T. Kashiwakura, Y. Okabe, and S. Sato, Physica C **160**, 567 (1989).
[10] J. W. Allen, C. G. Olson, M. B. Maple, J.-S. Kang, L. Z. Liu, J.-H. Park, R. O. Anderson, W. P. Ellis, J. T. Markert, Y. Dalichaouch, and R. Liu, Phys. Rev. Lett. **64**, 595 (1990).
[11] T. Takahashi, H. Matsuyama, H. Katayama-Yoshida, K. Seki, K. Kamiya, and H. Inokuchi, Physica C **170**, 416 (1990).
[12] C. C. Homes and T. Timusk, Phys. Rev. Lett. **71**, 1645 (1993); A. V. Puchkov, D. N. Basov, and T. Timusk, J. Phys.: Condens. Matter **8**, 10049 (1996).
[13] Y. Kim, A. Heeger, L. Acedo, G. Stucky, and F. Wudl, Phys. Rev. B **36**, 14034 (1987); G. Yu, C. Lee, A. Heeger, and S.-W. Cheong, Physica C **203**, 419 (1992).
[14] R. O. Anderson, R. Claessen, J. W. Allen, C. G. Olson, C. Janowitz, L. Z. Liu, J.-H. Park, M. B. Maple, Y. Dalichaouch, M. C. de Andrade, R. F. Jardim, E. A. Early, S.-J. Oh, and W. P. Ellis, Phys. Rev. Lett. **70**, 3163 (1993).
[15] C. Quitmann, J. Ma, R. J. Kelly *et al.*, Physica C **235-24**, 1019 (1994).
[16] A. Yurgens, D. Winhkler, N. V. Zavaritsky, and T. Claeson, Phys. Rev. Lett. **79**, 5122 (1997).
[17] H. S. Somal, B. J. Feenstra, J. Schützmann, J. Hoon Kim, Z. H. Barber, V. H. M. Duijn, N. T. Hien, A. A. Menovsky, M. Palumbo, and D. van der Marel, Phys. Rev. Lett. **76**, 1525 (1996).
[18] A. S. Alexandrov, V. V. Kabanov, and N. F. Mott, Phys. Rev. Lett. **77**, 4796 (1996).
[19] A. S. Alexandrov, Physica C **274**, 237 (1997).
[20] H. A. Blackstead, J. D. Dow, and D. B. Pulling, Physica C **265**, 143 (1996).
[21] J. C. Phillips, Physica C **221**, 327 (1994).
[22] P. W. Anderson, Science **235**, 1196 (1987); J. Phys.: Condens. Matter **8**, 10083 (1996).
[23] Yu. A. Izyumov, Usp. Fiz. Nauk **161**, 2 (1991) [Sov. Phys. Usp. **34**, 361 (1991)]; Usp. Fiz. Nauk **165**, 403 (1995) [Phys. Usp. **38**, 385 (1995)].
[24] N. M. Plakida, *High T_c Superconductivity*, Springer-Verlag, Berlin (1995).
[25] V. J. Emery and S. A. Kivelson, Phys. Rev. Lett. **74**, 3253 (1995).
[26] Q. Si, J. Phys.: Condens. Matter **8**, 9953 (1996).
[27] L. A. Openov, V. F. Elesin, and A. V. Krasheninnikov, Physica C **257**, 53 (1996).
[28] J. M. Tranquada, W. J. L. Buyers, H. Chou, T. E. Mason, M. Sato, S. Shamoto, and G. Shirane, Phys. Rev. Lett. **64**, 800 (1990).
[29] P. Monthoux and D. Pines, Phys. Rev. B **49**, 4261 (1994).
[30] B. P. Stojkovic and D. Pines, Phys. Rev. B **56**, 11931 (1997).
[31] M. Kohen, G. Gladstone, J. Jensen, and J. Schriefer, in *Superconductivity*, ed. by R. Parks, Marcel Dekker, New York (1969).
[32] B. T. Geilikman, Usp. Fiz. Nauk **88**, 327 (1966) [Sov. Phys. Usp. **9**, 142 (1966)].
[33] D. Pines, Phys. Rev. **109**, 280 (1958).
[34] A. I. Agafonov and E. A. Manykin, Phys. Rev. B **52**, 14571 (1995); Zh. Éksp. Teor. Fiz. **109**, 1405 (1996) [JETP **82**, 758 (1996)].
[35] A. I. Agafonov and E. A. Manykin, JETP Lett. **65**, 439 (1997).
[36] A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskii, *Methods of Quantum Field Theory in Statistical Mechanics* [in Russian], Fizmatgiz, Moscow (1962).
[37] F. Yonezawa and T. Matsubara, Prog. Theor. Phys. **35**, 357 (1966).

Translation provided by the Russian Editorial office.

Softening of phonon modes in C_{60} crystals induced by laser irradiation: Thermal effects

K. P. Meletov*)

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

E. Liarokapis

Physics Division, National Technical University, Athens 157 80, Greece

J. Arvanitidis, K. Papagelis, and S. Ves

Physics Department, Aristotle University of Thessaloniki, GR-540 06 Thessaloniki, Greece

G. A. Kourouklis

School of Technology, Aristotle University of Thessaloniki, GR-540 06 Thessaloniki, Greece

(Submitted 13 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1785–1794 (November 1998)

Reversible softening of the intramolecular $A_g(2)$ pentagonal pinch (PP) mode of a C_{60} single crystal in the face centered cubic phase has been studied as a function of laser power density by means of Raman scattering. The average temperature rise in the laser excitation spot has been determined using the Stokes to anti-Stokes integrated peak intensity ratio for the $H_g(1)$ phonon mode. Softening of the PP-mode was found to be due to heating of the sample resulting from laser irradiation, in good quantitative agreement with experimental results obtained for uniformly heated samples. These findings are in excellent agreement with results obtained by numerical calculations of the local temperature distribution and average temperature in the laser spot based on calculated integrated intensities of the Stokes and anti-Stokes bands of the PP-mode. These calculations were based on experimental data for the temperature dependence of phonon frequency and width, absorbance, and thermal conductivity in solid C_{60}. [S1063-7761(98)01611-4]

1. INTRODUCTION

Raman scattering has been, since the initial discovery of the fullerene family of compounds, a very useful tool for their characterization. In particular, the response of the $A_g(2)$ PP-mode to a variety of perturbations has been used to probe many diverse properties of solid fullerene and fullerene-based materials. These include temperature- and pressure-induced orientation-ordering phase transitions and effects due to intercalation of solid C_{60} with alkali metals.[1–5] Raman scattering has also been used to study photodimerization observed in solid C_{60} under conditions of intense laser illumination, and dimerization caused by high pressure and temperature treatment.[6,7] The latter effects are clearly manifested by the considerable softening of the $A_g(2)$ PP-mode.

The frequency of the $A_g(2)$ mode initially reported by Bethune *et al.*[8] for the room-temperature Raman spectrum of air-exposed C_{60} films is 1469 cm^{-1}. It was also reported that the room-temperature Raman spectrum of oxygen-free C_{60} contains a broad peak at 1459 cm^{-1}, which is more intense than the 1469-cm^{-1} peak. Exposure of the sample to oxygen leads to recovery of the 1469-cm^{-1} peak.[9] The 1459-cm^{-1} peak in the Raman spectrum of C_{60} was explained by Rao *et al.*[6] as a manifestation of the photoassisted dimerization of oxygen-free C_{60} films under intense laser illumination. It has also been shown that oxygen-exposed C_{60} films are more resistant to laser irradiation and need considerably higher laser power densities to initiate the photodimerization reaction.[6,10] The softening of the PP-mode associated with the photodimerization of C_{60} is irreversible. The phototransformed material is stable at room temperature and can be recovered only upon heating to temperature greater than 420 K. When the laser illumination level is below the photodimerization threshold, the PP-mode exhibits reversible softening down to 1461 cm^{-1}.[9–11]

Reversible softening of the PP-mode caused by laser illumination has attracted special interest in Raman scattering studies of solid C_{60}, especially concerning its origin.[12] Raman experiments performed at 40 K have divulged the existence of a second wide band downshifted by $\sim$3 cm^{-1} from the 1469-cm^{-1} peak at a laser power density $\sim$50 W/cm^2.[12] Increasing the laser power density leads to gradual softening and enhancement of the intensity of this band. At the same time, the intensity of the 1469-cm^{-1} peak goes down, and disappears at a laser power density of $\sim$300 W/cm^2 without any detectable change in peak position. A softening of the new band continues as the power density increases, and becomes irreversible at laser power densities exceeding 500 W/cm^2. Splitting and softening of the PP-mode is related to the high concentration of molecules in the lowest excited triplet state resulting from the high absorbance of laser radiation, high singlet-triplet intersystem crossing, and the

relatively high lifetime of the triplet state.[11–13] It is assumed under these conditions that each intramolecular phonon mode will split into two Raman components that correspond to the PP-mode frequencies in the ground and excited electronic states of the C_{60} molecule.[12]

This assumption, in our opinion, must be examined in light of the photophysics of large molecular systems. It is well established that the frequency of any intramolecular phonon mode of a large molecule, for example an aromatic hydrocarbon, is higher in the electronic ground state than in the excited states.[14] The frequency difference determined via vibrational analysis of the electron-phonon bands in the luminescence and absorption spectra varies from 5 to 10% of the corresponding ground state frequency. Moreover, the frequencies of the corresponding modes in the ground and excited states of a molecule have fixed values and do not depend on the populations of the various states, which are, in turn, related to the laser power density. In C_{60}, the frequency difference of the split Raman peaks[12] is very small in comparison to the PP-mode frequency, and the change in frequency can be attributed to population of the excited triplet state.[12]

From another point of view, laser illumination heats the sample, which, as a rule, is the main reason for phonon mode softening. The heating of solid C_{60} under laser irradiation, even at low laser power densities, is due to the relatively low thermal conductivity of this material.[15] This has been proposed as an alternative mechanism for PP-mode softening.[9,10]

The experimental study of sample heating due to laser irradiation involves a rather straightforward procedure, and is based on analysis of Raman peak intensities in the Stokes and anti-Stokes regions of the spectrum. We have performed a detailed study of the softening of the PP-mode of solid C_{60} as a function of the laser power. Our motivation was to examine the relationship between sample heating in the laser spot and PP-mode softening using the relationship between the total intensities of the Stokes $I_S(\omega)$ to anti-Stokes $I_{AS}(\omega)$ bands,

$$\frac{I_S(\omega)}{I_{AS}(\omega)}\left[\frac{\omega_L+\omega}{\omega_L-\omega}\right]^3=\exp\left(\frac{\hbar\omega}{k_BT}\right), \tag{1}$$

where T is the mean temperature in the laser spot, ω_L is the laser frequency, and ω is the phonon frequency, appropriately corrected for the ω^3 scattering efficiency factor.[16] Effects related to the frequency dependence of the optical coefficients have been neglected, as we are far enough from resonance conditions. We have measured detailed Raman spectra of C_{60} single crystals in the low-energy Stokes and anti-Stokes regions, as well as in the high-energy region where the PP-mode is located, at room temperature, and at various laser power densities. The results clearly indicate considerable overheating of the sample in the laser spot. They agree well with results on the uniform bath temperature dependence of the PP-mode frequency. We have also performed numerical calculations of the local temperature distribution and the mean temperature in the laser spot from the calculated integrated intensities of the Stokes and anti-Stokes bands of the PP-mode. These calculations were based on experimental data for the absorbance and thermal conductivity of solid C_{60}, and are in good agreement with the experimental results.

2. EXPERIMENTS

Single crystals of fullerite were grown from a solution of C_{60} in toluene. The primary C_{60} material, with purity better than 99%, was obtained by the Krätschmer method.[17] Data were recorded on crystals in the form of thin platelets with well-developed specular surfaces and dimensions $\sim 300 \times 300 \times 50\ \mu m^3$. The uniform-temperature data were taken using a nitrogen gas flow cryostat, for bath temperatures up to 470 K. In this case the samples were glued to the finger tip using a high-temperature and high thermal conductivity glue. Room-temperature data were taken on freely located air-exposed samples.

Raman spectra were recorded using a triple monochromator (DILOR XY-500) equipped with a CCD cryogenic detector system. The spectral width of the system was $\sim 5\ cm^{-1}$. The 514.5-nm line of an Ar^+ laser was used for excitation. The laser beam was focused to a spot either $\sim 1.25\ \mu m$ in diameter using an Olympus 100× objective, or $\sim 7.5\ \mu m$ in diameter using a Nikon 20× objective with a flux adapter. The laser spot diameter and half-width are critical parameters, along with the laser power, in assessing and comparing experimental results on laser-induced effects. Throughout this paper, we have adopted as the laser spot diameter and half-width the values measured at the 10% and $1/e$ intensity values relative to the peak, respectively. The spectra were recorded in back-scattering geometry using a $\lambda/4$ plate as a scrambler and an Olympus microscope system for image processing. The laser power at the sample varied from 0.06 to 0.3 mW for a laser spot diameter $\simeq 1.25\ \mu m$, and from 0.4 to 1.5 mW for a laser spot diameter $\simeq 7.5\ \mu m$. The data for temperature dependence were recorded at the lowest laser power necessary for recording spectra to minimize the effects of laser irradiation, and the temperature was stabilized for a long time to ensure uniformity over the entire sample volume.

Peak positions and total intensities were determined by fitting Lorentzians to the experimental data. The accuracy of the peak positions was about $0.25\ cm^{-1}$. To eliminate systematic errors in the peak positions, the experimental setup was calibrated before every measurement using the 17976.7-cm^{-1} plasma line of a Ne lamp, which is located near the PP-mode spectral position. The accuracy of the total peak intensities was limited by the scatter in the background values, and the error was therefore estimated to be no more than 10% of the intensity of the weakest anti-Stokes band. The temperature stabilization accuracy during uniform temperature measurements was ~ 1 K.

3. RESULTS AND DISCUSSION

The Raman spectrum of C_{60} single crystals, taken at room temperature and normal pressure, contains ten main intramolecular modes: $H_g(1)$-$H_g(8)$ and $A_g(1)$, $A_g(2)$. Their frequencies are very close to those previously determined: the differences do not exceed 2–3 cm^{-1}.[1,8] In addi-

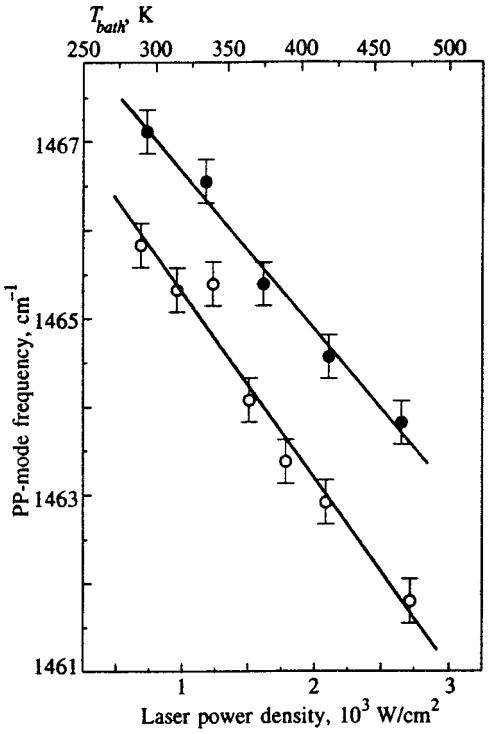


FIG. 1. Dependence of the PP-mode frequency of solid C_{60} on laser power density at room temperature, laser spot diameter ~7.5 μm (open symbols). Solid symbols show the dependence of the PP-mode frequency of solid C_{60} on the uniform temperature of the sample, T_{bath} for a fixed laser power density ~200 W/cm^2. Solid lines are linear least-square fits to the experimental data.

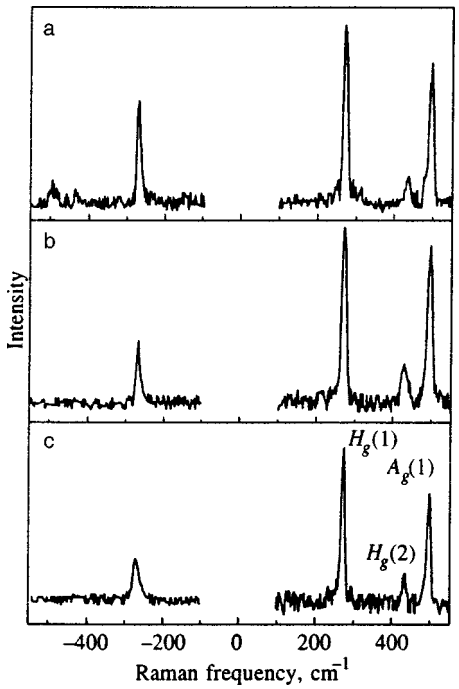


FIG. 2. Raman spectra of solid C_{60} in the low-energy Stokes and anti-Stokes regions at room temperature, for various laser power levels *W*: a) 0.3 mW, b) 0.15 mW, c) 0.06 mW. Laser spot diameter ≃1.25 μm, T_{bath}=300 K, $T_{\text{spot}}^{\text{exp}}$ is the mean temperature inside the laser spot determined from Eq. (1): a) $T_{\text{spot}}^{\text{exp}}$=529 K; b) $T_{\text{spot}}^{\text{exp}}$=401 K; c) $T_{\text{spot}}^{\text{exp}}$=340 K.

tion to these bands there are some very weak Raman peaks that may originate in second-order scattering.[18] The most intense $A_g(2)$ PP-mode, located under normal conditions at ≃1467.3 cm^{-1}, corresponds to the out-of-phase stretching of pentagonal and hexagonal carbon rings.

The frequency of all Raman peaks is sensitive to the laser power density: the majority of the modes soften when the laser power increases. The dependence of the PP-mode frequency on laser power density is shown in Fig. 1 by open symbols. The initial value of the PP-mode frequency ≃1465.8 cm^{-1} at laser power density ≃600 W/cm^2 decreases linearly to ≃1461.8 cm^{-1} as the laser power density increases to ≃2700 W/cm^2. The extrapolation of this dependence to vanishing laser power density yields a frequency for the PP-mode of ≃1467.5 cm^{-1}. This is close to the highest experimental value of 1467.3 cm^{-1} observed at minimal laser power density ≃200 W/cm^2.

The softening of the PP-mode under laser illumination is reversible when the laser power density increases to ≃3000 W/cm^2 (for an exposure time of ≃600 sec). At higher laser power densities it becomes irreversible, which case visible damage of the crystal surface at the illumination spot region is observed. The solid symbols in Fig. 1 show the dependence of the PP-mode frequency on the bath temperature, T_{bath}, for the uniformly heated samples. The laser power density for this measurement was kept constant at the lowest level, ≃200 W/cm^2, which corresponds, as will be shown, to a local temperature rise of ≃10 K and a shift in phonon frequency of ≃0.2 cm^{-1}. These values are close to the experimental accuracy in the peak position and estimated temperature in the laser spot. At room temperature, the frequency of the PP-mode is ≃1467.3 cm^{-1}, and it decreases to 1463.7 cm^{-1} at 470 K. The temperature dependence of the PP-mode frequency is very similar to its dependence on laser power density. This is a clear indication that softening of the PP-mode under laser illumination may be related to local overheating of the sample in the laser spot.

Figure 2 shows Raman spectra of the C_{60} crystal taken in the Stokes and anti-Stokes regions at three different laser power levels. The spectrum in the Stokes region contains three intramolecular phonon modes, $H_g(1)$, $H_g(2)$, and $A_g(1)$, with frequencies 273, 435, and 495 cm^{-1}, respectively. In the anti-Stokes region, all three spectra contain the prominent $H_g(1)$ Raman peak. The other two peaks are not detectable at laser powers 0.06 and 0.15 mW; they only become visible if the spectrum is recorded at laser power ≃0.3 mW. The peak intensities of the Raman bands in the Stokes region are essentially the same in all three spectra, whereas in the anti-Stokes region the intensity of $H_g(1)$ peak increases noticeably as the laser power increases. This is a clear indication that the temperature of the sample within the laser illumination spot gradually increases with laser power.

A comparison of the total intensities of the $H_g(1)$ Raman peak in the Stokes and anti-Stokes regions on the basis of Eq. (1), using temperature *T* as a fitting parameter, yields the average temperature $T_{\text{spot}}^{\text{exp}}$ in the spot. The data reveal considerable overheating of the sample within the laser illumination spot. The temperature $T_{\text{spot}}^{\text{exp}}$ reaches ≃530 K at laser power 0.3 mW and spot diameter ≃1.25 μm, which is about 270 K lower than the heater temperature for the sublimation

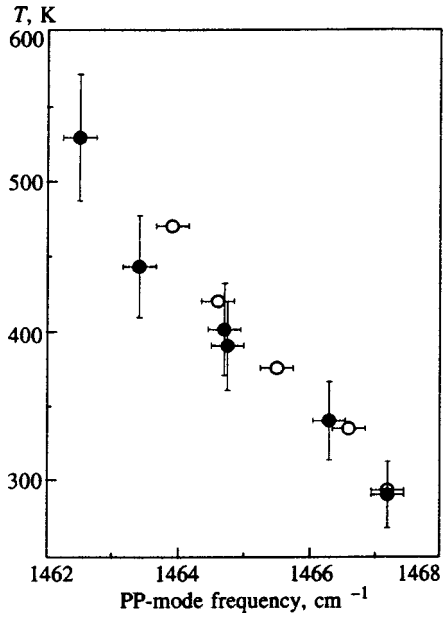


FIG. 3. Dependence of PP-mode frequency of solid C_{60} on the uniform temperature of the sample, T_{bath} (open symbols), and the local temperature inside the laser excitation spot $T_{\text{spot}}^{\text{exp}}$ (solid symbols).

of C_{60} powder during vapor growth of fullerite single crystals. These data are compatible with visually observed damage of the crystal surface at high laser powers, when sublimation of material and crater creation takes place due to extreme overheating of the material. As discussed below, the strong temperature rise is due to very low thermal conductivity, which strongly localizes the effect of the laser irradiation.

Figure 3 shows the dependence of PP-mode frequency on the uniform temperature of the sample T_{bath} (open symbols), and on the mean temperature $T_{\text{spot}}^{\text{exp}}$ in the laser spot at various laser powers, determined from Eq. (1) (solid symbols). The agreement between these data is within the experimental errors for the $T_{\text{spot}}^{\text{exp}}$ determination, which varies from 20 to 50 K in the various measurements. The results indicate that the temperature of the excited crystal region inside the laser spot, $T_{\text{spot}}^{\text{exp}}$, determined as described above, is significantly higher than room temperature, and increases with laser power. They also mean that the dominant effect of laser illumination is to overheat the sample inside the laser spot, and there is no need to turn to the excited triplet state of C_{60} to explain softening of the PP-mode. The relatively high overheating of fullerite with respect to other solids is related primarily to the relatively low thermal conductivity of this material.[15]

It can be shown that the resulting laser overheating is compatible with experimental data on light absorbance and thermal conductivity of fullerite.[15,19] As a check, we have calculated the overheating temperature distribution $\Delta T(R,Z)$ inside the laser spot from the steady-state solution given originally by Lax[20] for constant thermal conductivity:

$$\Delta T(R,Z)=\frac{I_0 wA(1-\Re)}{K_0}\times\int_0^\infty d\lambda J_0(\lambda R)F(\lambda)\,\frac{Ae^{-\lambda Z}-\lambda e^{-AZ}}{A^2-\lambda^2}, \quad (2)$$

where $\Re$ is the reflectivity, K_0 is the thermal conductivity at room temperature, and $R=r/w$, $Z=z/w$, $A=\alpha w$ are dimensionless parameters for the radius τ, the depth z, and the absorption coefficient α. In the above equation, $J_0(\lambda R)$ is the zeroth-order Bessel function and $F(\lambda)$ is the corresponding Bessel transform of the laser beam profile, which is assumed to be Gaussian, and $I(r)=I_0\exp(-r^2/w^2)$, where w is the beam half-width at $1/e$ of the maximum intensity I_0. In this axial symmetry, the temperature profile depends only on the radius r and depth z.

In the above solution, the material is considered semi-infinite, an approximation which is clearly not valid in our case, as the crystallites are of small dimensions. But due to the very low thermal conductivity, the calculations prove that heating is localized, and that it very closely follows the laser beam profile.

The temperature dependence of the thermal conductivity[21] and absorption coefficient[22] has also been considered in the literature. In the present case, the thermal conductivity can be considered constant at $0.4\ \text{W}\cdot\text{m}^{-1}\cdot\text{K}^{-1}$ above room temperature,[15] although more recently even lower values have been reported for the room-temperature thermal conductivity.[23] The optical coefficients can also be considered constant (absorption 2.7 μm^{-1}[19] and reflectivity 0.19[24]) in the temperature range under consideration. Those two assumptions simplify the calculations considerably. The beam half-width w was calculated from the value given by the manufacturer of the microscope objective for the excitation wavelength at optimum focusing and for 90% of the total intensity. As all measurements were obtained with the microscope very carefully focused, we can assume that the actual values are very close to optimal, i.e., $w=0.4$ μm for 100× magnification and 2.5 μm for 20× magnification. In Fig. 4 we present the temperature distribution obtained from Eq. (2) for several laser beam powers and spot diameters $d=1.25$ μm (a) and 7.5 μm (b).

