Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 433–440.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 483–490.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Exotic Baryon States in Topological Soliton Models¶

H. Walliser[a,*] and V. B. Kopeliovich[b,]**

[a]*Fachbereich Physik, Universität Siegen D57068, Siegen, Germany*
[b]*Institute for Nuclear Research, Russian Academy of Sciences, pr. Shestidesyatiletiya Oktyabrya 7a, Moscow, 117312 Russia*
**e-mail: walliser@physik.uni-siegen.de*
***e-mail: kopelio@al20.inr.troitsk.ru, kopelio@cpc.inr.ac.ru*

Received April 7, 2003

Abstract—The novel observation of an exotic strangeness $S = +1$ baryon state at 1.54 GeV is to trigger an intensified search for this and other baryons with exotic quantum numbers. This state was predicted long ago in topological soliton models. We use this approach together with the new datum in order to investigate its implications for the baryon spectrum. In particular, we estimate the positions of other pentaquark and septuquark states with exotic and with nonexotic quantum numbers. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In a recent paper [1], Nakano *et al.* report on an exotic strangeness $S = +1$ baryon state observed as a sharp resonance at 1.54 ± 0.01 GeV in photoproduction from neutrons. The confirmation of this finding would give formidable support to topological soliton models [2, 3] for the description of baryons in the nonperturbative regime of QCD. Higher multiplets containing states carrying exotic quantum numbers arise naturally in the $SU(3)$ version of these models. These were called exotic because such states cannot be built from only three valence quarks within quark models; additional quark-antiquark pairs must be added. The terms pentaquark and septuquark therefore characterize the quark content of these states. Strictly, there is nothing exotic about these states in soliton models; they simply come as members of the next higher multiplets.

Indeed, beyond the minimal {8} and {10} baryons, a $\{\overline{10}\}$ baryon multiplet was also mentioned earlier by Chemtob [4]. Within a simple $SU(3)$ symmetric Skyrme model, Biedenharn and Dothan [5] estimated the excitation energy of the $\{\overline{10}\}$ multiplet having a spin of $J = 1/2$ to be only 0.60 GeV [sic] above the nucleon. Both this multiplet and the {27} multiplet with a spin of $J = 3/2$ contain low-lying $S = +1$ states, called Z and Z^* in what follows. The first numbers for these exotic states, taking into account the configuration mixing caused by symmetry breaking, were given in [6], although they are around 0.1 GeV too high if the value found in [1] proves to be correct. Diakonov, Petrov, and Polyakov [7] postulated the experimental $P11(1.71)$ nucleon resonance of a member of the $\{\overline{10}\}$ multiplet, and hence the Z state again with a low excitation energy (0.59 GeV). Weigel [8] showed that similar low numbers (0.63 GeV) can be obtained in an extended Skyrme model calculation that includes a scalar field.

It should be added that the excitation energies of similar exotic states have been estimated for arbitrary baryonic numbers [9]. All these states appear to be above the threshold for the decay due to strong interactions. In general, the excitation energies for the $B > 1$ systems are comparable to those for baryons; e.g., the $S = 1$ dibaryon state belonging to the $\{\overline{35}\}$ multiplet was calculated to be only 0.59 GeV above the NN threshold [10].

In this paper, we address the following questions concerning the $B = 1$ sector. Is the exotic Z state at 1.54 GeV reported in [1] compatible with soliton models and the known baryon spectrum? If Z is actually located at this position, what does it imply for the other exotic states?

2. $SU(3)$ SOLITON MODEL

A large number of different soliton models exist: pure pseudoscalar ones, models with scalar fields and/or vector and axial-vector mesons, and even models that include quark degrees of freedom. There is also a vast number of possible terms in the effective action for each of these models, partly with free adjustable parameters. However, the $SU(3)$-symmetric part always leads to the same collective Hamiltonian with only two model-dependent quantities determining the baryon spectrum (Section 2.1). The situation with the symmetry breaking part is less advantageous, unfortunately, but still there appears one predominant standard symmetry breaker, which is to be the third model-dependent quantity needed (Section 2.2). Thus, instead of

¶This article was submitted by the authors in English.

referring to a specific model (which also involves a number of free parameters), we adjust these three quantities to the known {8} and {10} baryon spectra and to the recently reported Z state [1]. Using this input, we try to answer the questions posed in the Introduction. We also show that the values needed for the three quantities are not too far from what can be obtained in the standard Skyrme model.

In the baryon sector, the static hedgehog soliton configuration located in the nonstrange $SU(2)$ subgroup is collectively and rigidly rotated in the $SU(3)$ space. There are other approaches, such as the soft rotator approach and the bound state approach, but the rigid rotator approach is probably most appropriate for $B = 1$.

2.1. SU(3)-Symmetric Part

The $SU(3)$-symmetric effective action leads to the collective Lagrangian [11]

$$L^S = -M + \frac{1}{2}\Theta_\pi \sum_{a=1}^{3} (\Omega_a^R)^2 + \frac{1}{2}\Theta_K \sum_{a=4}^{7} (\Omega_a^R)^2 - \frac{N_C B}{2\sqrt{3}}\Omega_8^R \tag{1}$$

depending on the angular velocities Ω_a^R, $a = 1,\ldots, 8$. It is generic for all effective actions whose nonanomalous part contains at most two time derivatives, the term linear in the angular velocity depends on the baryon number B and the number of colors N_C, and it appears due to the