The Raman spectra are given by the convolution of the scattering from volumes $(2\pi r\,dr\,dz)$ of circular rings of equal temperature at the laser spot. Based on the temperature distribution along the r and z-axis obtained from Eq. (2), the Stokes spectra can be calculated from the cross section,[16] neglecting the frequency dependence of the second-order susceptibility and the optical constants, as we are far from resonance, i.e.,

$$\frac{dI_s}{d\omega}\propto\int_0^\infty dz\int_0^\infty dr(2\pi r)\times(1+\eta)\,\frac{(\omega_L-\omega)^3}{\omega_{ph}}\,\frac{\Gamma/2}{(\omega-\omega_{ph})^2+\Gamma^2/4}\,I(r)e^{-\alpha z}, \quad (3)$$

where the phonon full FWHM $\Gamma\approx 5.1+0.003(T-300)$ cm^{-1} and the phonon frequency $\omega_{ph}\approx 1467$

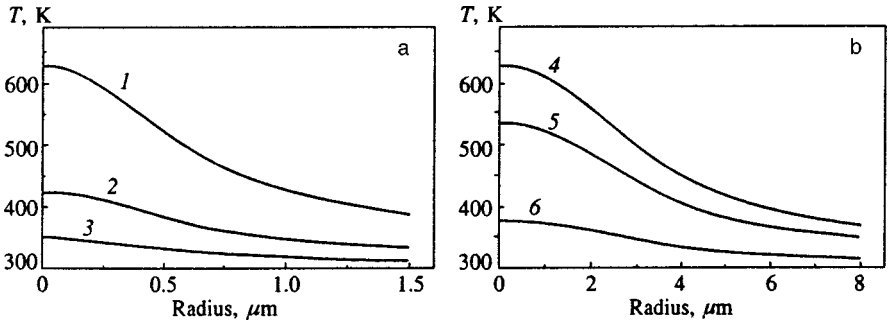


FIG. 4. Temperature distribution as a function of radius inside the laser spot calculated from Eq. (2) and assuming ambient sample temperature $T_0=300$ K and zero depth ($z=0$). a) Laser spot diameter $d=1.25$ μm and laser power is (*1*) 0.4; (*2*) 0.15, and (*3*) 0.06 mW; b) laser spot diameter $d=7.5$ μm and laser power is (*4*) 1.6; (*5*) 1.15, and (*6*) 0.38 mW.

$-0.02(T-300)$ cm^{-1} both depend on the temperature T in a way defined by the uniform heating (Fig. 1). The statistical factor η is defined in Ref. 16

$$\eta(T)=\frac{1}{\exp(\hbar\omega/k_BT)-1}. \tag{4}$$

For the anti-Stokes component, the factor $(1+\eta)(\omega_L-\omega)^3$ in Eq. (3) must be replaced by $\eta(\omega_L+\omega)^3$.

Based on these calculations, the peak position, width, average temperature, and ratio of the Stokes to anti-Stokes total intensities for the $A_g(2)$ PP-mode can be obtained for every temperature distribution. The calculated average temperature in the spot, $T_{\rm spot}^{\rm calc}$ as obtained from the calculated ratio of the Stokes to anti-Stokes total intensities for the PP-mode (corrected for the ω^3 dependence) are shown in Fig. 5 for $d=7.5$ μm. In the same figure, the data points indicate the experimental results for the average temperature $T_{\rm spot}^{\rm exp}$ obtained from the ratio of the Stokes to anti-Stokes components of the $H_g(1)$ phonon mode, also corrected for the ω^3 dependence. The agreement between the theoretical predictions and the experimental values is remarkable, despite the fact that no adjustable parameters were used in the calculations. This is a strong indication that the modifications induced by laser irradiation in the spectra are due to local heating, which can raise the temperature to 600 K at the center of the laser spot (zero depth), and a power of $\simeq$1.6 mW, as our calculations indicate for the value of $T(0,0)$ (shown by the dashed line in Fig. 5 and calculated from Eq. (2) for $r=z=0$).

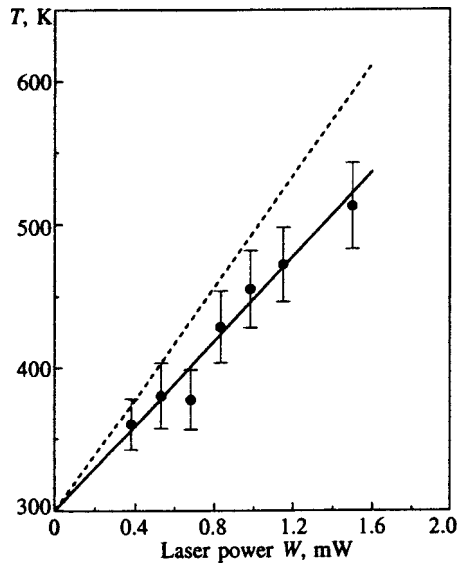


FIG. 5. Temperature in the laser excitation spot as a function of laser power for laser spot diameter $\simeq$7.5 μm. Solid symbols are experimental data for $T_{\rm spot}^{\rm exp}$ determined from the Stokes to anti-Stokes total intensity ratio for the $H_g(1)$ phonon mode. Solid line is the mean temperature $T_{\rm spot}^{\rm calc}$ calculated using Eq. (2) and Eq. (3). Dashed line is the temperature $T(0,0)$ at the center of the laser excitation spot and zero depth ($r=0$, $z=0$), calculated using Eq. (2).

To compare results obtained under various experimental conditions, namely the excitation spot diameter d and laser power density P, we found that the results for the average temperature over the spot can be described by the approximate expression

$$\langle T\rangle_{S/AS}^{\rm appr}\simeq T_0+1116\,\frac{W}{d}, \tag{5}$$

where the temperature is given in K, T_0 is ambient temperature, the laser power W is in mW, and d is in μm. Taking into account that $W=P\pi d^2/4$, expression (5) can be written

$$\langle T\rangle_{S/AS}^{\rm appr}\simeq T_0+876Pd. \tag{6}$$

Equation (6) implies that at larger spot diameters, one needs lower laser power densities to achieve the same average temperature rise. This result is consistent with our experimental data obtained for sample overheating at two different spot diameters, and is a very important consideration in trying to compare results obtained under different experimental conditions.[6,8–12] Equation (6) makes it clear that the power density is not sufficient to compare results on laser-induced overheating of materials with low thermal conductivity.

The calculated temperature profiles inside the laser spot show a considerable difference between the temperature at the center and periphery of the spot. The difference, starting with $\simeq$30 K at $W=0.2$ mW, reaches $\simeq$220 K at $W=1.6$ mW (see Fig. 4). The highly nonuniform temperature distribution inside the spot results from the low thermal conductivity of the material. This may be the reason for the splitting of the PP-mode in the Raman spectra taken at temperatures lower than 250 K, the temperature of the orientational-ordering phase transition from fcc to sc structure.[25] According to van Loosdrecht *et al.*,[1] the temperature dependence of the PP-mode frequency exhibits a jump of about 4 cm^{-1} at 250 K. We remark that in this case the transition temperature of 250 K might be located between the maximum and minimum temperatures within the excitation spot. This means that for a range of laser power densities within the excitation spot, one can have both the low and

high temperature phases. The Raman spectra taken under these conditions are expected to show peaks of the PP-mode coming from the two phases, with intensities proportional to the scattering volume corresponding to each phase. At lower and higher laser power densities, the area inside the excitation spot will correspond to a single phase of the material in the low or high temperature phase, respectively. It must be said that this is not exactly the case in Ref. 12, because the laser power densities used are not sufficient to produce the considerable temperature rise within the laser excitation spot. We also think that it might be interesting to extend the measurements of local overheating within the laser spot to the low-temperature region, taking special precautions in the focusing of the laser beam and measuring its power.

4. CONCLUSIONS

Reversible softening of the PP-mode in the Raman spectra of fullerite C_{60} under laser illumination is related to overheating of the material inside the laser spot. Estimates of the mean temperature inside the laser spot on the basis of experimental Raman spectra in the Stokes and anti-Stokes regions show that it can reach as high as 530 K at laser power $\simeq 0.3$ mW and laser spot diameter $\simeq 1.25$ μm. The experimental dependence of PP-mode frequency on the uniform temperature of the sample agrees well with the dependence of the PP-mode position on the mean temperature in the laser spot. Numerical calculations of the local temperature in the laser spot, based on experimental measurements of laser power density, optical absorption, and thermal conductivity of solid C_{60}, are in good quantitative agreement with the experimental results.

This work was partially supported by the General Secretariat for Research and Technology, Greece (Grant No. 96-1214), the Russian Fund for Fundamental Research (Grant No 96-02-17489), the Russian State Program ''Fullerenes and Atomic Clusters,'' and NATO HTECH. CRG No. 972317. K. P. Meletov acknowledges the support and hospitality of the Laboratory of Physics, Physics Division, School of Technology, Aristotle University of Thessaloniki, during the course of this work.

*)E-mail: mele@issp.ac.ru

[1] P. H. M. van Loosdrecht, P. J. M. van Bentum, and G. Meijer, Phys. Rev. Lett. **68**, 1176 (1992).
[2] N. Chandrabhas, M. N. Shashikala, D. V. S. Muthy, A. K. Sood, and C. N. R. Rao, Chem. Phys. **197**, 319 (1992).
[3] K. P. Meletov, D. Christofilos, G. A. Kourouklis, and S. Ves, Chem. Phys. **236**, 265 (1995).
[4] Ping Zhou, Kai-An Wang, Ying Wang, P. C. Eklund, M. S. Dresselhaus, G. Dresselhaus, and R. A. Jishi, Phys. Rev. B **46**, 2595 (1992).
[5] J. Winter and H. Kuzmany, Solid State Commun. **84**, 935 (1992).
[6] A. M. Rao, P. Zhou, K.-A. Wang *et al.*, Science **259**, 955 (1993).
[7] P.-A. Persson, U. Edlund, P. Jacobsson, D. Johnels, A. Soldatov, and B. Sundqvist, Chem. Phys. **258**, 540 (1996).
[8] D. S. Bethune, G. Meijer, W. C. Tang, H. J. Rosen, W. G. Golden, H. Seki, C. A. Brown, and M. S. de Vries, Chem. Phys. **179**, 181 (1991).
[9] Y. Hamanaka, S. Nakashima, M. Hangyo, H. Shinohara, and Y. Saito, Phys. Rev. B **48**, 8510 (1993).
[10] J. Sauvajol, F. Brocard, Z. Hricha, and A. Zahab, Phys. Rev. B **52**, 14839 (1995).
[11] P. H. M. van Loosdrecht, P. J. M. van Bentum, M. A. Verheijen, and G. Meijer, Chem. Phys. **198**, 587 (1992).
[12] P. H. M. van Loosdrecht, P. J. M. van Bentum, and G. Meijer, Chem. Phys. **205**, 191 (1993).
[13] Ping Zhou, Zheng-Hong Dong, A. M. Rao, and P. C. Eklund, Chem. Phys. **211**, 337 (1993).
[14] V. L. Broude, E. I. Rashba, and E. F. Sheka, *Spectroscopy of Molecular Excitons* [in Russian], Energoizdat, Moscow (1981).
[15] R. C. Yu, N. Tea, M. B. Salamon, D. Lorents, and R. Malhotra, Phys. Rev. Lett. **68**, 2050 (1992).
[16] W. Hayes and R. Loudon, *Scattering of Light by Crystals*, John Wiley & Sons, New York (1978).
[17] W. Krätschmer, K. Fostiropoulos, and D. Huffman, Chem. Phys. **170**, 167 (1990).
[18] Zheng-Hong Dong, Ping Zhou, J. M. Holden, P. C. Eklund, M. S. Dresselhaus, and G. Dresselhaus, Phys. Rev. B **48**, 2862 (1993).
[19] K. P. Meletov, V. K. Dolganov, O. V. Zharikov, I. W. Kremenskaya, and Yu. A. Ossip'yan, J. Phys. I **2**, 2097 (1992).
[20] M. Lax, J. Appl. Phys. **48**, 3919 (1977).
[21] M. Lax, Appl. Phys. Lett. **33**, 786 (1978).
[22] E. Liarokapis and Y. S. Raptis, J. Appl. Phys. **57**, 5123 (1985).
[23] O. Andersson, A. Soldatov, and B. Sundquist, Phys. Rev. B **54**, 3093 (1996).
[24] M. Patrini, F. Marabelli, G. Guizzetti, M. Manfredini, C. Castoldi, and P. Milani, in *Recent Advances in the Chemistry and Physics of Fullerenes and Related Materials*, ed. by K. Kadish and R. Ruoff, The Electrochemical Society Inc., Pennington (1994), p. 632.
[25] P. A. Heiney, J. E. Fisher, A. R. McGhie, W. J. Romanow, A. M. Denenstein, J. P. McCauley, A. B. Smith, and D. E. Cox, Phys. Rev. Lett. **66**, 2911 (1991).

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

Optical transitions and cyclotron resonance at Landau levels split by a periodic potential

V. Ya. Demikhovskiĭ*) and A. A. Perov†)

N. I. Lobachevskiĭ Nizhegorod State University, 603600 Nizhniĭ Novgorod, Russia
(Submitted 23 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1795–1803 (November 1998)

The structure of the electron spectrum is investigated and selection rules are found for transitions between magnetic subbands in a surface 2D superlattice of quantum dots in a perpendicular magnetic field. The photon absorption probabilities are calculated, and the profiles of the absorption lines are determined for allowed and forbidden direct dipole transitions between subbands split off from different Landau levels. © *1998 American Institute of Physics.* [S1063-7761(98)01711-9]

1. INTRODUCTION

The quantum states of Bloch electrons in a magnetic field have intrigued theoreticians[1–5] and experimentalists[6–8] for several decades now. So far, however, no one has observed phenomena to corroborate the existence of a split structure of Landau levels. This deficit is attributable to the fact that the observation of such phenomena in real crystals calls for as yet unattainable magnetic fields of the order of 1000 T. On the other hand, the magnetic subbands of Bloch electrons might be observed in artificial crystals, i.e., surface two-dimensional (2D) superlattices situated in a perpendicular magnetic field. The last decade has witnessed progress in the creation of such semiconductor structures[9] with a mean free path significantly exceeding the period of the potential. More recently, first attempts have been undertaken to observe spectra of the type known as Hofstadter's ''butterfly'' (Ref. 4) in 2D superlattices by means of magnetoresistance measurements.[10] Another possibility for observing the structure of the spectrum experimentally is to investigate the magnetooptics of such structures. In this paper we report a numerical study of the absorption of electromagnetic radiation in arrays of quantum dots (dot lattices) in a magnetic field.

2. BASIC EQUATIONS AND COMPUTATIONAL METHOD

In the proposed model the Hamiltonian of an electron in a periodic 2D potential in a static magnetic field and in an electromagnetic field has the form $\hat{H}=\hat{H}_0+\hat{H}_{\rm int}$. The unperturbed Hamiltonian is interpreted here as the Hamiltonian of the electron in a perpendicular magnetic field and in the field of a periodic potential:

$$\hat{H}_0=\frac{1}{2m^*}\left(\hat{\mathbf{p}}+\frac{|e|\mathbf{A}_0}{c}\right)^2+V(x,y), \tag{1}$$

where the function $V(x,y)=V_0\cos^2(\pi x/a)\cos^2(\pi y/a)$ models the periodic potential of a square lattice of quantum dots, $\mathbf{A}_0=H(0,x,0)$ is the vector potential of the static magnetic field, and m^* is the effective mass of the electron. The eigenfunction of the Hamiltonian (1), subject to the generalized Bloch conditions in a magnetic field,

$$\psi_{\mathbf{k}}(x,y)=\psi_{\mathbf{k}}(x+qa,y+a)\exp(-ik_xqa)\times\exp(-ik_ya)\exp(-2\pi ipy/a), \tag{2}$$

can be written as a series in oscillatory functions $\varphi_N(x)$ along the x direction and plane waves in the y direction[11–13]:

$$\psi_{\mathbf{k}}^{N_0,n_0}(x,y)=\sum_{N=0}^{\infty}\sum_{n=1}^{p}C_{Nn}^{N_0,n_0}(\mathbf{k})\sum_{j=-\infty}^{+\infty}\varphi_N\times\left(\frac{x-x_0-jqa-nqa/p}{l_H}\right)\exp\left(ik_x\left[jqa+\frac{nqa}{p}\right]\right)\exp\left(2\pi iy\frac{jp+n}{a}\right)\exp(ik_yy). \tag{3}$$

The quantum numbers N_0 and n_0 define the magnetic subband: $n_0=1,p$ labels the subband split off from the N_0th Landau level. The parameter $p/q=|e|Ha^2/2\pi\hbar c$ (p and q are integers) is equal to the number of quanta of magnetic flux through a lattice unit cell of area a^2, and $l_H=\sqrt{c\hbar/|e|H}$ is the magnetic length. We previously[11–13] proposed a numerical method for calculating these functions. To find the spectrum and wave functions, we write the Hamiltonian H_0 in the representation of symmetrized functions associated with the coefficients $C_{Nn}^{N_0,n_0}(\mathbf{k})$ in Eq. (3). In this representation the eigenvalues $E_{N_0,n_0}(\mathbf{k})$ of the Hamiltonian (1) are found by diagonalizing its matrix.

The effect on the system of an electromagnetic wave propagating along the vector **H** and polarized linearly along the x axis is taken into account by perturbation theory. The perturbation Hamiltonian is

$$\hat{H}_{\rm int}=-i\frac{|e|\hbar}{m^*c}A_1\exp(-i\omega t)\frac{\partial}{\partial x}. \tag{4}$$

In Eq. (4) A_1 is the vector potential of the electromagnetic field.

The number of photons absorbed per unit time per unit surface area is

$$\alpha=\frac{2\pi}{\hbar\Phi}\int |P_{\mathbf{kk}}^{i\to f}|^2\delta[E_f(\mathbf{k})-E_i(\mathbf{k})-\hbar\omega]\frac{2}{(2\pi)^2}d^2k, \tag{5}$$

where Φ is the photon flux density. We assume here that the wave vectors of the electron in the initial and final states coincide, i.e., we are concerned with direct interband transitions. The integration in Eq. (5) is carried out over all occupied initial states.

It follows from Eqs. (3) and (4) that the direct dipole transition matrix element has the form

$$P_{\mathbf{kk}}^{i\to f}=A_1\frac{|e|}{qam^*c}\sum_{N,M=0}^{\infty}\sum_{n,s=1}^{p}C_{Ms}^{f*}(\mathbf{k})C_{Nn}^{i}(\mathbf{k})\times\sum_j\int_{-qa/2}^{qa/2}\varphi_M\left[\frac{x-x_0-jqa-nqa/p}{l_H}\right]\times\hat{p}_x\varphi_N\left[\frac{x-x_0-jqa-nqa/p}{l_H}\right]dx. \tag{6}$$

In strong magnetic fields, where $qa\gg\sqrt{N}\,l_H\sqrt{M}\,l_H$ and $V_0\ll\hbar\omega_c$, the oscillatory functions for all effective N and M in (6) are highly localized in the magnetic unit cell. Consequently, the contribution of terms with large j to the integral (6) are exponentially small. Restricting the equation to terms with $j=0$ and extending the limits of integration over x to infinity, we obtain

$$P_{\mathbf{kk}}^{i\to f}=-i\frac{|e|\hbar A_1}{qam^*cl_H\sqrt{2}}\sum_{N=0}^{\infty}\sum_{n,s=1}^{p}[C_{N-1,s}^{f*}(\mathbf{k})C_{Nn}^{i}\times(\mathbf{k})\sqrt{N}-C_{N+1,s}^{f*}(\mathbf{k})C_{Nn}^{i}(\mathbf{k})\sqrt{N+1}\,]. \tag{7}$$

Here i and f are specified by the set of two quantum numbers characterizing the magnetic Landau subband.

3. RESULTS AND DISCUSSION

Of primary concern here is the calculated structure of the electron spectrum. All the parameters of the periodic potential and the magnetic fields are given in the figure captions. Figure 1a shows the energy levels for $\mathbf{k}=0$, corresponding to the edges of the magnetic subbands. The number of magnetic flux quanta through unit cell for $p/q\geqslant 2$ labels the vertical axis. Clearly, the levels tend to bunch up toward the unperturbed Landau levels in this magnetic field range. In the range $p/q<2$ (not shown in the figure) magnetic Landau subbands are not formed, and Hofstadter butterfly spectra emerge.[4] We can see that the total width of the split Landau levels decreases as the number N increases. The spectrum of magnetic subbands split off from the zeroth ($N=0$) Landau level is shown in magnified scale in Fig. 1b, where the number of magnetic flux quanta is now incremented by one: $p/q=15/1$; 61/4; 31/2; 63/4; 16/1. The electron wave vector in each subband falls within the limits of the first magnetic Brillouin zone: $-\pi/qa\leqslant k_x\leqslant\pi/qa$; $-\pi/a<k_y<\pi/a$. It is evident that p nonoverlapping magnetic subbands are formed for $p/q=15/1$ and 16/1. The same subbands are formed under each Landau level. In the interval between integral values of p/q the number of subbands varies sharply with the magnetic field. Each subband existing for integral p/q splits into a series of subbands so that their total number under each Landau level is equal to p. The narrow subbands for $p/q=61/4$; 31/2; 63/4 are indistinguishable in Fig. 1b.

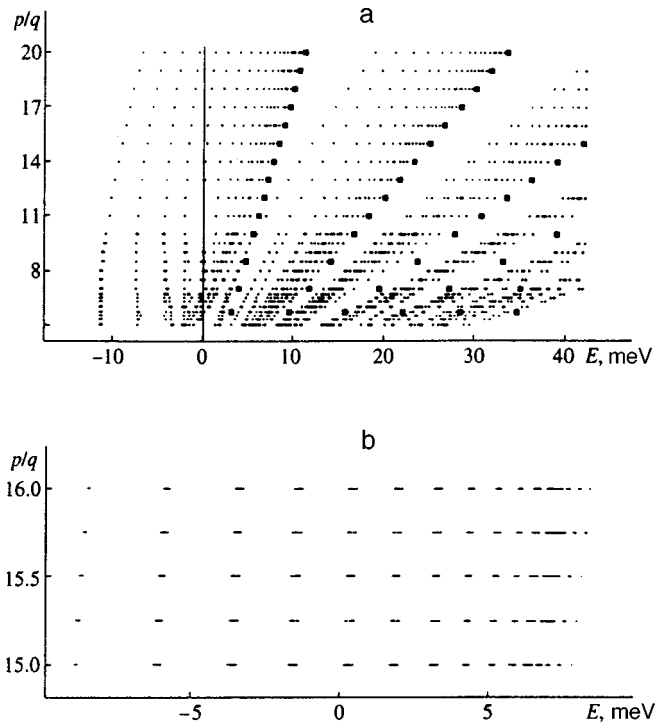


FIG. 1. a) Energy spectrum of a quantum dot lattice (Al, Ga)As ($a=80$ nm, $V_0=-20$ meV) in a magnetic field for $\mathbf{k}=0$. a) The positions of the Landau levels are indicated by heavy dots. b) Structure of the magnetic Landau subbands $(0, n)$ for $p/q=15/1$; 61/4; 31/2; 63/4; 61/1.

Figure 2 illustrates the distribution of widths of the Landau subbands for $p/q=15/1$. Clearly, the widths of subbands situated under one Landau level can differ by several orders of magnitude. The maximum width occurs for subbands near the middle of the split Landau level (see also Fig. 1b). The edges of these subbands should be experimentally resolvable in the absorption spectra. The widths of subbands situated at

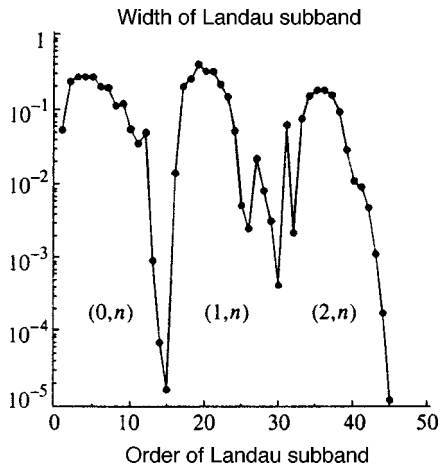


FIG. 2. Width of the Landau subbands (N, n) for $N=0$, 1, 2 ($n=1$, p) for $p/q=15/1$, $a=80$ nm; $V_0=-20$ meV.

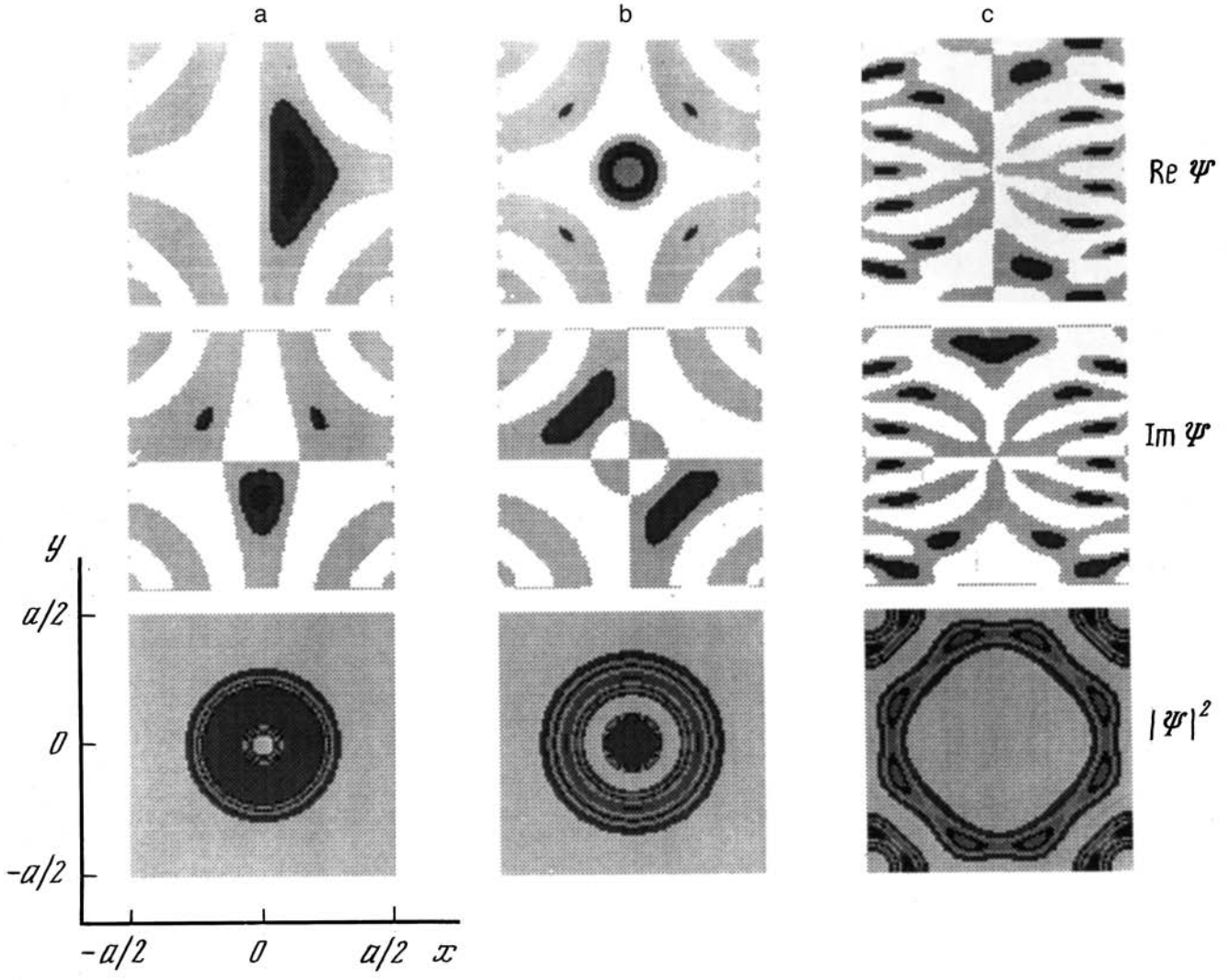


FIG. 3. Spatial distribution of the real and imaginary parts of the ψ function and the electron density $|\psi|^2$ for $\mathbf{k}=0$ in states: a) (1, 1); b) (1, 2); c) (1, 11). The magnetic field corresponds to the number of magnetic flux quanta $p/q=15/1$. Parameters of the superlattice: $a=80$ nm; $V_0=-20$ meV. The wave functions of states (1, 1) and (1, 11) are transformed under representation *B*, and those for (1, 2) are transformed under representation *A* of group C_2. The distribution of the electron density has C_{4V} group symmetry.

the edges of the region of the split Landau level are exponentially small, and they will be observed as discrete levels.

Figure 3 illustrates the spatial distributions of the real and imaginary parts of the electron wave functions and the electron density for states $(N_0, n_0)=(1,1); (1,2); (1,11)$, $\mathbf{k}=0$. The calculations are carried out for a magnetic field corresponding to $p/q=15/1$. The dark areas represent the regions of maximum values of the function. Only positive values of the real and imaginary parts of the ψ functions are plotted.

Our band structure determined here and the wave functions can be used to draw conclusions as to the nature of optical absorption in the investigated system. A group-theoretic analysis establishes the selection rules for transitions at an arbitrary point of the magnetic Brillouin zone. In particular, for $\mathbf{k}=0$ transitions between states transformed under different irreducible representations of the Hamiltonian symmetry group C_2, one symmetric (*A*) and one antisymmetric (*B*), are allowed in the dipole approximation. If the initial state *i* is transformed under representation *B* (*A*), a transition is allowed to all final states $f=i+[2j+1]$ (*j* is an integer) transformed under representation *B* (*A*). This rule is confirmed by numerical calculations of the matrix elements. All transitions are allowed at low-symmetry points of the Brillouin zone.

The transition probabilities 0 from different subbands split off from the same Landau level can differ substantially. For example, the probability is high for transitions from states in the middle of subbands ($\mathbf{k}=0$) that are split off from a Landau level with even (odd) *N* and are transformed under representation *A* (*B*). But if the initial state for $\mathbf{k}=0$ belongs to a subband split off from an even (odd) Landau level and is transformed under representation *B* (*A*), the transition probability becomes relatively low. This result stems from the fact that the wave function in the subband (N_0, n_0) is constructed mainly from N_0th-level oscillatory functions.

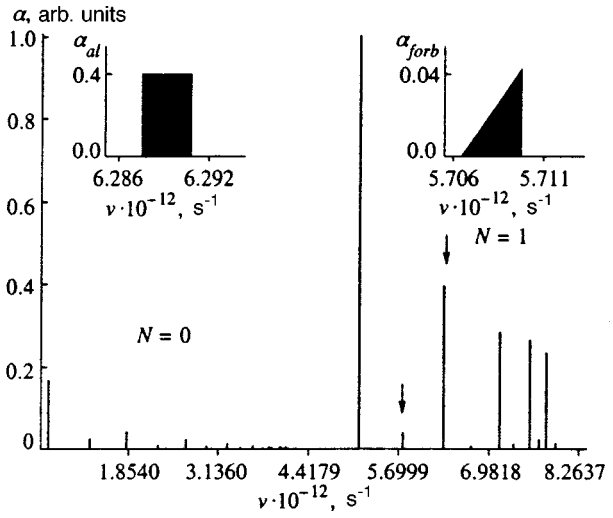


FIG. 4. Photon absorption probability for a superlattice of quantum dots with parameters $a=80$ nm and $V_0=-20$ meV in a magnetic field corresponding to $p/q=15/1$ vs. electromagnetic wave frequency. The lowest Landau subband $(0,\ 1)$ is partially filled. Transitions to states of subbands of the zeroth $(N=0)$ and first $(N=1)$ Landau levels are visible. Insets: fine structure of the absorption lines of allowed $\alpha_{\rm al}(\nu)$ and forbidden $\alpha_{\rm forb}(\nu)$ transitions, indicated by arrows in the main figure.

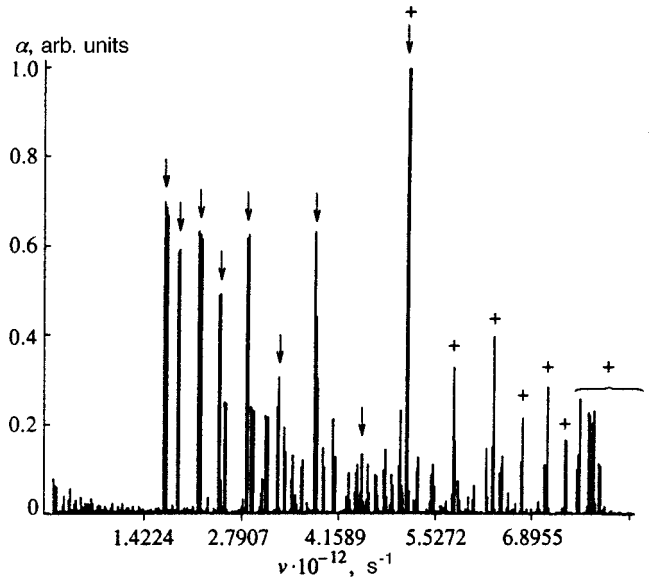


FIG. 5. Probability of absorption of a photon by a superlattice of quantum dots with parameters $a=80$ nm and $V_0=-20$ meV in a magnetic field corresponding to $p/q=15/1$ vs. electromagnetic wave frequency. The eight lowest Landau subbands $(0,\ 1)\ldots(0,\ 8)$ are completely filled, and the ninth subband $(0,\ 9)$ is partially filled. The arrows indicate transitions from the nine filled subbands to subband $(1,\ 1)$. The plus signs indicate transitions between subbands $(0,\ 1)\rightarrow(1,\ n)$ $(n=1,\ p)$.

It can be shown that the mixing of neighboring unperturbed $N_0\pm1$ Landau states in the subband wave functions of the N_0th split level is proportional to the small parameter

$$\beta=\frac{V_0}{\hbar\omega_c}\,\frac{2\pi l_H}{a},$$

which is defined as the ratio of matrix elements of the off-diagonal block of the matrix of the Hamiltonian $\hat{H}_0$ to elements of the diagonal block. (The structure of the matrix of the Hamiltonian is described in detail in Refs. 11 and 12). If the parameter $\beta\approx1$, the mixing of neighboring Landau states is high, and the probabilities of transitions from any initial state will be of the same order. This same parameter β is responsible for the weakness of transitions between states within a single split Landau level relative to transitions to neighboring split sublevels.

The probability of absorption by magnetic subbands is calculated according to Eq. (5). In a magnetic field corresponding to the number of magnetic flux quanta $p/q=15/1$ we have $a=9.71l_H$ for the indicated parameters of the superlattice, so that Eq. (7) can be used to calculate the transition matrix element.

To visualize the structure of the absorption spectrum in an instructive way, we calculate $\alpha(\nu)$ when only states of the lowest Landau subband $(0,1)$ near $\mathbf{k}=0$ are occupied. The corresponding absorption spectrum is shown in Fig. 4. The figure shows the absorption lines associated with transitions to subband states $(0,n)$ (low-frequency region) and $(1,n)$ (high-frequency region). Transitions to subband states $(1,n)$ are strongest; the transition probability to states $(0,n)$ is on the order of $\beta^2=0.5$ times lower. The probability of a photon being absorbed near $\mathbf{k}=0$ varies significantly as a function of the order number of the magnetic subband of the final state; alternating lines of allowed and forbidden transitions are visible. This pattern corresponds to our established selection rules. The insets to Fig. 4 show the fine structure of the absorption lines of allowed and forbidden transitions near the middle of the magnetic subbands indicated by the arrows in the main figure. The allowed transitions $(P_{\mathbf{kk}}^{i\rightarrow f}\neq0)$ near $\mathbf{k}=0$ correspond to characteristic plateaus of the absorption lines (see the left inset), whereas in the case of forbidden transitions $(P_{\mathbf{kk}}^{i\rightarrow f}=0)$ the absorption coefficient increases linearly with frequency (right inset). This result is attributable to the fact that the matrix element of a forbidden transition is proportional to $|\mathbf{k}|$. The width of the absorption lines is $\Delta^{i\rightarrow f}=|\Delta E_i-\Delta E_f|/\hbar$, where $\Delta E_{i,f}=|E_{i,f}(0)-E_{i,f}(k_F)|$ (k_F is the Fermi quasimomentum).

For typical densities of 2D electrons of the order of $10^{11}\,\mathrm{cm}^{-2}$, several subbands of the zeroth Landau level are filled (for $H\approx10^5$ Oe), adding to the complexity of the absorption spectrum. Figure 5 shows the photon absorption probability in transitions between Landau subbands when the eight lowest magnetic subbands $(0,1)\ldots(0,8)$ are completely filled and the ninth subband $(0,9)$ is partially filled. This situation corresponds to a carrier density of the order of $1.5\times10^{11}\,\mathrm{cm}^{-2}$. In Fig. 5 absorption lines associated with transitions between magnetic subbands split off from the Landau level $N=0$ are visible at lower frequencies. The first low-frequency absorption line shown in the figure is attributable to transitions between subband $(0,9)$ and subband $(0,10)$. The frequencies of this and the next-nearest transitions are in the millimeter microwave range. Such transitions between magnetic subbands of one Landau level can be detected by observing ordinary cyclotron resonance. The concept of cyclotron resonance is customarily identified with electron transitions between Landau levels. As shown above,

the presence of the periodic potential causes the Landau levels to split into magnetic subbands. It is significant here that the electron ψ functions in the subbands are mixed with all unperturbed Landau states [see Eq. (3)]. The contribution of different Landau levels to the states of subbands of the N_0th unperturbed level in the linear approximation with respect to the number of Landau levels included in the expansion (3) is proportional to the parameter β. In general, therefore, we encounter a nonzero transition probability between subbands contiguous with the same Landau level. This result follows from an analysis of Eq. (7) for the transition matrix element. For example, if several subbands of the zeroth level are filled, transitions to all other subbands of the given Landau level are characterized mainly by the second term in the bracketed expression of Eq. (7). It provides the largest contribution (of the order of β) to the transition matrix element. We can therefore speak of resonances (cyclotron resonances) associated with transitions between subbands rather than between Landau levels. The experimental investigation of such absorption spectra in the magnetic subbands affords a useful tool for studying the quantum states of Bloch electrons in a magnetic field.

Figure 5 also shows transitions from filled subbands of the zeroth Landau level to magnetic subbands contiguous with the first Landau level. Here we have omitted the enlarged-scale profiles of the absorption lines. We merely note that the line describing transitions from completely filled bands to an empty band has a logarithmic singularity in the middle (pagoda profile), owing to van Hove singularities in the density of states. The maximum absorption intensity corresponds to the transition between subbands $(0,1) \rightarrow (1,1)$.

The authors of an innovative study[8] of luminescence in a quantum antidot lattice ($a=200\,\mathrm{nm}$) in a magnetic field ($H \leqslant 2\times 10^4\,\mathrm{Oe}$) have observed oscillations of the luminescence intensity in connection with the existence of commensurate orbits with cyclotron radii $2R_c=(S-1/4)a$ (S is an integer). It is readily verified that the commensurability condition can be satisfied when the number of occupied Landau levels is much greater than unity. Moreover, special calculations of our own have shown that the magnetic subbands of individual Landau levels overlap considerably under the experimental conditions of Ref. 8. We note that the commensurability of orbits with a periodic potential cannot be observed for the parameters used in our work, because only one Landau level is partially occupied, and the situation is far removed from semiclassical status. In addition, the magnetic subbands of different Landau levels do not overlap.

In summary, we have established that the investigation of cyclotron resonance and optical absorption in the infrared can yield very valuable information about the structure of Bloch states in a 2D lattice in a magnetic field.

This work was carried out with financial support from the Russian Fund for Fundamental Research (Project 98-02-16412).

*)E-mail: demi@phys.unn.runnet.ru
†)E-mail: perov@phys.unn.runnet.ru

[1] M. Ya. Azbel', Zh. Éksp. Teor. Fiz. **46**, 929 (1964) [Sov. Phys. JETP **19**, 634 (1964)].
[2] J. Zak, Phys. Rev. A **134**, 1602 (1964); Phys. Rev. A **134**, 1607 (1964).
[3] A. Rauh, Phys. Status Solidi B **65**, K131 (1974).
[4] D. R. Hofstadter, Phys. Rev. B **14**, 2239 (1976).
[5] H. Silberbauer, J. Phys.: Condens. Matter **4**, 7355 (1992).
[6] D. Weiss, M. L. Roukes, A. Menschig *et al.*, Phys. Rev. Lett. **66**, 27 (1991).
[7] D. Weiss, K. Richter, A. Menschig *et al.*, Phys. Rev. Lett. **70**, 4118 (1993).
[8] I. V. Kukushkin, D. Weiss, G. Lütjering *et al.*, Phys. Rev. Lett. **79**, 1722 (1997).
[9] D. Weiss, D. Grambow, K. von Klitzing *et al.*, Appl. Phys. Lett. **58**, 2960 (1991).
[10] T. Schlösser, K. Ensslin, J. P. Kotthaus *et al.*, Semicond. Sci. Technol. **11**, 1582 (1996).
[11] V. Ya. Demikhovskiĭ and A. A. Perov, Fiz. Tverd. Tela (St. Petersburg) **40**, 1134 (1998) [Phys. Solid State **40**, 1035 (1998)].
[12] A. A. Perov, Preprint No. SMR.998d-15 (International Centre for Theoretical Physics, Trieste, 1997).
[13] V. Ya. Demikhovskiĭ and A. A. Perov, in *Megagauss and Megampere Pulse Technology and Applications* [in Russian], Sarov (1997).

Translated by James S. Wood

Noncollinear orientation of external magnetic field and flux lines penetrating an isotropic hard superconductor

S. E. Savel'ev and L. M. Fisher

State Research Center ''All-Russian Electrotechnical Institute,'' 111250 Moscow, Russia

V. A. Yampol'skiĭ

Institute for Radiophysics and Electronics, National Academy of Sciences of Ukraine, 310085 Khar'kov, Ukraine

(Submitted 31 March 1998)
Zh. Éksp. Teor. Fiz. **114**, 1804–1816 (November 1998)

Penetration by Abrikosov flux lines of an isotropic hard superconductor in the critical state induced by changes in the orientation of external magnetic field has been theoretically investigated. The analysis has been based on the microscopic nonlocal model taking into account forces of bulk and surface pinning, alongside magnetic forces of interaction of the row of penetrating vortices with existing flux lines, Meissner currents, and vortex images. New vortices penetrate a superconductor only when the angle through which the field is rotated is larger than a certain critical value. It has been determined that the alignment of entering vortices is essentially different from that of the applied magnetic field. The feasibility of detecting noncollinearity effects is discussed. © *1998 American Institute of Physics.* [S1063-7761(98)01811-3]

1. INTRODUCTION

The problem of Abrikosov flux lines penetrating bulk superconductors has a long history. This issue was first discussed in the well-known paper by Bean and Livingston,[1] who predicted on the basis of the thermodynamic approach the existence of a surface barrier which should impede penetration of flux lines into a sample. Effects related to this barrier in soft superconductors were studied by Ternovskiĭ and Shekhata.[2] Later, surface barriers of various types, their properties and manifestations in all types of superconductors were investigated by many researchers (see the review by Brandt[3] and references therein).

Although the number of publications on this topic has been quite considerable, some aspects of vortex penetration have not attracted researchers' attention. In particular, they usually considered a situation when only the magnitude of an external magnetic field varied, and all flux lines entering and leaving a sample were aligned with the external magnetic field **H**. At the same time, an electrodynamic description of hard superconductors in a magnetic field of variable amplitude and orientation has remained a topical issue for many years. For example, after the first studies[4–9] of the effect of crossing between vortices with differing orientations, a set of publications concerning this phenomenon were issued by different groups recently.[10–16] A change of the magnetic field orientation leads to several interesting effects, among which we direct the reader's attention to the collapse of a transport current caused by an alternating magnetic field parallel to this current[17,18] and suppression of the static magnetic moment of a sample by a transverse alternating magnetic field.[11,19]

It is clear that a proper description of these and similar phenomena demands information concerning alignment of vortices penetrating a sample when the external magnetic field **H** is rotated. It might seem that vortices penetrating an isotropic hard superconductor should always be aligned with **H**. This conclusion is often derived from the continuity of the tangential component of the magnetic induction at the sample surface. The continuity condition, however, does not yield the direction of vortices penetrating the sample. In fact, the magnetic induction in a superconductor is a sum of two components, one of which, $\mathbf{B}_v$, is due to the system of flux lines inside the superconductor, while the other, $\mathbf{B}_m$, is related to the Meissner current flowing near the sample surface:

$$\mathbf{B}=\mathbf{B}_v+\mathbf{B}_m. \tag{1}$$

It is well known[20] that the solenoidal component of magnetic induction $\mathbf{B}_v$ vanishes at the superconductor boundary, so the necessary boundary condition contains only the Meissner component $\mathbf{B}_m$. Hence, the orientation of vortices near the surface is undefined. Their orientation should be determined using other considerations. These might be the balance of forces on a penetrating vortex or a minimum of the Gibbs free energy. One can infer from energy considerations that magnetic flux lines both entering and leaving a superconductor are aligned with the external magnetic field only when the applied magnetic field direction has remained unchanged throughout the prior magnetic history of a sample. When the magnetic field orientation varies, the question about the orientation of vortices entering a superconductor is not so simple. This paper is dedicated to the theoretical analysis of this problem.

We have analyzed the simplest case, when an external magnetic field is rotated in the plane parallel to the boundary

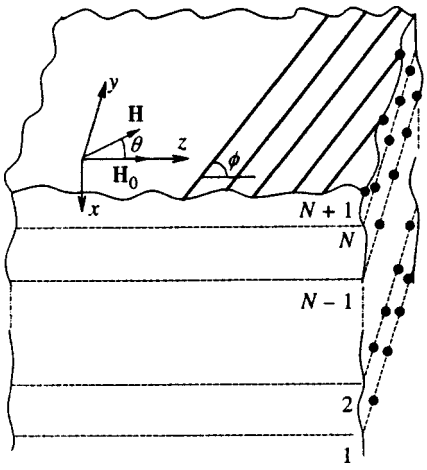


FIG. 1. Problem geometry.

of a superconducting half-space. In our analysis, we have used a microscopic model of the critical state.[21–23] In this model, a system of flux lines is described in terms of coordinates of isolated flux line rows. It is assumed that, irrespective of its position, a pinning force $f_{\rm pin}$ acts on each vortex. This force ranges between $-p_{\rm pin}$ and $p_{\rm pin}$, where $p_{\rm pin}$ is a constant phenomenological parameter. In the problem under discussion, an important role is played by nonlocal effects[24,25] due to the nonlocal relation between the magnetic induction and flux line density. The importance of nonlocality in problems concerning a magnetic field changing its orientation was first indicated by D'Anna.[15,16] His analysis, however, was based on a macroscopic approach to penetration of vortices into a superconductor and neglected some forces of a microscopic nature that act on rows of penetrating vortices.

The main result of the reported work is that the orientations of vector **H** and flux lines penetrating an isotropic bulk superconductor can be notably different. The onset of vortex penetration into a hard superconductor takes place when the angle of the external field exceeds a certain threshold angle $\theta_{\rm in} \ll 1$. The angle between the row of penetrating flux lines and external magnetic field turns out considerably larger than $\theta_{\rm in}$. Moreover, at certain parameter values, this angle can be close to $\pi/2$. We have analyzed the dependence of $\theta_{\rm in}$ and relative orientation of penetrating vortices on the magnitude of external magnetic field and other parameters of the problem. The threshold angle diminishes with H as $1/H$, whereas the tilt angle of penetrating vortices tends to a constant.

2. STATEMENT OF THE PROBLEM. MICROSCOPIC EQUATIONS

Consider penetration of the Abrikosov flux lines into a hard superconductor which occupies half-space $x>0$. The magnetic field is assumed first to be aligned with the z-axis and increased monotonically from zero to $H_0 \gg H_{c1}$, then rotated in the zy-plane through angle θ (its vector is denoted by **H**). Here H_{c1} is the low critical field of the superconductor. The analyzed configuration is depicted by Fig. 1. This figure also shows schematically rows of magnetic flux lines that penetrated the sample prior to the rotation of applied magnetic field. The numbers of these rows run from unit to N. As was stated above, all vortices in these rows are aligned with the z-axis.

Let us focus on the possible reaction of the system to the magnetic field rotation. In principle, the system of vortices would lower its energy by turning through the same angle θ with the magnetic field. This process takes place in soft superconductors. The situation is radically different in hard superconductors, since pinning forces prevent changes in the flux line alignment. The magnetic torque on each vortex is actually proportional to its length L, whereas the torque due to pinning forces is proportional to its length squared. Comparison between these two torques implies that, when

$$L \gg cH_0/2\pi J_c, \tag{2}$$

i.e., the sample is sufficiently long and has a sufficient critical current density J_c, realignment of the vortex system is impossible. In principle, vortices can be bent in a certain region near the sample boundary, where the y-component of magnetic field penetrates the sample, i.e., the vortex vector has a nonzero projection on the y-axis. A simple estimate[26] yields a dimension of this region of order of $cH_0/2\pi J_c$ and, in accordance with Eq. (2), this is a small fraction of L. In what follows, we assume that this condition, which is considered a definition of bulk hard superconductors, is satisfied.

2.1. Gibbs free energy

Thus, the only possibility of changing the orientation of vortices in a bulk hard superconductor is associated with penetration of new magnetic flux lines tilted with respect to the z-axis and crossing between vortices with different orientations. As a result of this crossing, long flux lines are cut into shorter segments, whose alignment can be changed.[10,27] To determine the conditions for penetration of these new vortices when the applied field is rotated, and to calculate the tilt angle ϕ of these vortices with respect to the z-axis, one must use force or energy approaches. In this procedure, one must take account of the magnetic interaction between a penetrating vortex, and its image,[1] the existing flux line lattice, Meissner currents, and bulk and surface pinning centers. This problem has been solved on the basis of a microscopic approach,[22–25] since it becomes necessary to investigate the entry of successive vortices into a sample. This approach takes accurate account of nonlocality effects due to the long-range interaction between vortices, whose range is of order of the London penetration depth λ.

In order to calculate the magnetic forces on each row of vortices, we have determined the electromagnetic component $G_{\rm em}$ of the Gibbs free energy. The cumbersome calculations needed for solving the problem are described in Appendix. Here the final result is given:

$$G_{\rm em}(x_1, x_2, \ldots, x_N, x_{N+1})$$

$$= \cos\phi \sinh\frac{x_{N+1}}{\lambda} \frac{\Phi_0^2}{4\pi b\lambda} \sum_{j=1}^{N} \exp\left(-\frac{x_j}{\lambda}\right) + \cos(\phi-\theta)$$

$$\times \frac{H_0\Phi_0}{4\pi}\left[\exp\left(-\frac{x_{N+1}}{\lambda}\right)-1\right] + G_s(x_{N+1}) + \frac{\Phi_0^2}{16\pi b\lambda}$$

$$\times\left[1-\exp\left(-\frac{2x_{N+1}}{\lambda}\right)\right]+\frac{\Phi_0^2}{16\pi b\lambda}$$
$$\times\sum_{i,j=1}^{N}\left[\exp\left(-\frac{|x_i-x_j|}{\lambda}\right)-\exp\left(-\frac{x_i+x_j}{\lambda}\right)\right]$$
$$+\frac{H_0\Phi_0}{4\pi}\cos\theta\sum_{i=1}^{N}\left[\exp\left(-\frac{x_i}{\lambda}\right)-1\right], \qquad (3)$$

where Φ_0 is the magnetic flux quantum, b is the distance between vortices in a row, $x_1, x_2, \ldots, x_N$ are coordinates of vortex rows that penetrated the sample before the external field was rotated, and x_{N+1} is the coordinate of the vortex row whose tilt angle ϕ with respect to the z-axis is to be calculated. In Fig. 1 the penetrating vortex row is labeled by the number $N+1$. The first four terms on the right-hand side of Eq. (3) describe the interaction of the penetrating row with all other vortices, the Meissner current, and their images. The third term $G_s(x_{N+1})$ is associated with the short-range component of the energy of interaction between a vortex and its image, which is responsible for the Bean–Livingston barrier, and the fourth term describes the long-range component of this energy. The rest of the terms describe interactions among vortex rows labeled by numbers from 1 to N, with their images, and also with the Meissner current.

2.2. Force balance equation for vortices in the bulk of the sample

Given Eq. (3), one can easily derive a force balance equation for each vortex row. To this end, one must equate the pinning force to the negative of the derivative of $G_{\rm em}$ with respect to the row coordinate. For vortices numbered 1 to N we have

$$f_{\rm pin}^{(n)}=-\frac{\partial G_{\rm em}}{\partial x_n}=\frac{\Phi_0^2}{8\pi b\lambda^2}\left[\sum_{j=n+1}^{N}\exp\left(\frac{x_j-x_n}{\lambda}\right)\right.$$
$$\left.-\sum_{j=1}^{n-1}\exp\left(\frac{x_n-x_j}{\lambda}\right)-\sum_{j=1}^{N}\exp\left(-\frac{x_j+x_n}{\lambda}\right)\right]$$
$$+\frac{H_0\Phi_0}{4\pi\lambda}\cos\theta\exp\left(-\frac{x_n}{\lambda}\right), \quad 1\leqslant n\leqslant N, \qquad (4)$$

where the pinning force $f_{\rm pin}^{(n)}$ can take any value between $-p_{\rm pin}$ and $p_{\rm pin}$. The quantity $p_{\rm pin}$ is a phenomenological parameter of the theory. This parameter will be assumed position-independent throughout the volume $x>0$.

2.3. Force balance equation for the penetrating vortex row

The force balance equation for the penetrating vortex row has the form

$$p_{\rm sh}^{\rm sur}=-\frac{\Phi_0^2}{8\pi b\lambda^2}\left[2\cos\phi\sum_{j=1}^{N}\exp\left(-\frac{x_j}{\lambda}\right)+1\right]$$
$$+\frac{H_0\Phi_0}{4\pi\lambda}\cos(\theta-\phi), \qquad (5)$$

where $p_{\rm sh}^{\rm sur}$ is the sum of the surface pinning force, which is no less than the bulk pinning force $p_{\rm pin}$, and the short-range component of the force between the penetrating vortex row and its image obtained by differentiating the term $G_s(x_{N+1})$ in the expression for the Gibbs energy (3). The force $p_{\rm sh}^{\rm sur}$ is strongly affected by the sample surface quality, and it is natural to treat this quantity as a second phenomenological parameter of the theory. In the limit of a perfect surface, this parameter, to order of magnitude, can be as great as $p_{\rm sh}^{\rm sur}\sim\Phi_0^2/16\pi^2\xi\lambda^2$, where ξ is the coherence length. This surface force dictates the peak height of the Bean–Livingston barrier. In real superconductors, this force is considerably reduced by surface roughness, thermal activation effects, etc.[3]

In addition to the short-range force $p_{\rm sh}^{\rm sur}$, Eq. (5) also contains the long-range force obtained by differentiating the fourth term on the right-hand side of Eq. (3) with respect to x_{N+1}. This force corresponds to the 1 in the brackets on the right-hand side of Eq. (5). Both surface forces are constant in θ and ϕ, so they can be conveniently combined into the single term $p^{\rm sur}=p_{\rm sh}^{\rm sur}+\Phi_0^2/8\pi b\lambda^2$. Using this notation, we can rewrite Eq. (5), which expresses the condition for penetration of a new vortex row into the sample, in the form

$$p^{\rm sur}=F_{\rm mag}(\phi,\theta)=-\frac{\Phi_0^2}{4\pi b\lambda^2}\cos\phi$$
$$\times\sum_{j=1}^{N}\exp\left(-\frac{x_j}{\lambda}\right)+\frac{H_0\Phi_0}{4\pi\lambda}\cos(\theta-\phi). \qquad (6)$$

An analysis of Eqs. (4) and (6), which describe the balance of forces, is quite complicated. We therefore first take a microscopic approach to vortex penetration of a sample in a fixed-orientation increasing magnetic field, and only then return to the problem of a rotated magnetic field.

3. VORTEX PENETRATION OF A HARD SUPERCONDUCTOR IN AN INCREASING MAGNETIC FIELD

We use Eqs. (4) and (6) to analyze vortex-row penetration of a hard superconductor in an increasing external magnetic field. It is convenient to use the following dimensionless variables:

$$p=\frac{8\pi b\lambda^2}{\Phi_0^2}p_{\rm pin}, \quad p_s=\frac{8\pi b\lambda^2}{\Phi_0^2}p^{\rm sur}>1+p,$$

$$f_{\rm mag}(\phi,\theta)=\frac{8\pi b\lambda^2}{\Phi_0^2}F_{\rm mag}(\phi,\theta),$$

$$h=\frac{2b\lambda H_0}{\Phi_0}, \quad \zeta_n=\frac{x_n}{\lambda}. \qquad (7)$$

Setting $\phi=0$ and $\theta=0$ in Eqs. (4) and (6), we rewrite the equations in the new variables

$$p=\sum_{j=n+1}^{N}\exp(\zeta_j-\zeta_n)-\sum_{j=1}^{n-1}\exp(\zeta_n-\zeta_j)$$
$$-\sum_{j=1}^{N}\exp(-\zeta_j-\zeta_n)+h\exp(-\zeta_n),$$
$$p_s=f_{\rm mag}=-2\sum_{j=1}^{N}\exp(-\zeta_j)+h. \quad (8)$$

This system of $N+1$ equations determines the x coordinates of the N vortex rows in the sample, and the field $h=h^{(N+1)}$ at which the last $[(N+1)\text{th}]$ row penetrates the superconductor. At $N=0$ these equations yield the field $h^{(1)}$ at which the first vortex row enters the sample.

The second equation of the system immediately implies $h^{(1)}=p_s$, or in dimensional units,

$$H_0^{(1)}=p_s^2H_{c1}>H_{c1}.$$

It is just as easy to find the field $h^{(2)}$:

$$h^{(2)}=\sqrt{p_s^2+4p}.$$

In the general case, when the sample contains N vortex rows, a complicated nonlinear problem must be analyzed. Surprisingly, this problem admits of an exact analytic solution.[21] The field at which the $(N+1)$th row penetrates the sample was calculated in Ref. 23:

$$h^{(N+1)}=\sqrt{p_s^2+4Np}. \quad (9)$$

In the same paper, the magnetic force $f_v(h)$ on an entering vortex row generated by vortices already present was also calculated (first term on the right-hand side of the second of Eqs. (8)):

$$f_v(h)=-2\sum_{j=1}^{N}\exp(-\zeta_j)=\sqrt{h^2-4pN}-h. \quad (10)$$

Here the number N of vortex rows already present is

$$N=\left[\frac{h^2-p_s^2}{4p}\right]+1, \quad (11)$$

where the notation denotes the integer part of its argument.

It follows from the above discussion that vortices penetrate the sample at certain discrete values of the external magnetic field. The next vortex row in turn, number N, enters the sample at field values $h=h^{(N)}$ such that the magnetic force $f_{\rm mag}$ on that row (Eq. (8)) exceeds the force p_s due to interaction with the surface. The magnetic force is determined by the current J_v generated by vortices in the sample (first term on the right-hand-side of Eq. (8)) and the Meissner current J_m (second term). Here J_v and J_m are the currents generating the solenoidal and Meissner components of the magnetic induction in Eq. (1). Immediately after penetration of the Nth vortex row, the force balance condition (8) no longer holds, since the force generated by each newly arrived row adds to the magnetic force and ''impedes'' entry of the next $[(N+1)\text{th}]$ row. Circumstances governing the entry of new vortices gradually improve as h continues to rise, since the force due to the Meissner current increases. When the field reaches $h=h^{(N+1)}$, the magnetic force reaches a value such that force balance again holds, and the $(N+1)$th row enters the sample. Figure 2 shows a segment of the curve of magnetic force $f_{\rm mag}$ plotted against dimensionless field h calculated by Eqs. (8), (10), and (11) at fixed p and p_s. The jumps in the curve correspond to penetration of the next vortex row when $f_{\rm mag}$ reaches p_s.

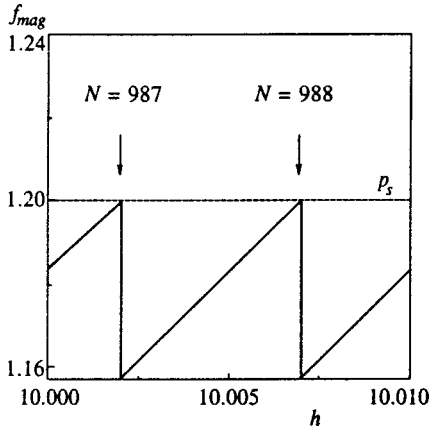


FIG. 2. Magnetic force on a penetrating vortex row as a function of dimensionless magnetic field h calculated at $p=0.025$, $p_s=1,2$. The arrows point to fields h at which the next vortex row penetrates the sample.

The distance between jumps in the magnetic force along the horizontal axis is given by Eq. (9). In dimensional variables, the magnetic field interval ΔH between neighboring jumps can be expressed as

$$\Delta H\simeq 2H_b\sqrt{H_{c1}/H_0}. \quad (12)$$

Here $J_c=cp_{\rm pin}/\Phi_0$ is the critical current density, and $H_b=4\pi J_c\lambda/c$ is the amplitude of the so-called nonlocal barrier.[24,25] In the usual local theory of the critical state of hard superconductors, where the London penetration depth is the smallest characteristic length, this small barrier is ignored, whereas in the nonlocal theory this barrier determines the range of external magnetic fields over which the vortex lattice in a superconductor remains immobilized. In Eq. (12) the field H_b is multiplied by the small parameter of the theory $\sqrt{H_{c1}/H_0}$. The discrete nature of vortex penetration of a hard superconductor can therefore scarcely be detected experimentally, even in perfectly prepared samples. This means that the macroscopic local theory neglecting the characteristic length λ and magnetic field increment ΔH is quite adequate for describing vortex penetration of a superconductor in an increasing magnetic field. At the same time, vortex penetration of a sample in a rotating magnetic field can be described only in terms of the microscopic theory. The point, as will be shown in the next section, is that the orientation of flux lines in a rotated field **H** is extremely sensitive to the location of H_0 within the interval ΔH defined by Eq. (12).

4. VORTEX PENETRATION OF A HARD SUPERCONDUCTOR UNDER ROTATION OF AN EXTERNAL MAGNETIC FIELD

Using the microscopic approach described above, we examine the penetration of vortices into a superconductor when

the external magnetic field can be rotated. The analysis will be based on the general equation system (4), (6). A rotation of field **H** can, generally speaking, induce in the vortex system changes of two types. First, a decrease in the magnetic force due to the factor $\cos\theta$ in the last term of Eq. (4) can destroy the balance of forces and eject vortices from the superconductor. Second, penetration of vortices with a different orientation can be induced.

We now determine which of these two processes begins first, i.e., at smaller θ. Supposing that the ejection of vortices from the sample begins first, one can calculate from Eq. (4) the threshold angle θ_{out} at which the Nth vortex row, which is most sensitive to changes in the external field, begins to depart. Recall that the pinning force $f_{\text{pin}}^{(n)}$ in Eq. (4) can take an arbitrary value in the range $-p_{\text{pin}}<f_{\text{pin}}^{(n)}<p_{\text{pin}}$, cancelling changes in the magnetic force on the right-hand side of the same equation. Therefore, the balance of forces is destroyed only when the magnetic force decreases by $2p_{\text{pin}}$. As a result, we obtain an equation for the threshold angle θ_{out} in dimensionless variables defined by Eq. (7):

$$1-\cos(\theta_{\text{out}})=\frac{2p\exp(x_N/\lambda)}{h}. \tag{13}$$

Hence, $\theta_{\text{out}}>2(p/h)^{1/2}$. This means that all flux lines in the bulk of the sample remain immobilized at rotation angles $\theta<2(p/h)^{1/2}$. As will be shown below, penetration of new vortices begins at angles θ_{in} that are significantly less than θ_{out}. Therefore the penetration of new vortices will be analyzed assuming that all vortices already in place are immobile.

Let us rewrite the force balance equation (6) for vortices penetrating the sample in dimensionless variables with due account of Eq. (10):

$$p_s=f_v\cos\phi+h\cos(\phi-\theta_{\text{in}}),\quad f_v=\sqrt{h^2-4pN}-h<0. \tag{14}$$

This equation contains two unknown angles ϕ and θ_{in}. An additional equation is needed to determine these angles. It can be easily obtained by noting that the first vortex row to enter the sample must be oriented at an angle ϕ that maximizes the driving magnetic force. By requiring that the derivative of the second of Eqs. (14) with respect to ϕ vanish, we obtain the required second equation:

$$f_v\sin(\phi)+h\sin(\phi-\theta_{\text{in}})=0. \tag{15}$$

This equation alone shows that the orientation ϕ of the penetrating vortices is different from that of the external field: $\phi>\theta_{\text{in}}$.

Equations (14) and (15) can be rewritten in terms of ϕ and $\phi-\theta_{\text{in}}$. Some simple algebra yields

$$\cos\phi=\frac{h^2-f_v^2-p_s^2}{2(-f_v)p_s},\quad \cos(\phi-\theta_{\text{in}})=\frac{h^2-f_v^2+p_s^2}{2hp_s}. \tag{16}$$

In the general case [arbitrary h in the interval $(h^{(N)},h^{(N+1)})$], the expressions for ϕ and θ_{in} are quite messy. For this reason, Fig. 3 shows numerical calculations of these angles as functions of the dimensionless field h. Clearly, the threshold rotation angle of the field **H** and the

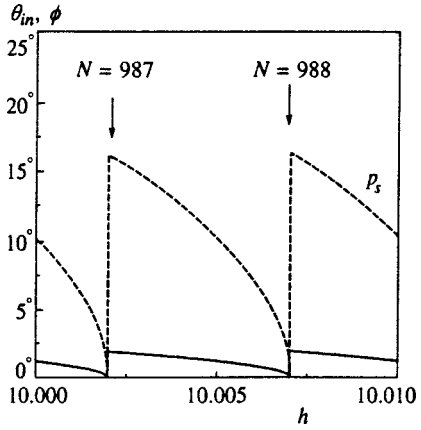


FIG. 3. Threshold angle θ_{in} (solid lines) and tilt angle ϕ of penetrating vortices (dashed lines) vs. dimensionless magnetic field h calculated at $p=0.025$ and $p_s=1, 2$.

alignment of penetrating vortices are extremely sensitive to the exact value within the interval $(h^{(N)},h^{(N+1)})$ at which the external magnetic field stopped increasing prior to reorientation. Characteristic values of the angle ϕ are substantially greater than the rotation angle θ_{in}. The angles ϕ and θ_{in} have their largest values $\phi_{\max}$ and $\theta_{\max}$ at the leftmost limits of the intervals $h^{(N)}<h<h^{(N+1)}$, and then monotonically drop to zero (this can be easily found by analyzing formulas (16)). The peak amplitudes $\phi_{\max}$ and $\theta_{\max}$ decrease with increasing h (as the number of vortex rows introduced into the sample before rotating the field increases). Curves of $\phi_{\max}(h)$ and $\theta_{\max}(h)$ are plotted in Fig. 4. It is clear that $\theta_{\max}(h)$ monotonically drops to zero, whereas $\phi_{\max}(h)$ tends to a certain constant. As a result, the ratio $\phi_{\max}(h)/\theta_{\max}(h)$ increases without bound.

This behavior of $\phi_{\max}(h)$ and $\theta_{\max}(h)$ at all admissible values of the phenomenological parameters p and p_s follows directly from Eqs. (9)–(11) and (16). It can be shown that when the field h is slightly greater than $h^{(N)}$, f_v equals $(p_s^2-4p)^{(1/2)}-h$. Thus, we derive from Eq. (16)

$$\cos\phi_{\max}=\frac{h\sqrt{p_s^2-4p}-p_s^2+2p}{p_s(h-\sqrt{p_s^2-4p})}, \tag{17}$$

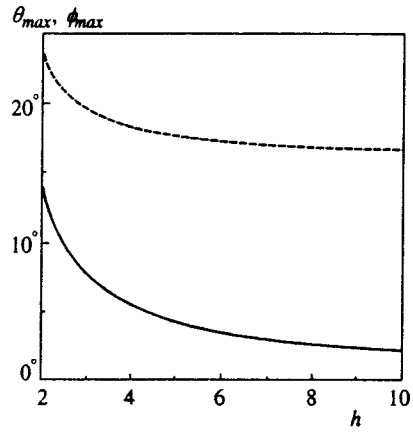


FIG. 4. Angles $\theta_{\max}$ and $\phi_{\max}$ (dashed line) vs. dimensionless magnetic field h calculated at $p=0.025$ and $p_s=1,2$.

$$\cos(\phi_{\max}-\theta_{\max})=\frac{h\sqrt{p_s^2-4p}+2p}{hp_s}. \tag{18}$$

Asymptotically, in the limit of strong magnetic fields $h \gg p_s$,

$$\phi_{\max}(h)\simeq\arcsin(2\sqrt{p}/p_s), \quad \theta_{\max}(h)\simeq 2\sqrt{p}/h, \quad h\gg p_s. \tag{19}$$

Note that $\theta_{\max}(h)$ is smaller by a factor of $(h)^{1/2}\gg 1$ than the threshold angle θ_{out} defined by Eq. (13), at which vortices are ejected by rotation of the external magnetic field.

5. DISCUSSION

The most important result of this work is that we have established an interesting feature of vortex penetration into a superconductor induced by rotation of an external magnetic field. The vortices generated in the sample are aligned noncollinearly with the magnetic field. The angle between them depends on external parameters and can be quite considerable. For example, this angle reaches $\pi/2$ in the limit of strong magnetic fields h at $p\sim p_s$. The continuity of the tangential magnetic field component is preserved here, since the field generated by vortices is always zero on the interface, while continuity is ensured by the field component due to the Meissner current.

The physical reason for this noncollinearity between the generated vortices and applied magnetic field is related to the interaction between the vortex system and superconductor surface. Therefore, it may seem surprising that the misalignment angle does not increase with the parameter p_s, which characterizes this interaction, but, on the contrary, decreases notably. The asymptotic behavior given by Eq. (19) indicates that in fact both angles $\phi_{\max}(h)$ and $\theta_{\max}(h)$ decrease with p_s and approach one another. This result can be naturally explained in terms of the nonlocal microscopic model of the critical state. At large $p_s\gg p$, the vortices that have penetrated the sample and overcome all surface barriers are acted upon by a uncompensated magnetic force due to the Meissner currents (the motion of these vortices is impeded by the small bulk pinning force p instead of p_s). As a result, the vortices ''drop'' deep into the bulk through a distance of about λ, and the so-called vortex-free region is created.[22,23] In the long run, the role of the term f_v, which is associated with the magnetic force on entering vortices due to already-present vortices, is not very important in the force balance equation (14), and consequently in Eq. (15). Therefore the problem of vortex penetration becomes similar to that of vortex penetration in a fixed-orientation field[1,2] in this situation, i.e., the entering vortices are essentially collinear with the applied field.

The results of the reported work are interesting not only from the standpoint of the physical model of vortex penetration, but can also be instrumental in developing a kinetic model of the vortex system of a superconductor in a magnetic field with variable orientation. Such models were described in our previous publications.[10–12] We suggested that vortices at each point could have a small angular jitter. It follows from the results of the present work, however, that vortices of an arbitrary orientation can penetrate a superconductor when the applied field is rotated through a relatively small angle. This means that the equations obtained previously[10–12] describe the vortex lattice only in those regions of the sample where the angular jitter is inessential owing to multiple crossings among vortices, i.e., far from the superconductor boundaries. The reported results can be considered a first step toward formulation of correct boundary conditions for the kinetic theory.[10–12]

Finally, let us discuss the feasibility of experimental detection of a strong misalignment of vortices penetrating a superconductor at small rotation angles of applied magnetic field. If a sample (parallelepiped) is placed in a magnetic field aligned with its long axis and then the field **H** is rotated in the plane of the largest face through a small angle $\theta > \theta_{\text{in}}$, one can attempt to observe, using the decoration technique, vortex cores ending on the side surface at large distances from the end faces of the parallelepiped. This observation would provide direct confirmation of the theory developed here.

This work was financially supported by the Russian Fund for Fundamental Research (Project 97-02-16399) and by the State Superconductivity Program (Project 96046).

APPENDIX A: GIBBS ENERGY OF CROSSING VORTEX ROWS

Consider the Gibbs free energy of vortices in the London approximation

$$G_{\text{em}}=\frac{1}{8\pi}\int dx\,dy[\mathbf{B}^2+\lambda^2(\operatorname{curl}\mathbf{B})^2-2\mathbf{H}\cdot\mathbf{B}].$$

The Meissner $\mathbf{B}_m$ and vortex $\mathbf{B}_v$ components of magnetic induction **B** satisfy the following equations and boundary conditions:

$$\mathbf{B}_m+\lambda^2\operatorname{curl}\operatorname{curl}\mathbf{B}_m=0, \quad \mathbf{B}_m(x=0)=\mathbf{H},$$

$$\mathbf{B}_v+\lambda^2\operatorname{curl}\operatorname{curl}\mathbf{B}_v=\Phi_0\sum_i \mathbf{e}_i\delta(x-x_i) \times\sum_{k=-\infty}^{\infty}\delta\left[y-\left(k-\frac{i}{2}\right)b\right]\equiv\Phi_0\mathbf{n},$$

$$\mathbf{B}_v(x=0)=0.$$

Here $\mathbf{e}_i$ is the unit vector aligned with the vortex, index i labels vortex rows, and k is the vortex number in a given row. The equations take into account the fact that vortices form a triangular lattice. After simple transformations and omission of terms independent of vortex row coordinates, we have

$$G_{\text{em}}=\frac{1}{8\pi}\int dx\,dy(\Phi_0\mathbf{B}_v\cdot\mathbf{n}-2\mathbf{H}\cdot\mathbf{B}_v). \tag{A1}$$

We write the vortex component of the magnetic field as a sum of fields generated by vortex rows and their images:

$$\mathbf{B}_v = \sum_i \mathbf{e}_i B_i , \tag{A2}$$

where

$$B_i \mathbf{e}_i + \lambda^2 \operatorname{curl}\operatorname{curl}(B_i \mathbf{e}_i) = \Phi_0 \mathbf{e}_i \delta(x - x_i) \sum_{k=-\infty}^{\infty} \delta\left[y - \left(k - \frac{i}{2} \right) b \right]. \tag{A3}$$

The field $B_i(x,x_i)$ generated by the ith vortex row is determined by a sum of modified Bessel functions of the second kind. It can be shown[28] that B_i has a component decaying rapidly over a distance smaller than b and a component dropping exponentially over the length $\lambda \gg b$:

$$\begin{aligned} B_i = {} & \frac{\Phi_0}{2b\lambda}\left[\exp\left(-\frac{|x-x_i|}{\lambda} \right) - \exp\left(-\frac{x+x_i}{\lambda} \right) \right] - \frac{\Phi_0}{4\pi\lambda^2} \\ & \times \ln\left[1 - 2\exp\left(-\frac{2\pi|x-x_i|}{b} \right) \cos\frac{2\pi(y-y_i)}{b} \right. \\ & \left. + \exp\left(-\frac{4\pi|x-x_i|}{b} \right) \right] + \frac{\Phi_0}{4\pi\lambda^2} \ln\left[1 - 2 \right. \\ & \times \exp\left(-\frac{2\pi(x+x_i)}{b} \right) \cos\frac{2\pi(y-y_i)}{b} \\ & \left. + \exp\left(-\frac{4\pi(x+x_i)}{b} \right) \right], \end{aligned} \tag{A4}$$

where $y_i = 0$ for even i and $y_i = b/2$ for odd i.

The rapidly decaying component can always be neglected, unless the penetrating row interacts with its image. This interaction generates an additional contribution G_s to the Gibbs free energy, which determines the surface barrier. Substituting Eqs. (A2)–(A4) into Eq. (A1), we obtain an expression for the Gibbs free energy:

$$\begin{aligned} & G_{\mathrm{em}}(x_1, x_2, \ldots, x_N, x_{N+1}) \\ & = \frac{\Phi_0^2}{16\pi b\lambda} \sum_{i,j=1}^{N+1} \cos(\phi_i - \phi_j)\left[\exp\left(-\frac{|x_i - x_j|}{\lambda} \right) \right. \\ & \left. - \exp\left(-\frac{x_i + x_j}{\lambda} \right) \right] + \frac{H_0 \Phi_0}{4\pi} \sum_{i=1}^{N+1} \cos(\phi_i - \theta) \\ & \times \left[\exp\left(-\frac{x_i}{\lambda} \right) - 1 \right] + G_s(x_{N+1}). \end{aligned} \tag{A5}$$

Here H_0 and θ are the magnitude of vector $\mathbf{H}$ and its tilt angle with respect to the z axis, and ϕ_i is the vortex tilt angle in the ith row. Assuming that all vortices are oriented along the z axis and only the $(N+1)$th row deviates by the angle ϕ, Eq. (A5) yields (3).

[1] C. P. Bean and J. D. Livingston, Phys. Rev. Lett. **12**, 14 (1964).
[2] F. F. Ternovskiĭ and L. N. Shekhata, Zh. Éksp. Teor. Fiz. **62**, 2297 (1972) [Sov. Phys. JETP **35**, 1202 (1972)].
[3] E. H. Brandt, Rep. Prog. Phys. **58**, 1456 (1995).
[4] R. Gauthier and M. A. R. LeBlanc, IEEE Trans. Magn. **MAG-13**, 560 (1977).
[5] E. H. Brandt, J. R. Clem, and D. G. Walmsley, J. Low Temp. Phys. **37**, 43 (1979).
[6] E. H. Brandt, J. Low Temp. Phys. **39**, 41 (1980).
[7] J. R. Clem, J. Low Temp. Phys. **38**, 353 (1980).
[8] J. R. Clem, Physica **107B**, 453 (1981).
[9] R. Boyer, G. Fillion, and M. A. R. LeBlanc, J. Appl. Phys. **51**, 1692 (1980).
[10] I. F. Voloshin, A. V. Kalinov, S. E. Savel'ev *et al.*, Zh. Éksp. Teor. Fiz. **111**, 1071 (1997) [JETP **84**, 592 (1997)].
[11] L. M. Fisher, A. V. Kalinov, S. E. Savel'ev *et al.*, Physica C **278**, 169 (1997).
[12] S. E. Savel'ev, L. M. Fisher, and V. A. Yampol'skii, Zh. Éksp. Teor. Fiz. **112**, 936 (1997) [JETP **85**, 507 (1997)].
[13] S. K. Hasanain, Sadia Manzoor, and A. Amirabadizadeh, Supercond. Sci. Technol. **8**, 519 (1995).
[14] F. Perez-Rodriguez, A. Perez-Gonzalez, J. R. Clem, G. Gandolfini, and M. A. R. LeBlanc, Phys. Rev. B **56**, 3473 (1997).
[15] G. D'Anna, Phys. Rev. Lett. **76**, 1924 (1996).
[16] G. D'Anna, Phys. Rev. B **54**, R780 (1996).
[17] I. V. Baltaga, L. M. Fisher, N. V. Il'in *et al.*, Phys. Lett. A **148**, 213 (1990).
[18] I. V. Baltaga, N. M. Makarov, V. A. Yampol'skii *et al.*, Fiz. Nizk. Temp. **21**, 411 (1995) [Low Temp. Phys. **21**, 320 (1995)].
[19] L. M. Fisher, A. V. Kalinov, I. F. Voloshin *et al.*, Solid State Commun. **97**, 833 (1996).
[20] V. V. Shmidt and G. S. Mkrtchyan, Usp. Fiz. Nauk **112**, 459 (1974) [Sov. Phys. Usp. **17**, 170 (1974)].
[21] V. V. Bryksin and S. N. Dorogovtsev, Zh. Éksp. Teor. Fiz. **102**, 1025 (1992) [JETP **75**, 558 (1992)].
[22] V. S. Gorbachev and S. E. Savel'ev, Zh. Éksp. Teor. Fiz. **109**, 1387 (1996) [JETP **82**, 748 (1996)].
[23] S. E. Savel'ev and V. S. Gorbachev, Zh. Éksp. Teor. Fiz. **110**, 1032 (1996) [JETP **83**, 570 (1996)].
[24] I. F. Voloshin, V. S. Gorbachev, S. E. Savel'ev *et al.*, JETP Lett. **59**, 55 (1994).
[25] L. M. Fisher, V. S. Gorbachev, S. E. Savel'ev *et al.*, Physica C **245**, 231 (1995).
[26] K. Funaki, M. Noda, and K. Yamafuji, Jpn. J. Appl. Phys., Part 1 **21**, 1580 (1982).
[27] J. R. Clem, Phys. Rev. B **26**, 2463 (1982).
[28] V. V. Shmidt, Zh. Éksp. Teor. Fiz. **61**, 398 (1971) [Sov. Phys. JETP **34**, 211 (1971)].

Translation provided by the Russian Editorial office.

''Unusual'' domain walls in multilayer systems: ferromagnet + layered antiferromagnet

V. D. Levchenko and Yu. S. Sigov

M. V. Keldysh Institute of Applied Mathematics, Russian Academy of Sciences, 125047 Moscow, Russia

A. I. Morozov and A. S. Sigov

Moscow State Institute of Radio Engineering, Electronics, and Automation (Technical University), 117454 Moscow, Russia

(Submitted 6 April 1998)
Zh. Éksp. Teor. Fiz. **114**, 1817–1826 (November 1998)

The structure and conditions for the onset of a new type of domain wall in multilayer systems comprising a ferromagnet and a layered antiferromagnet is investigated by numerical simulation. Domain walls occur as the result of frustrations produced by interface roughness, i.e., by the existence of atomic steps on them. The domain walls are investigated both in a ferromagnetic film on a layered antiferromagnetic substrate and in multilayer structures. It is shown that a domain wall broadens with increasing distance from the interface; this trend is attributed to the nontrivial dependence of the wall energy on the thickness of the layer. The structure of the domain walls in multilayer ferromagnet–layered antiferromagnet systems varies dramatically as a function of the energies of interlayer and in-layer exchange interactions between adjacent layers. © *1998 American Institute of Physics.* [S1063-7761(98)01911-8]

1. INTRODUCTION

The discovery of giant magnetoresistance has stimulated interest in multilayer structures consisting of alternating ferromagnetic (Fe or Co) and nonmagnetic (Cr, Cu, or Ag) metal layers.

The interaction between the magnetic layers is described by the Ruderman–Kittel–Kasuya–Yosida (RKKY) potential, which oscillates as a function of position. The oscillatory character of the interaction for a certain thickness of the nonmagnetic spacer layer imparts antiferromagnetic orientation to the magnetizations of the adjacent layers. When an external magnetic field is applied, the orientation changes to ferromagnetic, and this process is accompanied by a drop of up to tens of percent in the electrical resistance (hence the appellation ''giant'').

However, the interfaces are not perfectly smooth, and atomic steps exist on them. Under certain conditions, a change in the thickness of the spacer layer by one monatomic layer causes the exchange interaction between the layers to change sign. If the characteristic spacing of the steps on the interface is greater than a certain critical value, the situation becomes energetically favorable to partitioning of the magnetic layers into domains such that the magnetizations of the adjacent magnetic layers exhibit parallel and antiparallel orientations.[1] The width δ of the resulting domain walls is dictated by the competition of exchange interactions within the layers $J_{\parallel}$ and between magnetic layers $J_{\perp}$:

$$\delta \sim b\sqrt{J_{\parallel} l / J_{\perp} b}, \tag{1}$$

where b is the interatomic distance and l is the thickness of the magnetic layer. The width δ can be substantially less than the width of ordinary domain walls in a ferromagnet, but because $J_{\perp}$ decreases with the thickness d of the nonmagnetic spacer layer as d^{-2}, we have $J_{\parallel} \gg J_{\perp}$, and the characteristic values of δ are equal to hundreds of angstroms, greatly exceeding the layer thicknesses. In this case the variation of the width of the wall with distance from the interface is negligible.

If chromium is chosen for the nonmagnetic spacer layers, which have a thickness greater than 32 Å, then as the temperature is lowered in the chromium layers, antiferromagnetic ordering sets in with the generation of a plane spin-density wave.[2] When $d < 50$ Å, the wave is commensurate (AF_0 structure). A similar commensurate layered antiferromagnetic structure occurs when iron atoms ($>2\%$) are introduced into the chromium.[3] According to some data, a layered antiferromagnetic structure can occur in manganese layers.[4]

The onset of long-range order in the chromium layers has the effect that the interaction between the ferromagnetic layers no longer diminishes as d^{-2}, but is essentially independent of d.

The roughness of the interfaces, i.e., the existence of atomic steps on them, can render a uniform distribution of the ferromagnetic and antiferromagnetic order parameters in the layers energetically unfavorable, because the relative orientations of adjacent spins separated by an interface are opposite on opposite sides of a step (Fig. 1a).

If the distance between the steps is large enough, the formation of a domain wall is energetically favorable (Fig. 1b).[5,6] Owing to the increase in $J_{\perp}$, the width δ_0 of the domain wall near an interface can be of the order of the interatomic distance. In that event the variation of the width of the domain wall with increasing distance from the interface can become significant. The objective of the present study is to investigate the structure of the domain walls pro-

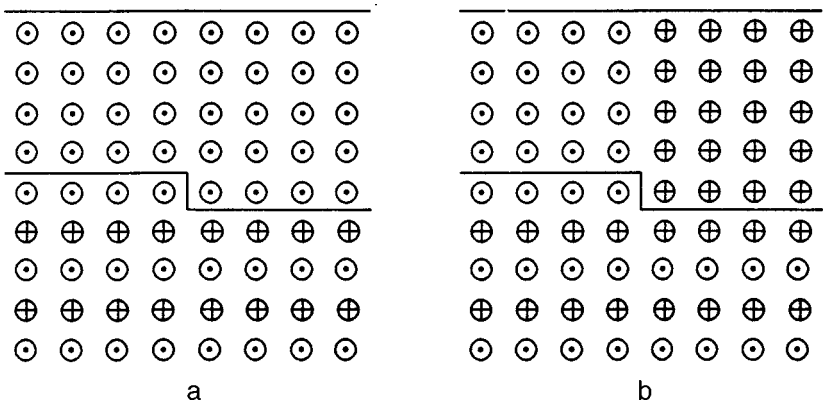


FIG. 1. Orientation of spins near the ferromagnet–layered antiferromagnet interface. a) Homogeneous distribution of the order parameters; b) in the presence of a domain wall.

duced by interface roughness and to determine the conditions underlying their onset.

2. FERROMAGNETIC THIN FILM ON A LAYERED ANTIFERROMAGNET

We consider a ferromagnetic thin film on a layered antiferromagnetic substrate. We investigate a solitary atomic step on the film–substrate interface. Let the step be parallel to the y axis of a Cartesian coordinate system, and let the z axis be perpendicular to the layer. This arrangement sets up a two-dimensional problem.

We investigate the exchange interaction between localized spins in the film–substrate system in the approximation of nearest-neighbor interaction. The Hamiltonian of the system has the form

$$\hat{\mathscr{H}} = -\frac{1}{2}\sum_{i,\delta} \mathscr{J}_{i,i+\delta}\hat{\mathbf{S}}_i\hat{\mathbf{S}}_{i+\delta}, \tag{2}$$

where $\hat{\mathbf{S}}_i$ is the ith localized spin operator, the index δ enumerates all nearest neighbors, and the exchange integral $\mathscr{J}_{i,i+\delta}$ is

$$\mathscr{J}_{i,i+\delta} = \begin{cases} \mathscr{J}_1 & \text{if } i,\ i+\delta \text{ refer to the film,} \\ \mathscr{J}_2(\delta) & \text{if } i,\ i+\delta \text{ refer to the substrate,} \\ \mathscr{J}_{12} & \text{otherwise,} \end{cases} \tag{3}$$

where $\mathscr{J}_1>0$, $\mathscr{J}_2(\delta)<0$ if i and $i+\delta$ are associated with different layers, and $\mathscr{J}_2(\delta)>0$ if i and $i+\delta$ are associated with the same layer. For definiteness we let $\mathscr{J}_{12}>0$.

By replacing $\mathscr{J}_2$ and the spin vectors associated with one of the sublattices of the antiferromagnet with their opposites, we can reduce the problem to the interaction of two ferromagnetic layers with frustrated interaction between the layers ($\overline{\mathscr{J}}_{12}=\mathscr{J}_{12}\operatorname{sgn}x$). The problem of a layered antiferromagnetic thin film on a ferromagnetic substrate is reducible to exactly the same problem.

When the thickness l of the substrate is much greater than the thickness d of the film (or, more precisely, when $|\mathscr{J}_2|l \gg \mathscr{J}_1 d$), a domain wall forms only in the bulk of the film, and the order parameter in the substrate remains homogeneous. We can therefore investigate the distribution of spins in the film for a given boundary condition at the interface.

Since we are investigating the spin distribution in the exchange approximation, their orientation in space and the plane of rotation in the domain wall are immaterial. In thin films the spins are oriented parallel to the layers. Consequently, we assume for definiteness that the spins to the right and left of the domain wall are parallel and antiparallel to the y axis, and the spin vectors rotate in the xy plane. Accordingly, in the substrate layer nearest the interface $\theta_i=0$ for $x>0$ and $\theta_i=\pi$ for $x<0$ (Fig. 1).

Replacing the spin operators in Eq. (2) by their mean values, assuming that the moduli of the latter are fixed, we obtain an expression for the energy in the mean-field approximation. After variation with respect to the quantities θ_i we go from a discrete to a continuous representation, assuming that the film thickness is much greater than the interatomic distance b ($d \gg b$). As a result, we obtain the following equation for the volume of the film:

$$\theta''_{xx}+\theta''_{zz}=0. \tag{4}$$

On the free boundary of the film with the vacuum at $z=a=d/b$, we have

$$\theta''_{xx}-\theta'_z=0. \tag{5}$$

At the film–substrate interface ($z=0$) we have

$$\theta''_{xx}+\theta'_z=\alpha\operatorname{sgn}x\sin\theta, \tag{6}$$

where $\alpha=\mathscr{J}_{12}\langle S_2\rangle/\mathscr{J}_1\langle S_1\rangle$, all derivatives in Eqs. (4)–(6) are evaluated with respect to dimensionless coordinates (normalized by the interatomic distance b), and $\langle S_1\rangle$ and $\langle S_2\rangle$ are the moduli of the mean spins of the film and substrate, respectively.

The boundary conditions (5) and (6) differ from those given in Ref. 7 by the term θ''_{xx}. This means, for example, that we can make a smooth transition from (6) to (4) when the substrate and film are made of the same material.

The solution of Eq. (4) subject to the boundary conditions (5) and (6) determines the distribution of the magnetization in the roughness-generated domain wall. It depends on the two dimensionless parameters α and a, where a is the dimensionless thickness of the film. In contrast, the width of ordinary domain walls is determined by the ratio of the exchange energy to the anisotropy energy, which we can disregard.

3. COMPUTATIONAL METHOD

The system of equations (4)–(6) is the Laplace equation for a function of two coordinates $\theta(x, z)$ with the nonlinear boundary conditions (6). We require in addition that the function θ be continuous in the range $0 \leqslant z \leqslant a$ and satisfy the conditions $\theta(x \to +\infty) \to 0$ and $\theta(x \to -\infty) \to \pi$.

To obtain a suitable equation for numerical solution, we first use a method analogous to the method of integral transforms.[8]. We transform the original set of differential equations to a single integral equation for a one-dimensional grid function $\psi(x_i)$. To do so, augmenting the definition of the function $\theta(x+2L, z) = \pi - \theta(x, z)$ in a sufficiently large region $x \in [-L, L]$, we take the discrete Fourier expansion of (4)–(6) along the x axis and then solve the decoupled equations for the Fourier harmonics analytically. We obtain

$$\psi(x_i) = \sum_k K(k, 0) \cos\left(\frac{\pi}{L} k x_i\right) \sum_j \cos\left(\frac{\pi}{L} k x_j\right) \alpha \operatorname{sgn} x_j \times \sin\psi(x_j) \equiv S(\psi, x_i), \tag{7}$$

where

$$K(k, z) = -4\left(\frac{L}{k\pi}\right)^2 \exp\left(-\frac{zk\pi}{2L}\right) \left\{ 2\left[\exp\left(\frac{ak\pi}{L}\right) + \exp\left(\frac{zk\pi}{L}\right)\right] \frac{L}{k\pi} + \left[\exp\left(\frac{ak\pi}{L}\right) - \exp\left(\frac{zk\pi}{L}\right)\right] \right\} \times \left\{ 4\left[\exp\left(\frac{ak\pi}{L}\right) - 1\right] \left(\frac{L}{k\pi}\right)^2 + 4\left[\exp\left(\frac{ak\pi}{L}\right) + 1\right] \frac{L}{k\pi} + \left[\exp\left(\frac{ak\pi}{L}\right) - 1\right] \right\}^{-1}. \tag{8}$$

To solve this equation numerically, we write the simple iterative scheme

$$\psi^{m+1}(x_i) = \psi^m(x_i) + C_F(x_i)[S(\psi^m, x_i) - \psi^m(x_i)], \tag{9}$$

where

$$C_F(x_i) = c_0 + \exp[-(x_i/x_T)^2], \tag{10}$$

with the initial condition $\psi^0(x_i) = \pi/2 - \tanh(10x_i/L)$. The free parameters c_0 and x_T are chosen experimentally in the course of the calculations in such a way as to maximize the rate of convergence of the iterative procedure without sacrificing its stability. We iterate until the residual η attains a value such that

$$\eta = \max_{x_i} |\psi^{m+1}(x_i) - \psi^m(x_i)| > \eta_0 = 10^{-5}. \tag{11}$$

We obtain a solution over the entire region from the expression

$$\theta(x_i, z) = \sum_k K(k, z) \cos\left(\frac{\pi}{L} k x_i\right) \times \sum_j \cos\left(\frac{\pi}{L} k x_j\right) \alpha \operatorname{sgn} x_j \sin\psi(x_j). \tag{12}$$

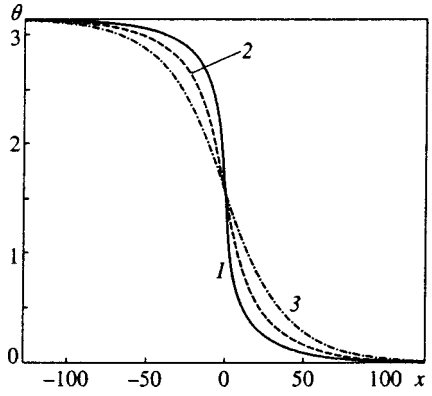


FIG. 2. Plots of $\theta(x)$ in a domain wall in various cross sections for $\alpha = 1/16$ and $a = 64$. *1*) $z = 0$; *2*) $z = 16$; *3*) $z = 64$.

We have used the foregoing method to perform a series of calculations for various combinations of the parameters $a = 4, 8, 16, 32, 64$ and $\alpha = 1/64, 1/16, 1/4, 1, 4$.

A similar method has been used in modeling the spin vortex described in Sec. 4, as well as multilayer structures (Sec. 5).

4. DISCUSSION

A typical plot of the function $\theta(x)$ is shown in Fig. 2. Note that the second derivative θ''_{xx} suffers a discontinuity at $x = z = 0$, while the first derivative θ'_z remains continuous there. The width of the domain wall $\delta(x)$ is interpreted as the distance between the points with coordinates (x_1, z) and (x_2, z), which correspond to $\theta_1 = \pi/4$ and $\theta_2 = 3\pi/4$, respectively.

The main distinguishing feature of the investigated domain walls is the fact that they increase with increasing distance from the interface. A graph of $\delta(z)$ for $\alpha a \gg 1$ is shown in Fig. 3. Clearly, the function is linear near the substrate and becomes essentially constant near the free surface. For $\alpha a \ll 1$ the variation of the width of the domain wall is negligible.

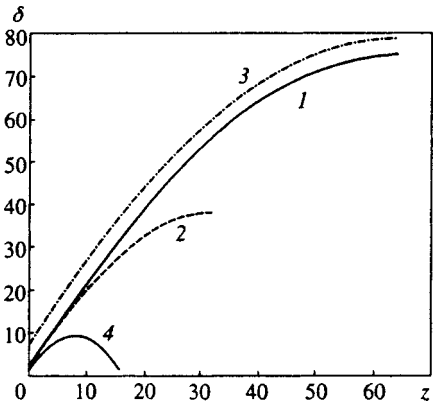


FIG. 3. Width of a domain wall vs. distance from the interface for $\alpha a \gg 1$ in the case of one free surface for *1*) $\alpha = 1$, $a = 64$; *2*) $\alpha = 1$, $a = 32$; *3*) $\alpha = 1/4$, $a = 64$, and in a three-layer structure for *4*) $\alpha = 3$ and $a = 16$.

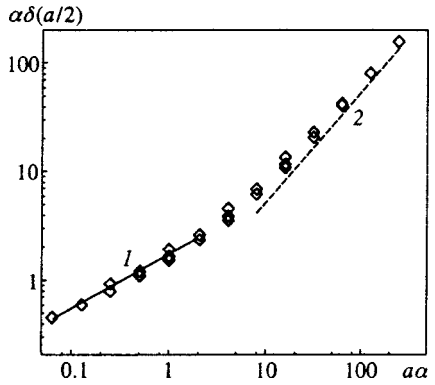


FIG. 4. Graph of $\alpha\delta(a/2)$ vs. the parameter $a\alpha$ (dots) and its approximation for *1)* $a\alpha \ll 1$, $\alpha\delta(a/2) \simeq 2\sqrt{a\alpha}$; *2)* $a\alpha \gg 1$, $\alpha\delta(a/2) \simeq a\alpha/2$.

The dimensionless width of the domain wall $\delta_0 = \delta(z=0)$ and a certain thickness-averaged δ_z', denoted from now on by β, can be estimated from simple energy considerations. We approximate $\theta(x,z)$ as follows:

$$\theta(x,z) = \begin{cases} 0 & \text{for } x \geqslant \delta(z), \\ \dfrac{\pi}{2}[1 - x/\delta(x)] & \text{for } -\delta(z) < x < \delta(z), \\ \pi & \text{for } x \leqslant -\delta(z), \end{cases} \tag{13}$$

where

$$\delta(z) = \delta_0 + \beta z, \quad 0 \leqslant z \leqslant a. \tag{14}$$

The energy contribution from the inhomogeneity of the order parameter in the domain wall, calculated per meter of its length along the y axis, is

$$W_1 = \frac{\mathscr{J}_1 \langle S_1 \rangle^2}{2b} \int_0^a dz \int_{-\infty}^{\infty} dx [(\theta_x')^2 + (\theta_z')^2]$$

$$\sim \frac{\pi^2 \mathscr{J}_1 \langle S_1 \rangle^2}{4b} \left(\frac{1}{\beta} + \frac{\beta}{3} \right) \ln \frac{\beta a + \delta_0}{\delta_0}. \tag{15}$$

The existence of a step causes the film–substrate interaction energy to increase by

$$W_2 = \frac{2\mathscr{J}_{12} \langle S_1 \rangle \langle S_2 \rangle}{b} \int_0^{\infty} dx [1 - \cos \theta(x,0)]$$

$$\sim \frac{2\mathscr{J}_{12} \langle S_1 \rangle \langle S_2 \rangle}{b} \delta_0. \tag{16}$$

We evaluate the parameters β and δ_0 by minimizing the energy W_1 with respect to the former and then minimizing the total energy of the domain wall

$$\tilde{W} = W_1 + W_2 \tag{17}$$

with respect to the latter.

For $a\alpha \ll 1$ we have

$$\beta \sim \sqrt{a\alpha}, \tag{18}$$

$$\delta_0 \sim \sqrt{a/\alpha}, \tag{19}$$

and for $a\alpha \gg 1$

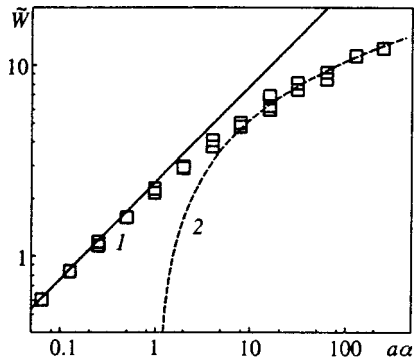


FIG. 5. Domain wall energy $\tilde{W}$ vs. $a\alpha$ (dots) and its approximation: *1)* $\tilde{W} \simeq 2\sqrt{a\alpha}$ for $a\alpha \ll 1$; *2)* $\tilde{W} \simeq 2\ln(a\alpha)$ for $a\alpha \gg 1$.

$$\beta \sim 1, \tag{20}$$

$$\delta_0 \sim 1/\min(1, \alpha). \tag{21}$$

The continuous approximation is valid if $\delta_0 \gg \alpha/(1+\alpha)$.

For the characteristic width of the domain wall $\delta(a/2)$ we obtain

$$\delta\left(\frac{a}{2}\right) \sim \begin{cases} \delta_0 \sim \sqrt{a/\alpha} & \text{for } a\alpha \ll 1, \\ a & \text{for } a\alpha \gg 1. \end{cases} \tag{22}$$

These estimates are in good agreement with the results of the calculations (Fig. 4). One exception is the expression for δ_0 when $a\alpha \gg 1$. For $\alpha > 1$ the results of the calculations differ appreciably from the estimate (21). In this case the neighborhood of the singularity at $x = z = 0$ specifically plays a major role, and the simple approximation (13) is invalid.

For the energy of the domain wall we obtain

$$\tilde{W} \sim \begin{cases} \dfrac{\mathscr{J}_2 \langle S_1 \rangle^2}{b} \sqrt{a\alpha} & \text{for } a\alpha \ll 1, \\ \dfrac{\mathscr{J}_1 \langle S_1 \rangle^2}{b} \ln(a\alpha) & \text{for } a\alpha \gg 1. \end{cases} \tag{23}$$

The calculated dependence of $\tilde{W}$ on $a\alpha$ is shown in Fig. 5.

The broadening of the domain wall has the effect that for $a\alpha \gg 1$ its energy increases only logarithmically as the film thickness is increased.

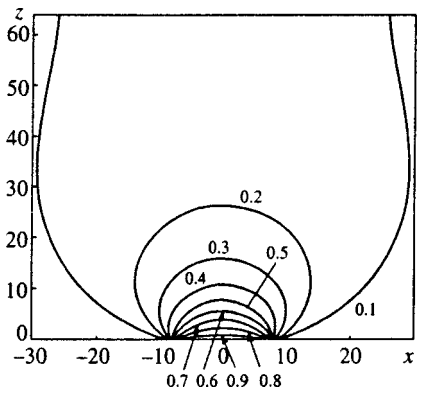


FIG. 6. Spin vortex generated for $R \ll a$ in the case $\alpha = 4$, $a = 64$, $R = 16$. The value of θ in units of π is indicated alongside each constant-θ line.

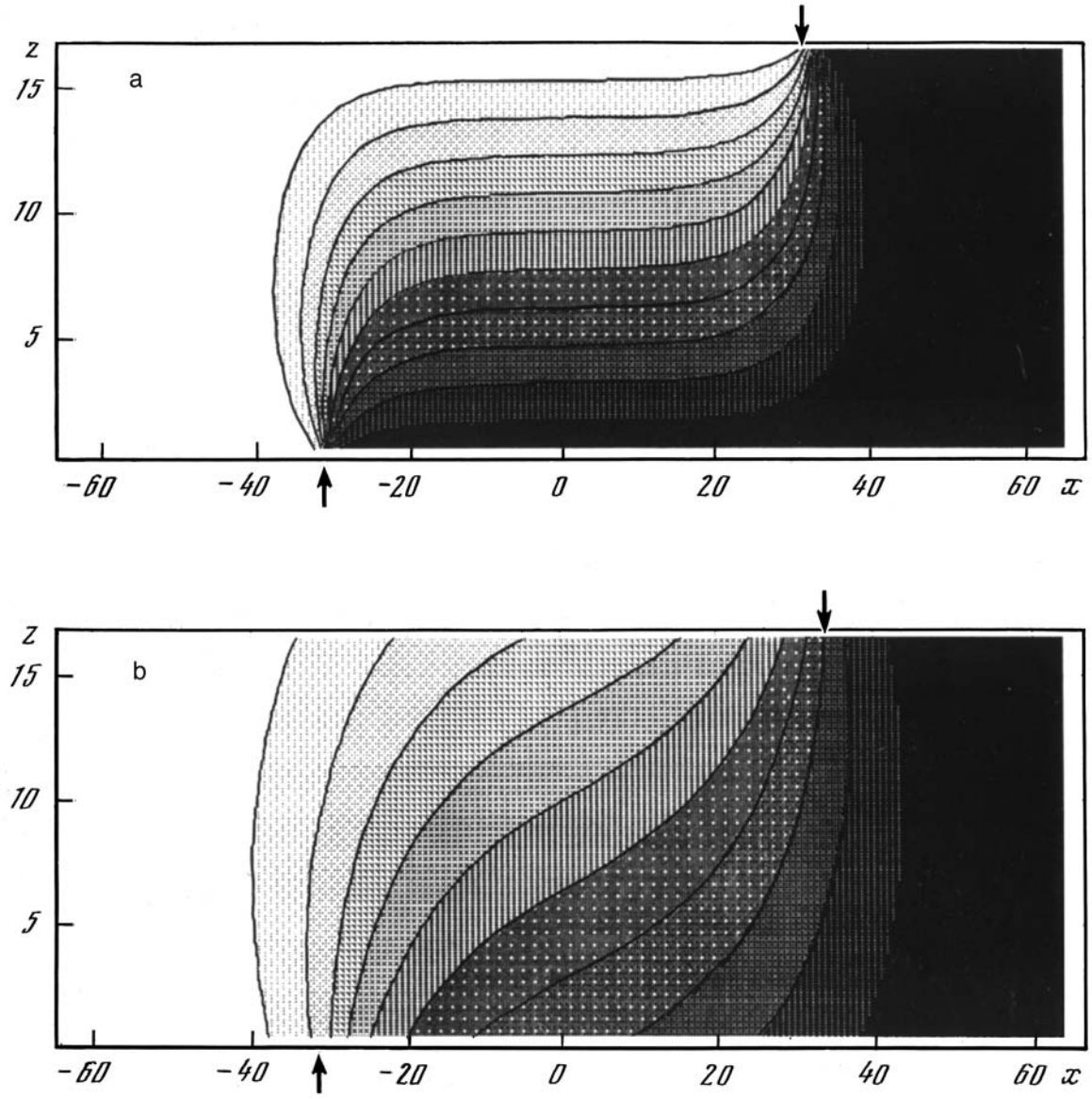


FIG. 7. Domain wall in a three-layer structure for $\alpha=3$, $a=16$, $R=64$ (a) and $\alpha=0.1$, $a=16$, $R=64$ (b). The step positions are indicated by arrows, and the constant-θ lines are drawn at intervals of $\pi/10$.

We now estimate the step spacing for which the uniform distribution of the order parameter in the film becomes energetically unfavorable. Let R be the dimensionless distance between two parallel steps, where the orientation of the spins in the film in the zone between steps corresponds to the maximum rather than to the minimum energy of interaction with the substrate. As a result, the energy of the system increases by $2\mathscr{J}_{12}\langle S_1\rangle\langle S_2\rangle R/b$ per meter of step length. If this energy is greater than the energy of two domain walls, it is more favorable for the film to acquire a domain with the opposite value of the order parameter. For $a\alpha\ll 1$ this condition is equivalent to the constraint $R>\delta_0$.

Two possibilities exist in the case $a\alpha\gg 1$. For $R\gg a$ two domain walls form, penetrating the entire thickness of the film. For $\delta_0\ll R\ll a$, on the other hand, a kind of spin vortex appears near the substrate (Fig. 6), penetrating the film to a depth of the order of R. Throughout the rest of the film the uniform distribution of the order parameter is undisturbed.

5. MULTILAYER STRUCTURES

Multilayer structures can be modeled when exchange interaction in the ferromagnetic layers (iron) is much stronger than in the antiferromagnetic layers (CrFe). Distortions of the uniform distribution of the order parameter occur only in the antiferromagnetic layers, where the distortions in one layer do not affect those in the other.

The behavior of a domain wall in an antiferromagnetic layer is determined entirely by the boundary conditions on its surface. As mentioned, the simultaneous sign reversal of both the exchange interaction between the antiferromagnetic sublattices and the direction of the spins of one of the sublattices reduces the problem of finding the distribution of the antiferromagnetic order parameter to finding the magnetization distribution in a ferromagnetic layer with frustrated interaction at the boundary.

The domain walls interconnect the nearest atomic steps,

which can be associated either with one boundary of the layer or with the opposite boundary. In contrast with the investigated case of one free boundary, the thickness of the domain wall is a maximum in the middle of the layer and not at the boundary (Fig. 3, curve *4*).

Modeling shows that the quantity α plays a major role. For $\alpha>1$ the spins closest to the boundaries of a layer are oriented parallel or antiparallel to the spins of the adjacent layers, depending on the sign of the exchange interaction between layers, to minimize the interlayer coupling energy. When the distance R between closest steps is much greater than the layer thickness ($R \gg a$), a domain wall formed on an atomic step rotates at a distance of the order of a and simultaneously broadens in such a way as to occupy the entire width of the layer and then run parallel to its boundaries (Fig. 7a). In this region, θ varies linearly with increasing z, from zero at one boundary of the layer to π at the other.

On the other hand, if $\alpha<1$ and $R \gg a$, the structure of the domain wall is more complex, because now the situation is energetically favorable for the spin vectors to rotate at the boundary of the layer through a certain angle relative to the direction corresponding to minimum surface energy (Fig. 7b). For $a\alpha<1$ we do not observe the zones, so conspicuous for $a\alpha>1$, where the domain wall rotates, and all distortions take place within a characteristic scale of the order of R. As in the case $a\alpha>1$, the domain wall fills essentially the entire volume of the layer in the space between steps.

In Ref. 9 the distribution of spins in a two-layer ferromagnet–layered antiferromagnet system is modeled numerically on the basis of the Ising model. However, the Ising model corresponds to very strong anisotropy of the easy-axis type and is invalid for the description of multilayer structures of the Fe/Cr type, for which the anisotropy energy in the plane of the layers is much lower than the exchange interaction energy. The domain walls formed within the framework of the Ising model have atomic widths, so that the unique properties of the domain walls in the layers of a ferromagnet–layered antiferromagnet were not observed in the cited paper.

The conditions for the emergence of the investigated domain walls can be tested experimentally and their structure analyzed by using a microscope to examine the magnetic forces of ferromagnetic films deposited on an antiferromagnetic substrate. The critical value of the characteristic step width R at which domain walls appear and its dependence on the film thickness can be found by investigating films of various thicknesses and with various widths R. Moreover, the dependence of the width of a domain wall at the surface on the film thickness can be determined. To within a factor of order unity, the theoretical dependence agrees with the graph in Fig. 4.

6. CONCLUSIONS

1. The existence of atomic steps on the interface between a layered antiferromagnet and a ferromagnetic film induces domain structure in the film if the distance R between steps is greater than the width δ_0 of the domain wall at the interface.

2. The width of the domain wall increases with increasing distance from the substrate. For thicker films, therefore, the energy of the wall increases logarithmically as the thickness of the film increases.

3. Instead of a domain structure, static spin vortices having a characteristic scale R occur near the substrate in thick films with $\delta_0 \ll R \ll a$.

4. In multilayer ferromagnet–layered antiferromagnet structures, the domain walls interconnect adjacent steps on the layer boundaries, occupying the entire space of the layer between the steps. The structure of such domain walls depends significantly on the relation between in-layer and interlayer exchange.

This work has received partial financial support from the Russian Fund for Fundamental Research (Grant No. 97-02-17627).

[1] A. I. Morozov and A. S. Sigov, JETP Lett. **61**, 911 (1995).
[2] E. E. Fullerton, S. D. Bader, and J. L. Robertson, Phys. Rev. Lett. **77**, 1382 (1997).
[3] E. Fawcett, H. L. Albert, V. Yu. Galkin *et al.*, Rev. Mod. Phys. **66**, 25 (1994).
[4] S. Bouarab, H. Nalt-Lazis, M. A. Khan *et al.*, Phys. Rev. B **52**, 10127 (1995).
[5] A. Berger and H. Hopster, Phys. Rev. Lett. **73**, 193 (1994).
[6] E. E. Fullerton, C. H. Sowers, and S. D. Bader, Phys. Rev. B **56**, 5469 (1997).
[7] J. C. Slonczewski, Phys. Rev. Lett. **67**, 3172 (1991).
[8] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vols. 1 and 2, Interscience, New York (1953, 1962).
[9] A. Berger and E. E. Fullerton, J. Magn. Magn. Mater. **165**, 471 (1997).

Translated by James S. Wood

Temperature dependence of the kinetic coefficients of interference ballistic structures

M. V. Moskalets

(Submitted 25 November 1997)
Zh. Éksp. Teor. Fiz. **114**, 1827–1835 (November 1998)

The effect of the temperature on the kinetic coefficients of a mesoscopic sample in contact with two electron reservoirs is considered for the case in which the electron transmission coefficient of the sample undergoes oscillations near the Fermi energy. © *1998 American Institute of Physics.* [S1063-7761(98)02011-3]

1. INTRODUCTION

One of the important properties of mesoscopic systems[1] at low temperatures is the preservation of phase coherence during propagation of electrons. Therefore, in such systems it is possible to observe effects that are sensitive to the phase of the electron wave function. As an example, we consider the Aharonov–Bohm effect,[2] which leads to oscillations with period $\Phi_0=h/e$ of the physical characteristics of nonsuperconducting doubly-connected samples immersed in a magnetic flux Φ (Refs. 3 and 4). This effect is observed in both the kinetic[5] and thermodynamic[6] properties of the samples.

The creation of structures in which electrons propagate ballistically has made it possible to imagine a number of devices based on the phenomenon of interference.[7–12] A peculiarity of such interference ballistic structures is the fact that the electron transmission coefficient $g(\varepsilon)$ of these structures varies substantially, $\Delta g \simeq g$, in response to an insignificant change in the energy $\Delta\varepsilon/\varepsilon \simeq \lambda/L$ (λ is the electron wavelength, L is the characteristic length of the device).

This fact is important in any consideration of electron and heat transport in an interference ballistic structure in contact with electron reservoirs. On the one hand, in the given case the range of temperature T and voltage V over which the linear-response approximation is valid is substantially narrowed: $\Delta T,\ V \ll \varepsilon_F \lambda_F/L$, where ε_F is the Fermi energy and λ_F is the wavelength of a Fermi electron. On the other hand, in such a system the electron thermal voltage coefficient should be significant.[13] This leads, in particular, to a violation of the Wiedemann–Franz law despite the fact that an interference ballistic structure is a purely elastic scatterer.

Transport phenomena in mesoscopic samples in contact with electron reservoirs at $T \neq 0$ are examined in Refs. 13 and 14, which consider both particle transport between electron reservoirs and electron transport only through a mesoscopic structure. In the first case, the calculation of the current (heat flux) takes into account both the resistance of the mesoscopic sample itself and the additional contact resistance (spreading resistance) between the leads and the electron reservoirs.[13,15] In the present paper, we consider only the first case (transport between electron reservoirs) as it more closely corresponds to the standard experimental situation (two-probe measurements). The feasibility of measuring the transport coefficients associated with the mesoscopic sample itself is discussed in Refs. 13 and 14.

We apply the results of Refs. 13 and 14 to the case in which the electron transmission coefficient $g(\varepsilon)$ of the mesoscopic sample oscillates as ε varies near the Fermi energy. In Sec. 2 we obtain an expression for the current I and heat flux Q at nonzero reservoir temperatures, with allowance for the oscillating nature of the dependence $g(\varepsilon)$. Section 3 considers the linear response regime. Section 4 derives expressions for I and Q for a large temperature difference between the reservoirs. Section 5 considers a one-dimensional ring enclosing magnetic flux Φ in terms of an interference ballistic structure, and shows that the presence of this flux can substantially alter the temperature dependence of the transport coefficients.

2. STATEMENT OF THE PROBLEM AND BASIC RELATIONS

Let a ballistic mesoscopic sample be connected by ballistic leads to two electron reservoirs, ''L'' and ''R'', The size of the sample is assumed to be small in comparison with the phase coherence length $L_\varphi(T)$. We denote the temperatures and chemical potentials of the reservoirs respectively by T_L, μ_L and T_R, μ_R. In the present paper, we consider a one-dimensional sample and one-dimensional leads. Generalization to the case of several conducting subzones in the absence of channel mixing is trivial. With channel mixing taken into account such a generalization can be made according to the results of Refs. 13 and 15.

Expressions for the current I and dissipative heat loss Q between electron reservoirs in the single-channel approximation have the form

$$I=\frac{2e}{h}\int d\varepsilon\, g(\varepsilon)\left[f_0\left(\frac{\varepsilon-\mu_L+e\varphi}{T_L}\right)-f_0\left(\frac{\varepsilon-\mu_R+e\varphi}{T_R}\right)\right], \quad (1)$$

$$Q=\frac{2}{h}\int d\varepsilon\, g(\varepsilon)\left[(\varepsilon-\mu_L+e\varphi)f_0\left(\frac{\varepsilon-\mu_L+e\varphi}{T_L}\right)\right.$$
$$\left.-(\varepsilon-\mu_R+e\varphi)f_0\left(\frac{\varepsilon-\mu_R+e\varphi}{T_R}\right)\right]. \quad (2)$$

Here $g(\varepsilon)$ is the electron transmission coefficient of the mesoscopic sample as a function of electron energy ε, $f_0(x)=(1+e^x)^{-1}$ is the Fermi distribution function, φ is the potential of the sample relative to the electron reservoirs. The quantity φ should be determined self-consistently from the Poisson equation.[16–18]

Below we assume that the energy dependence of the transmission coefficient, $g(\varepsilon)$, derives from interference processes. Therefore, the transmission coefficient is an oscillating function of electron wavelength $\lambda = 2\pi/k$. We write

$$g(\varepsilon) = \overline{g} + \sum_n g_n, \tag{3}$$

$$g_n = A_n \cos(nkL) + B_n \sin(nkL).$$

Here $\overline{g}$ is the mean transmission coefficient, $\varepsilon = (\hbar k)^2/2m^*$, and m^* is the effective mass of the electron. If

$$L \gg \lambda_F, \tag{4}$$

then near the Fermi energy the quantity $g(\varepsilon)$ can accurately be assumed to be an oscillating function of the electron energy. Substituting the expansion (3) into expressions (1) and (2) in this case and assuming that

$$T_L \ll \mu_L,\ T_R \ll \mu_R, \quad \Delta\mu = \mu_R - \mu_L \ll \mu_L, \mu_R, \tag{5}$$

we obtain

$$I = -\frac{2e}{h}\left\{ \overline{g}\Delta\mu + \pi^2 T^{*2} \sum_n \frac{1}{n} (\psi_n(T_L)\partial_\varepsilon g_n|_{\mu_L - e\varphi} - \psi_n(T_R)\partial_\varepsilon g_n|_{\mu_R - e\varphi}) \right\}, \tag{6}$$

$$Q = \frac{2\pi^2}{h}\left\{ \overline{g}\frac{T_L^2 - T_R^2}{6} + T^{*2}\sum_n \frac{\Delta\mu}{n}(\psi_n(T_L)\partial_\varepsilon g_n|_{\mu_L - e\varphi} + \psi_n(T_R)\partial_\varepsilon g_n|_{\mu_R - e\varphi}) + T^{*2}\sum_n \left(\cosh\frac{nT_L}{T^*}\psi_n^2(T_L) g_n(\mu_L - e\varphi) - \cosh\frac{nT_R}{T^*}\psi_n^2(T_R) g_n(\mu_R - e\varphi) \right) \right\}. \tag{7}$$

Here $\psi_n(T) = (T/T^*)[\sinh(nT/T^*)]^{-1}$, $T^* = \Delta_F/2\pi^2$, $\Delta_F = 2\varepsilon_F \lambda_F/L$, and we use the notation $\partial_\varepsilon \equiv \partial/\partial\varepsilon$. The Fermi energy $\varepsilon_F = \mu - e\varphi_0$ is defined in equilibrium: $\mu = \mu_L = \mu_R, T = T_L = T_R$. The quantity $\Delta\mu$ is determined both by the difference of electrostatic potentials V applied to the electron reservoirs and the difference of chemical potentials of the electron reservoirs due to their temperature dependence. Note that the dependence $\mu(T)$ is absent for a two-dimensional reservoir.

Expressions (6) and (7) are central to the present work and will be used below to determine the temperature dependence of the kinetic coefficients. We assume that inelastic processes are absent in the sample, $L \ll L_\varphi(T)$, and that the effect of the temperature reduces merely to energy averaging of the transmission coefficients $g(\varepsilon)$ (3).

3. KINETIC COEFFICIENTS IN THE LINEAR REGIME

For small differences of the reservoir chemical potentials $\Delta\mu$ and temperatures $\Delta T = T_R - T_L$

$$\Delta\mu \ll \Delta_F, \quad \Delta T \ll \min(T, T^*), \tag{8}$$

where Δ_F is the distance between electron energy levels near the Fermi level, the expressions for the current I and heat flux Q can be represented in matrix form:[13]

$$\begin{pmatrix} I \\ Q \end{pmatrix} = -\begin{pmatrix} L_0 & L_1/T \\ L_1 & L_2/T \end{pmatrix}\begin{pmatrix} \Delta\mu/e \\ \Delta T \end{pmatrix}. \tag{9}$$

In the present paper, we consider kinetic coefficients describing charge and energy transport between the electron reservoirs. For this case, as emphasized in Ref. 13, the Onsager relations are satisfied, as reflected in Eq. (9).

Expressions for the kinetic coefficients L_i can be obtained from expressions (6) and (7). The temperature T^* divides the low-temperature from the high-temperature range; in the former, the effect of the temperature can be neglected, while in the latter, averaging over energy leads to compensation of the interference contributions to the transmission coefficient, as a result of which $g(\varepsilon) \simeq \overline{g}$.

We now write out asymptotic expressions for the coefficients L_i.

1) Low temperatures: $T \ll T^*$,

$$L_0 = \frac{2e^2}{h} g(\varepsilon_F), \tag{10a}$$

$$L_1 = \frac{2e\pi^2}{3h} T^2 \partial_\varepsilon g|_{\varepsilon_F}, \tag{10b}$$

$$L_2 = \frac{2\pi^2}{3h} T^2 g(\varepsilon_F). \tag{10c}$$

2) High temperatures: $T \gg T^*$,

$$L_0 = \frac{2e^2}{h}\left[\overline{g} + \frac{2T}{T^*}\exp\left(-\frac{T}{T^*}\right) g_1(\varepsilon_F) \right], \tag{11a}$$

$$L_1 = \frac{4e\pi^2}{h} T^2 \exp\left(-\frac{T}{T^*}\right)\partial_\varepsilon g_1|_{\varepsilon_F}, \tag{11b}$$

$$L_2 = \frac{2e^2}{3h} T^2 \left[\overline{g} - \frac{6T}{T^*}\exp\left(-\frac{T}{T^*}\right) g_1(\varepsilon_F) \right]. \tag{11c}$$

We also write out expressions for the electron thermal voltage coefficient $\alpha = -\Delta\mu/e\Delta T$ and the thermal conductivity $\kappa = -Q/\Delta T$, which are measured at $I = 0$:

1) $T \ll T^*$,

$$\alpha = \frac{\pi^2}{3e} T \partial_\varepsilon \ln(g)|_{\varepsilon_F}, \tag{12a}$$

$$\kappa = \frac{2\pi^2}{3h} T g(\varepsilon_F)\left[1 - \frac{3(e\alpha)^2}{\pi^2} \right]. \tag{12b}$$

2) $T \gg T^*$,

$$\alpha = \frac{\pi^2}{e} T^* (\overline{g})^{-1} \partial_\varepsilon g_1|_{\varepsilon_F}, \tag{13a}$$

$$\kappa = \frac{2\pi^2}{3h} T \left[\overline{g} - \frac{6T}{T^*}\exp\left(-\frac{T}{T^*}\right) g_1(\varepsilon_F) \right]. \tag{13b}$$

Expressions (10), obtained in Ref. 13, and expressions (12) have the usual form for the theory of metals (see, e.g., Ref. 19). However, in contrast to normal metals, for which $e\alpha_0 \simeq T/\varepsilon_F$ (without allowance for phonon entrainment), in the case under consideration the electron thermal voltage is not small: $\alpha/\alpha_0 \simeq \varepsilon_F/T^* \gg 1$. This leads, in particular, to a breakdown of the Wiedemann–Franz law:

$$\frac{3e^3}{\pi^2}\frac{\kappa}{GT}=\begin{cases}1-3(e\alpha)^2/\pi^2, & T\ll T^*,\\ 1-\dfrac{8T}{T^*}\exp\left(-\dfrac{T}{T^*}\right)\dfrac{g_1(\varepsilon_F)}{\overline{g}}, & T\gg T^*.\end{cases} \tag{14}$$

Here $G=L_0$ is the conductance of the system. The greatest deviation should be observed for $T\simeq T^*$; it decreases at both lower and higher temperatures. We emphasize that in the present work we derive the deviation from the Aharonov–Bohm law for the transport coefficients between the electron reservoirs, whereas in Refs. 13 and 14 such a deviation was derived for the coefficients associated with the mesoscopic sample itself.

Comparing expressions (12a) and (13a), we see that $\alpha(T)$ has a maximum (in absolute value) at $T\simeq T^*$. To order of magnitude, the maximum is $|e\alpha(T^*)|\simeq 1$. Note also that the sign of α can be different for $T\ll T^*$ and $T\gg T^*$.

The conductance G of the system and the thermal conductivity κ for $T\ll T^*$ depend on the Fermi energy ε_F, which in turn depends on the chemical potentials of the electron reservoirs. For $T\gg T^*$ the small interference terms in the conductance and the thermal conductivity exhibit such a dependence (see (11a) and (13b)).

4. CURRENT AND HEAT FLUX FOR A LARGE TEMPERATURE DIFFERENCE BETWEEN THE ELECTRON RESERVOIRS

We fix the temperature of one of the reservoirs, e.g., T_L, and consider the dependence of the current and heat flux in the system on the temperature of the second reservoir (T_R). As follows from Eq. (6), in the limit

$$T_R \gg T^* \tag{15}$$

the current I is independent of T_R. Moreover, if we write

$$Q=Q_0+\delta Q, \quad Q_0=\frac{\pi^2}{6e^2}\overline{g}(T_L^2-T_R^2), \tag{16}$$

then it follows from Eq. (7) that the interference term δQ in the heat flux under condition (15) is also independent of T_R.

Inasmuch as inelastic processes take place only in the reservoirs and not in the sample, temperature averaging of the contributions to the current (and heat flux Q) due to electrons propagating from the left reservoir to the right reservoir (with temperature T_L) and in the reverse direction (with temperature T_R) takes place independently, as is immediately clear from (6) and (7).

In what follows we set $\mu_L=\mu_R$. If condition (15) is met, the dependence on T_L of the thermoelectric current I_T and the interference term δQ in the heat flux have the following form:

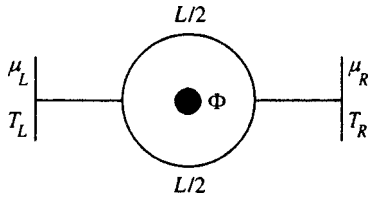


FIG. 1. Model of a one-dimensional ballistic ring of length L enclosing a magnetic flux Φ. The ring is connected by one-dimensional wires to two electron reservoirs at temperatures T_L and T_R and chemical potentials μ_L and μ_R.

1) $T_L \ll T^*$,

$$I_T=\frac{2e}{h}\int_{\varepsilon_0}^{\varepsilon_F} d\varepsilon[g(\varepsilon)-\overline{g}], \tag{17a}$$

$$\delta Q=\frac{2}{h}\int_{\varepsilon_0}^{\varepsilon_F} d\varepsilon(\varepsilon-\varepsilon_F)[g(\varepsilon)-\overline{g}]. \tag{17b}$$

Here $\varepsilon_0=\Delta_F[\varepsilon_F/\Delta_F]$ ($[x]$ is the integer part of x). In this regime, the thermoelectric current and the interference term in the heat flux are independent of the temperatures of the electron reservoirs T_L and T_R.

2) $T_L \gg T^*$, $T_R > T_L$,

$$I_T=-\frac{4e\pi^2}{h}T_L T^*\exp\left(-\frac{T_L}{T^*}\right)\partial_\varepsilon g_1|_{\varepsilon_F}, \tag{18a}$$

$$\delta Q=\frac{4\pi^2}{h}T_L^2\exp\left(-\frac{T_L}{T^*}\right)g_1(\varepsilon_F). \tag{18b}$$

Thus, the difference between the regime with a large temperature difference and the linear response regime ($\Delta T \ll T$) is that the thermoelectric current and the interference term in the heat flux are independent of the temperature of the hotter reservoir (T_R), and fail to vanisih in the limit $T_L \to 0$.

5. TEMPERATURE DEPENDENCE OF THE KINETIC COEFFICIENTS OF A ONE-DIMENSIONAL BALLISTIC RING WITH MAGNETIC FLUX

Consider a one-dimensional ballistic ring connected by one-dimensional leads to two electron reservoirs (Fig. 1). The ring encloses a magnetic flux Φ. The length of the ring L is assumed to be small in comparison with the phase coherence length: $L \ll L_\varphi$. Within the framework of quantum waveguide theory[20] in the approximation of noninteracting electrons, the transmission coefficient g for such a system for symmetrically located contacts is[20,21]

$$g=\frac{(1-\cos(kL))(1+\cos(2\pi\Phi/\Phi_0))}{(\cos(2\pi\Phi/\Phi_0)-1.25\cos(kL)+0.25)^2+\sin^2(kL)}. \tag{19}$$

The first two terms in expansion (3) have the form

$$\overline{g}=\frac{2(1+\cos(2\pi\Phi/\Phi_0))}{3+2\cos(2\pi\Phi/\Phi_0)}, \tag{20}$$

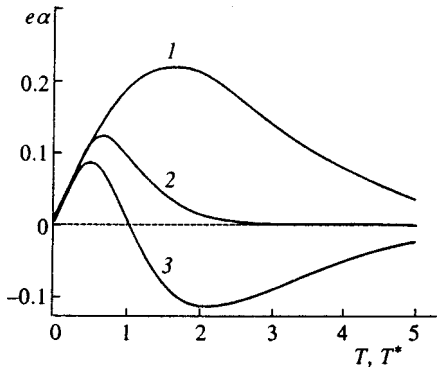


FIG. 2. Dependence of the thermoelectric coefficient α of a one-dimensional ring on the temperature T in the linear response regime ($\Delta T \ll T$) for $\Phi=0$ (*1*), Φ^* (*2*), and $0.15\Phi_0$ (*3*). $\{L/\lambda_F\}=0.75$.

$$g_1=\frac{2}{9}g\left(4\cos\left(\frac{2\pi\Phi}{\Phi_0}\right)-3\right)\cos(kL). \tag{21}$$

The kinetic coefficients of the system depend on the magnetic flux Φ, and this dependence can be manifested in a nontrivial way. We show this in the case of the thermoelectric coefficient α.

Figure 2 plots the dependence $\alpha(T)$ for three values of Φ and $\{L/\lambda_F\}=0.75$, where $\{x\}$ is the fractional part of x. The quantity α is governed by the transmission coefficient g for $T<T^*$ (see Eq. (12a)), and by its first harmonic g_1 for $T>T^*$ (see Eq. (13a)). The signs of the derivatives $\partial_\varepsilon g$ and $\partial_\varepsilon g_1$ depend on the magnetic flux Φ (and on the product k_FL). For some values of Φ these quantities have the same sign (Fig. 2, curve *1*) while for others their signs are different (Fig. 2, curve *3*). Therefore, by varying the magnetic flux Φ it is possible to go from a fixed-sign dependence of $\alpha(T)$ to a dependence that changes sign with increasing or decreasing temperature.

Furthermore, for $\Phi=\pm\Phi^*$ (where $\Phi^*/\Phi_0=(2\pi)^{-1}\arccos(3/4)\simeq 0.115$) the quantity g_1 (see Eq. (21)) governing the high-temperature ($T>T^*$) asymptotic limit of the thermoelectric coefficient α vanishes. Therefore, in expansion (6) it is necessary to retain the $n=2$ term. As a result, for $\Phi=\pm\Phi^*$ and $T>T^*$ we obtain

$$e\alpha\simeq\exp\left(-\frac{2T}{T^*}\right)\frac{2T}{T^*}\sin(2k_FL). \tag{22}$$

In the given case α falls off more rapidly with temperature (than for $\Phi\neq\pm\Phi^*$) (Fig. 2, curve *2*). Moreover, the temperature separating the low-temperature and high-temperature regions is reduced by a factor of two.

The interference terms in the conductance $G=L_0$ (see Eq. (11a)) and the thermal conductivity κ (see Eq. (13b)) have an analogous dependence on Φ and T for $T>T^*$. Specifically, they change sign with varying Φ; moreover, their magnitude falls off exponentially with increasing temperature, with a characteristic temperature of $T^*/2$ for $\Phi=\pm\Phi^*$ and T^* for other values of magnetic flux Φ.

For $\Phi\neq 0$ the conductance of the ring (19) vanishes near integer values of the ratio $k_FL/2\pi=L/\lambda_F$. Therefore, near these values the thermoelectric coefficient α has a significant value (at low temperatures). It can be seen from Fig. 3 that as the ratio L/λ_F approaches an integer, the peak in the dependence $\alpha(T)$ becomes sharper and shifts toward lower temperatures. Note that when we make the substitution $\{L/\lambda_F\}\to 1-\{L/\lambda_F\}$ the thermoelectric coefficient changes sign.

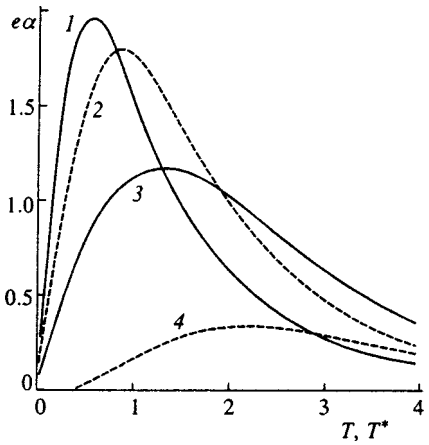


FIG. 3. Dependence of the thermoelectric coefficient α of a one-dimensional ring on the temperature T in the linear response regime ($\Delta T \ll T$) for $\{L/\lambda_F\}=0.05$ (*1*), 0.1 (*2*), 0.2 (*3*), and 0.4 (*4*). $\Phi=0.3\Phi_0$.

6. CONCLUSION

In the present paper we have considered the effect of the temperature on the kinetic coefficients of a mesoscopic sample in contact with two electron reservoirs. Charge and energy transport are described by the electron transmission coefficient $g(\varepsilon)$ of the sample. It is assumed that inelastic processes are absent in the sample and that the effect of temperature T reduces simply to energy averaging of the electron transmission coefficient g in a neighborhood $\Delta\varepsilon\simeq T$ about ε_F. Expressions have been obtained for the current I (6) and dissipative heat flux Q (7) which take account of oscillations in g near ε_F. Such oscillations are characteristic of ballistic mesoscopic systems in which interference has a substantial influence on the transmission of an electron wave.[7–12] Note that in contrast to the Sommerfeld expansion customarily used in the theory of normal metals, which takes account of variations of physical quantities of the system on a scale ε_F (or more precisely, on a scale $\Delta\varepsilon \gg T$), the expansion proposed here takes account of variation of physical characteristics of the system near ε_F in the energy range $\Delta\varepsilon\ll T,\varepsilon_F$.

We have shown that the thermoelectric coefficient α significantly exceeds the value α_0 characteristic of ordinary metals (without allowance for phonon entrainment and in the absence of magnetic impurities), $\alpha\simeq\alpha_0\varepsilon_F/T^*$, which is due to oscillations in the electron transmission coefficient near ε_F (Ref. 13). We have calculated the dependence $\alpha(T)$. In the linear regime ($\Delta T\ll T^*,T$) the dependence $\alpha(T)$ has a maximum (in absolute value) at $T\simeq T^*$. Note that the sign of α can be different for $T<T^*$ and for $T>T^*$. At high temperatures ($T>T^*$) the thermoelectric coefficient falls exponentially with temperature, $\alpha\propto\exp(-2\pi^2T/\Delta_F)$, and this dependence can be used to determine Δ_F. It should be noted

that the present work treats the one-channel approximation. For a multichannel sample without channel mixing, the high-temperature asymptotic limit of the dependence $\alpha(T)$ has a form analogous to that of the dependence $\Delta_F = \max(\Delta_F(n))$, where n is the channel number.

In the case of the Aharonov–Bohm effect for doubly-connected, one-dimensional samples, we predict a nontrivial dependence of the thermoelectric coefficient and the interference terms in the conductance and the thermal conductivity at high temperatures ($T > T^*$) on the temperature and magnetic flux.

[1] Y. Imry, in *Physics of Mesoscopic Systems: Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenco (eds.), World Scientific, Singapore (1986), p. 101.
[2] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 484 (1959).
[3] N. Byers and C. N. Yang, Phys. Rev. Lett. **7**, 46 (1961).
[4] F. Bloch, Phys. Rev. B **2**, 109 (1970).
[5] S. Washburn and R. A. Webb, Adv. Phys. **35**, 375 (1986).
[6] A. A. Zvyagin and I. V. Krive, Fiz. Nizk. Temp. **21**, 687 (1995) [Low Temp. Phys. **21**, 533 (1995)].
[7] S. Datta, Superlattices Microstruct. **6**, 83 (1989).
[8] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, J. Appl. Phys. **66**, 3892 (1989).
[9] F. Capasso and S. Datta, Phys. Today **43**, No. 2, 74 (1990).
[10] S. Datta and M. J. McLennan, Rep. Prog. Phys. **53**, 1003 (1990).
[11] W. Porod, Z. Shao, and C. S. Lent, Phys. Rev. B **48**, 8495 (1993).
[12] A. M. Jayannavar and P. Singha Deo, Mod. Phys. Lett. B **8**, 301 (1994).
[13] U. Sivan and Y. Imry, Phys. Rev. B **33**, 551 (1986).
[14] H.-L. Engquist and P. W. Anderson, Phys. Rev. B **24**, 1151 (1981).
[15] H. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).
[16] M. Büttiker and T. Christen, *Mesoscopic Electron Transport*, L. Kowenhoven, G. Schoen, and L. Sohn (eds.), NATO ASI Series E, Kluwer Academic Publishers, Dordrecht, to be published.
[17] M. V. Moskalets, JETP Lett. **62**, 719 (1995).
[18] M. V. Moskalets, Fiz. Nizk. Temp. **23**, 857 (1997) [Low Temp. Phys. **23**, 644 (1997)].
[19] A. A. Abrikosov, *Fundamentals of the Theory of Metals*, North-Holland, Amsterdam (1988).
[20] J. Xia, Phys. Rev. B **45**, 3593 (1992).
[21] H. Büttiker, Y. Imry, and M. Ya. Azbel, Phys. Rev. A **30**, 1982 (1984).

Translated by Paul F. Schippnick

NMR spectra under skin-effect conditions at ultralow temperatures

K. O. Khutsishvili*) and N. P. Fokina†)

I. Dzhavakhishvili Tbilisi State University, 380028 Tbilisi, Georgia

G. Eska‡)

Physikalisches Institut, Universitä Bayreuth, D-95440 Bayreuth, Germany

(Submitted 9 December 1997)
Zh. Éksp. Teor. Fiz. **114**, 1836–1847 (November 1998)

An investigation based on the coupled Maxwell–Bloch equations for a system of equivalent exchange-coupled spins is performed in order to explain a number of features of NMR spectra obtained in metals by Fourier-transforming of the free-induction decay at ultralow temperatures. Small angles of tilting of the nuclear magnetization by the exciting rf field are considered. It is shown that the free precession inherits the nonuniformity in the distribution of the rf field and the magnetization produced at the excitation stage inside the sample on account of the skin effect. As a result, the NMR spectrum is found to consist of a set of peaks—signals due to standing spin waves. However, such a spectrum can be observed only when the detuning of the exciting rf field is sufficiently large relative to the Larmor frequency of the spins. Otherwise, the rf field does not penetrate into the sample because of strong absorption by the spins. If the detuning is large, the dispersion signal and part of the NMR absorption signal are proportional to the equilibrium magnetization to the power 3/2. Such behavior is expected at low temperatures so that the coupling of the magnetization with the rf field is strong. The results obtained qualitatively explain the experimentally observed characteristics of the NMR spectra: the presence of kinks and structure of the NMR lines, the dependence of the shape and intensity of the spectrum on the detuning of the exciting rf field, and the nonlinear dependence of the nuclear susceptibility on the reciprocal of the sample temperature. [S1063-7761(98)02111-8]

1. INTRODUCTION

A series of experiments on the detection of NMR from nuclear spins in metals at ultralow temperatures was performed recently.[1–3] The NMR spectra were obtained by Fourier-transforming of the free induction decay and had a number of distinctive features: the dependence of the spectrum shape and intensity on the amplitude of the exciting rf field and the detuning of the field from the Larmor frequency of the nuclear spins, nonlinear dependence of the nuclear susceptibility on the reciprocal of the sample temperature, and presence of kinks and structure in the NMR lines. These features have still not been explained. In Ref. 2 it was indicated that the metallic properties of the samples (the skin effect) play a decisive role in the appearance of the characteristics of the NMR spectra obtained in such experiments. Specifically, the opinion was stated that the above-noted features of the NMR spectra are due to standing spin waves, on account of the skin effect, excited as the free induction decays.

The existence of spin waves is possible because when the rf field penetrates nonuniformly into a conductor (the skin effect) the spatial distribution of the magnetization is also nonuniform. The Ruderman–Kittel interaction between nuclear spins, which is of an exchange character, can therefore be written in the form of a nonuniform exchange energy.[4] However, as shown in Ref. 4, even in the presence of such an energy a uniform rf field can excite in the metal only a uniform precession of the magnetic moments, i.e., a spatial mode with wave vector $\mathbf{K}=0$ (we emphasize that in Ref. 4 a time-independent rf field was studied). On the other hand, Kittel[5] established that if surface anisotropy is present in the sample, then a stationary uniform field can excite observable standing spin waves with wave vectors $\mathbf{K}=0$ as well as $\mathbf{K}\neq\mathbf{0}$. The mathematical formulation of this problem is based on the fact that under the surface anisotropy conditions the magnetization satisfies different equations of motion at the surface and in the interior of the sample, and this gives rise to the distortion of the spin waves required for their observation. Therefore, in samples into which a uniform rf field penetrating unhindered and is uniform, the only reason that spin waves can be observed under conditions of stationary irradiation is that the equations of motion of the magnetization vector at the surface and in the interior of the sample have a different form.

In contrast to stationary NMR, free induction decay is a two-step process consisting of a preparatory stage (excitation of transverse magnetization) and free precession. In our opinion an external uniform rf field and the transverse magnetization excited by it propagate into the sample during the preparation stage in the form of two waves (represented in the form of ''field-like'' and ''spin-like'' branches) that penetrate into the sample nonuniformly because of the skin effect. The subsequent free precession inherits the nonunifor-

mity produced, since free oscillations of the magnetization and of the magnetic field due to its motion occur with spatially nonuniform initial conditions. This could be responsible for the fact that the amplitudes of the standing spin waves of magnetization, which are present during free precession, with ''quantized'' values of **K** in the sample are different from zero for $\mathbf{K}\neq 0$ also. Since according to the corresponding dispersion law different resonance frequencies correspond to different values of **K**, the Fourier-transforming of the free induction decay should give a set of resonance peaks referring to different modes of the excited spin waves. Moreover, the penetration of the spin-like branch of coupled waves into the sample in the preparation stage depends on which part of the exciting field is absorbed by the spins. The indicated spatial nonuniformities of the initial conditions for free induction decay thus depend on the detuning of the exciting field, which causes the spectrum obtained to depend on the detuning.

The experiments which we are examining, on obtaining NMR spectra by Fourier-transforming of the free induction decay, were performed on metals with two isotopes possessing magnetic moments, (^{63}Cu and ^{65}Cu in copper and ^{203}Tl and ^{205}Tl in thallium). The study of these systems gives rise to very complicated equations.

To determine the main features of the predicted phenomena we would therefore like to study a system of equivalent exchange-coupled nuclear magnetic moments and to calculate the Fourier-transforming of the free induction decay, taking into consideration the nonzero conductivity of the sample in the light of the above arguments.

2. GENERAL EXPRESSIONS FOR THE FREE INDUCTION DECAY SIGNAL IN METALS

We shall consider the behavior of the magnetization **M** of a sample in an effective field

$$\mathbf{B}_{eff}=\mathbf{B}_0+\mathbf{B}+\alpha\Delta\mathbf{M}, \tag{1}$$

where $\mathbf{B}_0$ is the constant magnetic field, $\alpha\Delta\mathbf{M}$ is the field due to nonuniform exchange whose microscopic origin is the Ruderman–Kittel interaction

$$-\sum_{i,j} J_{i,j}\mathbf{I}_i\cdot\mathbf{I}_j$$

between nuclear-spin operators (the indices i and j denote lattice sites) or the same thing in the form of the macroscopic exchange energy

$$-J_0\int d\mathbf{r}\int d\mathbf{r}'\mathbf{M}(\mathbf{r},t)\cdot\mathbf{M}(\mathbf{r}',t)$$

(below, following the arguments presented on p. 22 of Ref. 4, we arrive at a local field $\alpha\Delta\mathbf{M}$; the constants J_0 and α will be given below); $\mathbf{B}=\mu_0(\mathbf{H}+\mathbf{M})$ is the residual part of the ac magnetic induction ($\mu_0=4\pi\times10^{-7}$ H/m, **H** is the intensity of the magnetic field). We shall consider a sample in the form of a plate of thickness L, bounded by the planes $z=0$ and $z=L$ and arranged perpendicular to the z axis. (The axes are chosen so that the constant magnetic field is parallel to the z axis, while the time-varying one during excitation is directed along the surface of the sample in the direction of the x axis, which is parallel to the long side of the plate). Then all variables will depend only on the coordinate z.

It should be noted that in the experimental situation of Refs. 1–3, which we shall have in mind in what follows, the sample is a normal metal with no localized electronic spins. The constant magnetic field (0.4 T for NMR of thalium at a frequency of about 10 MHz) is weak in the sense of the curvature of the conduction-electron trajectories: The conduction-electron mean-free path, $l_e\approx10^{-5}$ m, in the region of residual resistance is smaller than the Larmor radius $R_L\approx6\times10^{-5}$ m, so that in such a field the conditions for the appearance of weakly damped helicoidal waves are not satisfied. A constant magnetic field could possibly alter the conductivity of the sample by a very small amount. But since, unfortunately, the conductivity is unknown, taking account of this small change is not feasible. Therefore the effects due to the curvature of the electron trajectories by a constant magnetic field in the present problem are unimportant and will be neglected.

We shall treat $M^+=M^x+iM^y$ and $H^x=H^x+iH^y$ as dynamical variables. The equation for M^+ can be obtained from

$$\frac{d\mathbf{M}}{dt}=\gamma\mathbf{M}\times\mathbf{B}_{eff}, \tag{2}$$

where γ is the gyromagnetic ratio of the spins. The equation for H^+ follows from Maxwell's equations

$$\nabla\times\mathbf{E}=-\frac{\partial\mathbf{B}}{\partial t},\quad \nabla\times\mathbf{H}=\mathbf{j}. \tag{3}$$

To clarify the dependence $\mathbf{j}(\mathbf{E})$ it is necessary to start from the relation between the electron mean-free path length and the conventional skin-depth. In the region of residual resistance, NMR frequency of about 10 MHz, and thallium conductivity $\sigma=6\times10^8\Omega^{-1}\cdot\mathrm{m}^{-1}$, we find $l_e\approx\delta\approx10^{-5}$ m. Thus the condition for the normal skin effect, $l_e\ll\delta$, is not satisfied.

However, even the extremely anomalous skin effect, $l_e\gg\delta$, likewise does not occur. To demonstrate the possibilities that taking account of the conductivity of the sample opens up in the standard picture of the free induction decay of nuclear spins we shall assume that skin-effect is a normal type, i.e., we shall write Ohm's law $\mathbf{j}=\sigma\mathbf{E}$ for the current density, where σ is the conductivity of the sample and is assumed to be independent of the wave vector. We note that the anomalous skin effect would lead to more uniform penetration of an ac magnetic field into the sample, which would strongly decrease the influence of the nonuniformity of the penetration on the free induction decay, and the NMR spectra obtained by Fourier transforming of the free induction decay would be identical to the spectra in dielectrics. Since the NMR spectra are unusual, it remains to assume that there exists a pronounced surface, i.e., the normal skin effect occurs (compare with the Kittel case[5]).

We shall confine our attention in the present paper to weak excitation of transverse magnetization by an external rf field (small tipping angle). We therefore take $M_z=M_0$,

where M_0 is the equilibrium nuclear magnetization. Eliminating **E** from Eq. (3), we find for H^+ an equation which is related to the equation for M^+ from Eq. (2):

$$\frac{\partial M^+}{\partial t} = -i\left(\omega_0 M^+ - \gamma\alpha M_0 \frac{\partial^2 M^+}{\partial z^2} \right) + i\gamma\mu_0 M_0 H^+ - \frac{M^+}{T_2}, \tag{4}$$

$$\frac{\partial H^+}{\partial t} = \frac{1}{\sigma\mu_0} \frac{\partial^2 H^+}{\partial z^2} - \frac{\partial M^+}{\partial t}, \tag{5}$$

where $\omega_0 = \gamma B_0$ is the Larmor frequency of the spins and T_2 is the spin–spin relaxation time introduced phenomenologically following Bloch.

We note first that all interesting effects described below arise because of the interaction of the magnetization with the field (the term proportional to $\gamma\mu_0 M_0$ in Eq. (4)). Therefore M_0 must be large enough to make these effects observable. In the experiments described in Refs. 1–3 this is achieved by using ultralow temperatures (ranging from values of the order of tenths of a millikelvin up to 10 mK), which are, however, much greater than the nuclear magnetic ordering temperature.

First we shall study the effect on the sample of an external ac field with intensity

$$H^x(z=0,t) = 2h_0 \cos \Omega t. \tag{6}$$

at the sample surface. Since the pulse length in free induction decay experiments is much less than T_2, we shall neglect the last term in Eq. (4) during the time-action of the pulse acts. We seek the solution of the system (4) and (5) in the form of two traveling waves: neglecting their ''entanglement,'' an ac magnetic field would penetrate in the form of one of them, while a transverse magnetization, produced by the penetration of the field, would penetrate in the form of the other (the frequencies of these waves equal Ω because the oscillations of the field and magnetization are forced):

$$\begin{aligned} H^+_{\text{pulse}} &= h^+ \exp(-i\Omega t + ikz), \\ M^+_{\text{pulse}} &= m^+ \exp(-i\Omega t + ikz). \end{aligned} \tag{7}$$

Substituting the expressions (7) into Eqs. (4) and (5) we find

$$(\omega_0 - \Omega + \gamma\alpha M_0 k^2) m^+ - \gamma\mu_0 M_0 h^+ = 0, \tag{8}$$

$$i\Omega m^+ + \left(i\Omega - \frac{k^2}{\sigma\mu_0} \right) h^+ = 0. \tag{9}$$

Note that these equations are suitable for the case where the spins approach ferromagnetic ordering ($J_0 > 0$). If J_0 is negative, two magnetic sublattices must be considered.[4] This problem falls outside the scope of the present paper.

The dispersion equation obtained from the requirement that the system (8) and (9) possess a nontrivial solution is

$$k^4 - k^2 \left(\frac{\Omega - \omega_0}{\gamma\alpha M_0} + i\frac{2}{\delta^2} \right) + i\frac{\sqrt{2}}{\delta} \frac{\Omega - \omega_0 - \gamma\mu_0 M_0}{\gamma\alpha M_0} = 0, \tag{10}$$

where $\delta = \sqrt{2/\sigma\mu_0\Omega}$ is the standard skin-depth in the metal. The solutions of Eq. (10) are

$$k_1^2 = k_0^2 \left(1 + i\frac{2\eta}{\delta^2 k_0^2} \right), \quad k_2^2 = i\frac{2}{\delta^2}(1-\eta), \tag{11}$$

where

$$k_0^2 = \frac{\Omega - \omega_0}{\gamma J_0 M_0 l^2}, \tag{12}$$

$$\eta = \frac{\zeta}{k_0^2\delta^2/2 - i\left(1 + \sqrt{1 + \dfrac{i2\zeta}{(k_0^2\delta^2/2 - i)^2}} \right)^{-1}},$$

$$\zeta = \frac{\mu_0 \delta^2}{J_0 l^2}, \quad J_0 = \frac{z_n J}{n_s \gamma^2 \hbar^2}. \tag{13}$$

Here J_0 is the most important quantity characterizing the interaction of the nuclei (the Ruderman–Kittel interaction of a given nucleus only with z_n nearest neighbors is taken into account, J is the constant characterizing this interaction, and n_s is the spin density) and l is an effective length in α. It should be noted that an estimate of α in ferromagnetic dielectrics yields[4] $\alpha = J_0 a^2/2$, where a is the lattice constant. However, the Ruderman–Kittel interaction length in principle can be assumed to extend over the electron mean-free path length l_e, so that there are grounds for taking l_e instead of a in the estimate of α. But, since this assumption has not been confirmed experimentally, l must be treated (just like α) as an adjustable parameter. For convenience we shall write α in the form $\alpha = J_0 l^2$. It follows from Eq. (8) that the partial characteristic frequency of the spin-like branch is $\omega_0 + \gamma\alpha M_0 k^2$. Since we have $J_0 > 0$, only a rf field with frequency $\Omega \geqslant \omega_0$ can excite precession of the magnetization, and therefore k_0^2 is positive.

Thus, the rf field and the transverse magnetization produced by it propagate into the plate in the form of two waves, one of which (with wave number k_2) is due to the skin effect in metals while the other (with wave number k_1) is due to the nonuniform exchange between nuclear spins. Since the relation

$$m^+ = -\left(1 + i\frac{\delta^2 k^2}{2} \right) h^+ \tag{14}$$

between the amplitudes of the two waves follows from Eq. (9), coupled traveling waves of the field and the magnetization can be written as a superposition of the two branches in the following form:

$$\begin{aligned} H^+_{\text{pulse}}(z,t) &= h_1^+ \exp(-k_1'' z)\exp(ik_1' z - i\Omega t) \\ &\quad + h_2^+ \exp(-k_2'' z)\exp(ik_2' z - i\Omega t), \end{aligned} \tag{15}$$

$$\begin{aligned} M^+_{\text{pulse}}(z,t) &= -\left(1 + i\frac{\delta^2 k_1^2}{2} \right) h_1^+ \exp(-k_1'' z) \\ &\quad \times \exp(ik_1' z - i\Omega t) - \left(1 + i\frac{\delta^2 k_2^2}{2} \right) h_2^+ \\ &\quad \times \exp(-k_2'' z)\exp(ik_2' z - i\Omega t). \end{aligned} \tag{16}$$

Waves of this form are solutions of the coupled equations for H^+ and the electronic magnetization in metals taking ac-

count of electron diffusion.[6] Equations (15) and (16) take into account the inequality $L \gg \delta$ in the experimental situations mentioned above. In this case neither the shape nor the dimensions of the sample are important, and we have penetration of H^+ and M^+ through the boundary into a metal filling a half-space.[4] Then, only the solutions $k_{1,2}$ satisfying $\mathrm{Im}\, k_{1,2} > 0$, i.e., the solutions which give the spatial damping in Eqs. (15) and (16), should be chosen.

The amplitudes $h_{1,2}^+$ can be found from the boundary conditions, of which one is the relation (6), while the other is the condition that there be no magnetization flux through the sample surface:[6]

$$\frac{\partial M^+}{\partial z}(z=0, t=0)=0. \tag{17}$$

The expressions found for the amplitudes are

$$h_1^+ = 2h_0 \frac{\varphi_2}{\varphi_2-\varphi_1}, \quad h_2^+ = -2h_0 \frac{\varphi_1}{\varphi_2-\varphi_1}, \tag{18}$$

where

$$\varphi_{1,2} = k_{1,2}\left(1 + i\frac{\delta^2 k_{1,2}^2}{2}\right). \tag{19}$$

We shall now examine the free precession of the magnetization after the rf pulse. We start from Eqs. (4) and (5) taking account of T_2^{-1}. In contrast to the preparatory stage (excitation by an external rf field), now H^+ and M^+ are not known in advance as functions of time and are to be determined. On the other hand, the spatial dependence of the variables considered is obvious: The nonuniform precession of the transverse magnetization occurring inside the sample produces a nonuniform magnetic field. Combining the corresponding waves, moving from one surface of the sample to the other and back, with anti-node boundary conditions

$$\frac{\partial M^+}{\partial z}(z=0, L)=0,$$

form cosine standing waves with wave numbers $K=\pi n/L$, $n=0, 1, 2, \ldots$.

Thus we shall seek the solution of Eqs. (4) and (5) during free precession in the form

$$\begin{aligned} M_{pr}^+(z,t) &= \sum_{n=0}^{\infty} m^+(t)\cos\frac{\pi n z}{L}, \\ H_{pr}^+(z,t) &= \sum_{n=0}^{\infty} h^+(t)\cos\frac{\pi n z}{L}. \end{aligned} \tag{20}$$

Substituting the expressions (20) into Eqs. (4) and (5) we find the following expression for $m^+(t)$:

$$\frac{d^2 m^+}{dt^2} + i\left(\omega_0 + \gamma\alpha M_0 K^2 + \gamma\mu_0 M_0 - \frac{i}{T_2} - i\frac{K^2}{\sigma\mu_0}\right)\frac{dm^+}{dt} + \frac{iK^2}{\sigma\mu_0}\left(\omega_0 + \gamma\alpha M_0 K^2 - \frac{i}{T_2}\right)m^+ = 0. \tag{21}$$

Therefore $m^+(t) = Ae^{i\lambda t}$. Since the coupling between the magnetization and the field is assumed to be small ($\gamma\mu_0 M_0/\omega_0 \ll 1$), we find the following approximate expressions for the characteristic roots:

$$\lambda_1 \approx -(\omega_0 + \gamma\alpha M_0 K^2 + \gamma\mu_0 M_0) + i\left(\frac{1}{T_2} + \frac{K^2}{\sigma\mu_0}\frac{\gamma\mu_0 M_0}{\omega_0 + \gamma\alpha M_0 K^2}\right), \tag{22}$$

$$\lambda_2 \approx \frac{iK^2}{\sigma\mu_0}\left(1 - \frac{\gamma\mu_0 M_0}{\omega_0 + \gamma\alpha M_0 K^2}\right). \tag{23}$$

On the other hand, the time-dependent parts of the expressions (20) are related as

$$h^+ = \frac{\omega_0 + \gamma\alpha M_0 K^2 - i/T_2 + \lambda}{\gamma\mu_0 M_0} m^+ \equiv \psi_t m^+. \tag{24}$$

Thus, the time-dependent parts of the expressions (20) can be written in the form of oscillations of two coupled linear oscillators. Therefore M^+ and H^+ during free precession can be written in the form

$$\begin{aligned} M_{pr}^+(z,t) &= \sum_{n=0}^{\infty} [A_{1n}\exp(i\lambda_1 t) + A_{2n}\exp(i\lambda_2 t)]\cos\frac{\pi n z}{L}, \\ H_{pr}^+(z,t) &= \sum_{n=0}^{\infty} [\psi_{1t}A_{1n}\exp(i\lambda_1 t) + \psi_{2t}A_{2n}\exp(i\lambda_2 t)]\cos\frac{\pi n z}{L}, \end{aligned} \tag{25}$$

where $\psi_{1,2t}$ denote ψ_t from Eq. (24) with the values of λ_1 and λ_2, respectively, substituted. Note that, since λ_2 is imaginary, there is actually only one oscillator (the spin system) possessing a set of characteristic frequencies

$$\omega_{\mathrm{res}}(n) = \omega_0 + \gamma\alpha M_0(\pi n/L)^2 + \gamma\mu_0 M_0 \tag{26}$$

and decay rates

$$\frac{1}{T_2'(n)} = \frac{1}{T_2} + \frac{1}{\sigma\mu_0}\left(\frac{\pi n}{L}\right)^2 \frac{\gamma\mu_0 M_0}{\omega_0 + \gamma\alpha M_0(\pi n/L)^2}, \tag{27}$$

which are, respectively, the usual spin-wave spectrum of a ferromagnet[4] $\omega_0 + \gamma\alpha M_0(\pi n/L)^2$ and the usual spin–spin decay with corrections due to the coupling of the magnetization with the magnetic field. The other ''oscillator'' (the field produced by the motion of the spins) decays monotonically with the decay constant λ_2 as a result of the resistance of the metal.

The amplitudes of free precession can be found, as usual, by matching the values of the components of the field and the magnetization at the initial moment of free precession ($t=0$ in Eq. (25)) with the same components at the end of the evolution process under the action of the exciting pulse. Recall that under the action of a pulse the boundary condition propagation into the sample took place.[7] We will study this wave-like propagating coordinate system rotating

with frequency Ω around the z axis with the spatial phase varying with distance z as $\exp(ik_1'z)$ for the spin-like wave and as $\exp(ik_2'z)$ for the field-like wave (mathematically this means that in Eqs. (15) and (16) we must set $k_1'z-\Omega t=0$ and $k_2'z-\Omega t=0$).

In such a coordinate system (we will call it the wave coordinate system), under the action of the exciting rf pulse, an observer will see static spatial distributions of the magnetic field and magnetization similar to the static distribution of the deviations from equilibrium of two types of hypothetical pendulums at various angles inside the sample. It is assumed that such a pattern is produced rapidly compared with the length of the rf pulse and then persists while the external field acts.

The matching conditions can therefore be written

$$M^+, H_{pr}^+(t=0)=M^+, H_{\text{pulse}}^+(z). \tag{28}$$

We take the left-hand side of expression (28) from Eq. (25), and we expand the right-hand side, which is given by expressions (15) and (16) with $k_{1,2}'z-\Omega t=0$, in a Fourier cosine series. A simple system of equations for $A_{1,2n}$ is found by comparing the expressions obtained. According to the above mentioned, the terms of Eq. (25) corresponding to the characteristic root λ_2 cannot be expected to exhibit resonant behavior, so we present only the computational result for $A_{1n}^{x,y}$ $\equiv \mathrm{Re(Im)}A_{1n}$:

$$A_{1n}^{x,y}=\psi_{1z}^{x,y}h_1^{x,y}f_1+\psi_{2z}^{x,y}h_2^{x,y}f_2, \tag{29}$$

where $h_1^{x,y}$ can be calculated from Eq. (18) and

$$\psi_{1,2z}^{x,y}\approx\pm\left(1-\frac{\delta^2(k_{1,2}^2)''}{2}\right)-\frac{\delta^2(k_{1,2}^2)'}{2}\frac{h_{1,2}^{y,x}}{h_{1,2}^{x,y}}, \tag{30}$$

$$f_{1,2}=\frac{2}{L}\frac{k_{1,2}''}{(k_{1,2}'')^2+(\pi n/L)^2}[1-\exp(-k_{1,2}''L)\cos\pi n]. \tag{31}$$

Finally, we have for the part of the x component of the magnetization that is resonant during the free induction decay

$$M_{pr}^x(\mathrm{res},n)=\exp\left(-\frac{t}{T_2'(n)}\right)[A_{1n}^x\cos(\omega_{\mathrm{res}}(n)t) - A_{1n}^y\sin(\omega_{\mathrm{res}}(n)t)]\cos\frac{\pi nz}{L}. \tag{32}$$

The expression in front of $\cos(\pi nz/L)$ in Eq. (32) is similar to the stationary nonuniform NMR signal in ferromagnets obtained by Kittel.[5] It yields the contribution of the nth mode of standing spin waves to the free induction decay. As usual,[8] the part of expression (32) that precesses in phase with the linearly polarized applied field yields a dispersion signal in the response of the spin system to a one-pulse excitation $A_{1n}^x/2h_0$, while the part shifted in phase by $\pi/2$ relative to the exciting field yields the absorption signal $-A_{1n}^y/2h_0$. Fourier transforming expression (32) gives the corresponding Lorentzian line of the nth mode of the spin waves with the resonance frequency and decay rate given by Eqs. (26) and (27).

It is interesting to note that the distortion of the standing spin waves that is due to the spatial nonuniformity inherited from the preparatory stage made the quantity

$$f_{1,2}\equiv\frac{2}{L}\int_0^L\exp(-k_{1,2}''z)\cos\frac{\pi nz}{L}dz$$

(and consequently $A_{1n}^{x,y}$) different from zero: With no spatial damping due to the spin effect, i.e., with $k_{1,2}''=0$, we would have $f_{1,2}=0$.

In summary, we have shown that the NMR spectrum found by Fourier-transforming of the free induction decay in metals at sufficiently low temperatures consists of a set of contributions of standing spin waves. If the Fourier transform is performed so that the free induction decay at the start of the Fourier-transform process passes through a maximum, then the Fourier cosine transform (F_c) gives an absorption signal, while the sine transform (F_s) gives a dispersion signal. In practice, the resulting Fourier transform $F_r=\sqrt{F_c^2+F_s^2}$ is usually calculated:

$$F_r(\omega,\lambda,n)=\frac{1}{\sqrt{2\pi}}\frac{\sqrt{[(A_{1n}^x)^2+(A_{1n}^y)^2][(\lambda_1'')^2+(\omega+\lambda_1')^2]}}{(\lambda_1'')^2+(\omega+\lambda_1')^2}, \tag{33}$$

where ω is the traveling frequency of the Fourier transform. Expression (33) together with Eqs. (29)–(31) describes the experimental results on recording NMR in metals at ultralow temperatures by Fourier-transforming of the free induction decay.

3. RESULTS

We analyze the results obtained. First, a collection of peaks should be expected near the frequency $\omega_0(\omega\approx\omega_0)$ in the Fourier-transforming of the free induction decay. According to Eq. (26) these peaks are due to spin waves. Consequently, in experiments studying free induction decay in metals, in contrast to dielectrics, when the rf field outside the metal is spatially uniform, standing spin waves can be detected because of the skin-effect-induced nonuniform distribution of the external field and the transverse magnetization produced by it inside the metal during the preparation (excitation) stage.

We shall now examine the dependence of the output of the free induction decay signal on the detuning of the external rf field from the Larmor frequency. First we shall discuss the results for exact tuning ($\Omega=\omega_0$). Then the quantities (11) are

$$k_1^2=\frac{2i}{\delta^2}\eta, \quad k_2^2=\frac{2i}{\delta^2}(1-\eta),$$

where

$$\eta=\frac{-i\zeta}{1+\sqrt{1+2i\zeta}}, \quad \zeta=\frac{\mu_0\delta^2}{J_0l^2},$$

and do not depend on M_0. Moreover, all quantities in the final results are independent of M_0 with the exception of those $\sim\gamma\mu_0M_0/\omega_0$, which can be neglected. Therefore, in metals at the exciting rf pulse tuned precisely to the Larmor

frequency there is no magnetic resonance signal that could be detected as a change in the complex impedance (active resistance and inductance) of the coil of the NMR spectrometer by placing the sample inside the coil compared with an empty coil. This behavior is explained by the fact that the rf field simply does not penetrate into the sample because of strong absorption by the spins: the sample is opaque for this field.

We now proceed to the case of large detunings, $\Omega-\omega_0 \gg \gamma J_0 M_0 l^2/\delta^2$, or in other words, $k_0\delta \gg 1$. (However, there must be a limit on the detuning, given by the condition for a continuous medium:

$$\gamma J_0 M_0 l^2/(\Omega-\omega_0) \gg a^2.$$

Note that this condition is easily satisfied even for ''large'' detunings.) Then, during the excitation stage the field-like branch of the coupled (field + magnetization) system propagates and decays, to a good approximation, over a distance on the order of the skin-depth $\delta(k_2'=k_2''\approx\delta^{-1})$. In contradistinction to this, the wavelength of the spin-like branch is much less than $\delta(k_1'\approx k_0)$, while the penetration depth of this wave into the sample is much greater than δ:

$$k_1''\approx \delta^{-1}\zeta/(k_0\delta)^3 \ll \delta^{-1}. \tag{34}$$

Let

$$\delta_{\text{spin}} \gg L \gg \delta \tag{35}$$

($\delta_{\text{spin}}=(\delta/\zeta)(k_0\delta)^3$ is the depth of the ''spin'' skin layer). The first inequality means that the static pattern of the distribution of the magnetization deviations, which was discussed prior to expression (28), is nonuniform over the entire thickness of the plate, while the second inequality presupposes (as is the case in an experiment) that the skin-depth is much less than this thickness.

Calculations lead to the conclusion that since the inequality

$$\frac{L}{\sigma}\frac{\zeta}{(k_0\delta)^3} \ll 1 \tag{36}$$

follows from Eq. (35), the intensity of the even modes of the standing spin waves which exist in a plate at free induction decay (i.e., with even n: an even number of half-waves of the standing spin waves fit within the thickness of the plate) is much higher than the intensity of the odd modes.

The corresponding NMR signals in a phase-sensitive Fourier-transforming of the free induction decay is described by the expressions

$$\chi'(\omega,\Omega,2n)\approx \frac{1}{\sqrt{2\pi}}\frac{2\gamma\mu_0 M_0}{\Omega-\omega_0}\left\{\sqrt{\frac{\gamma J_0 M_0 l^2}{(\Omega-\omega_0)\delta^2}}\times\left[1+\left(\frac{2\pi n}{L}\frac{\delta(k_0\delta)^3}{\zeta}\right)^2\right]^{-1}+\frac{\delta^{-1}L^{-1}}{\delta^{-2}+(2\pi n/L)^2}\right\}\times\frac{[\omega-\omega_{\text{res}}(2n)]}{[T_2'(2n)]^{-2}+[\omega-\omega_{\text{res}}(2n)]^2}, \tag{37}$$

$$\chi''(\omega,\Omega,2n)\approx \frac{1}{\sqrt{2\pi}}\frac{2\gamma\mu_0 M_0}{\Omega-\omega_0}\sqrt{\frac{\gamma J_0 M_0 l^2}{(\Omega-\omega_0)\delta^2}}\times\left[1+\left(\frac{2\pi n}{L}\frac{\delta(k_0\delta)^3}{\zeta}\right)^2\right]^{-1}\times\frac{[T_2'(2n)]^{-1}}{[T_2'(2n)]^{-2}+[\omega-\omega_{\text{res}}(2n)]^2}. \tag{38}$$

The expressions (37) and (38) depend on the equilibrium magnetization. They therefore indeed represent the NMR signals detected by the rf coil. Thus, the second basic result of this work is that the signals corresponding to standing spin waves can be detected only if the rf field exciting free induction decay is applied at a frequency sufficiently removed from ω_0. Then the spin-like branch of the coupled waves is excited over the entire thickness of the sample and makes NMR observable. Specifically, in a phase-sensitive Fourier transform, Lorentzian-shaped peaks of width $T_2'(2n)^{-1}$ should appear in the NMR absorption signal at the frequencies $\omega_{\text{res}}(2n)$. Such peaks are observable if the frequency spacing between the $(n+1)$st and neighboring $(n-1)$st peaks ($n=1, 3, \ldots$) exceeds their width, i.e.,

$$\gamma J_0 M_0 (4\pi n l/L)^2 > T_2^{-1}.$$

Substituting the data for ^{205}Tl (relative content 70%) with gyromagnetic ratio $\gamma=2.5\times10^7$ Hz/T, equivalent to the magnetization $M_0=202p$ A/m, where p is the polarization of the nuclear spins, using the experimental result of Ref. 9 $|J_0|=28\mu_0$ h/m, and substituting $T_2^{-1}\approx 2\pi\cdot 4\times10^4$ s^{-1}, measured in Ref. 3 with $p\simeq 0.5 (T\simeq 1$ mK), we find that the $(n+1)$st and $(n-1)$st peaks of the spin waves can be resolved if n satisfies the inequality $(nl/L)^2>10^{-2}$, which holds for $n>1$, if the sample thickness $l=10^{-4}$ m and l is interpreted as the electron mean-free path length in the residual-resistance region, $l_e\approx 10^{-5}$ m. In Ref. 3 the sample was cylindrical with diameter 3×10^{-3} m, so that it is not surprising that there was no ''comb'' of spin-wave peaks, but kinks and hints of structure were observed in the spectra obtained with the exciting rf field applied at a frequency much higher than the Larmor frequency of the isotopes. Unfortunately, it is impossible to make a direct comparison with the observed spectra, since our calculations were performed for nuclear spins of one kind, while the experiments of Refs. 1–3 were performed on samples with the two thallium isotopes in the natural abundance, strongly coupled with one another by the Ruderman–Kittel interaction.

The third basic result is the nonlinear dependence of the NMR signals on the equilibrium magnetization, or in other words, on the reciprocal of the sample temperature: The absorption signal and part of the dispersion signal (37) and (38) are proportional to $M_0^{3/2}$. The reason for this behavior is that, as one can see from Eqs. (34) and (35), for large detunings only the spin-like branch of the coupled waves penetrates into the entire thickness of the sample, and the amplitude of

this branch in the expression for the field, calculated according to Eqs. (18), is a nonlinear function of M_0. The transverse magnetization with free induction decay inherits this nonlinearity. The corresponding dependence of the intensity of the NMR lines on the reciprocal of the sample temperature has been observed experimentally.[3]

We note that analytical expressions derived for the absorption and dispersion signals in the present paper for ''large'' detunings are applicable for ^{205}Tl described above for $\Omega-\omega_0\approx 10^6$ s^{-1}: For $B_0=0.4$ T ($\omega_0\approx 2\pi\times 10^7$ s-1) and residual resistivity $\rho^{-1}\approx 6\times 10^8\Omega^{-1}\cdot$m^{-1} we have $\delta\approx 10^{-5}$ m, $k_0\delta\approx 10^{1/2}$, and $\delta_{\rm spin}\approx 10^{-3/2}$ m. In this case the inequalities (35), expressing the optimal conditions for the appearance of standing spin waves accompanying at free induction decay are satisfied for sufficiently thin samples ($L\approx 10^{-4}$ m), but the nonlinear dependence of the NMR signals on the reciprocal of the temperature undoubtedly also exists under conditions which are less favorable for manifestations of spin waves.

In summary, in the present work we have shown that the skin effect in metals at ultralow temperatures qualitatively changes the NMR picture observed as the Fourier-transforming of the free induction decay, even for small tipping angles of the magnetization [linear coupling of the (field + magnetization) system]. The results obtained are in qualitative agreement with experiments.

This work was supported in part by the program Deutsche Forschungsgemeinschaft.

*)E-mail: faculty@tsu.ge
†)E-mail: faculty@tsu.ge
‡)E-mail: girgl@btp9xa.phy.uni-bayreuth.de

[1] G. Eska in *Proceedings of the Conference on Quantum Fluids and Solids*, edited by G. Ihas and Y. Takano, Gainesville, Florida, USA, 1989, AIP Conf. Proc. No. 194 (1989), p. 316.
[2] W. Baeuml, G. Eska, and W. Pesch, Physica B **194–196**, 321 (1994).
[3] J. Leib, M. Huebner, S. Goetz, Th. Wagner, and G. Eska, J. Low Temp. Phys. **101**, 253 (1995).
[4] A. I. Akhiezer, V. G. Bar'yakhtar, and S. V. Peletminskiĭ, *Spin Waves*, North-Holland Pub. Co., Amsterdam, 1968 [Russian original, Nauka, Moscow, 1967].
[5] C. Kittel, Phys. Rev. **110**, 1295 (1958).
[6] J. Winter, *Magnetic Resonance in Metals*, Clarendon Press, Oxford, 1971.
[7] A. I. Tikhonov and A. A. Samarskiĭ, *Equations of Mathematical Physics*, Pergamon Press, New York, 1964 [Russian original, Nauka, Moscow, 4th edition, 1972].
[8] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford, 1961 [Russian translation, Inostr. Lit., Moscow, 1963].
[9] Yu. S. Karimov and I. F. Shchegolev, Zh. Éksp. Teor. Fiz. **41**, 1082 (1961) [Sov. Phys. JETP **14**, 772 (1962)].

Translated by M. E. Alferieff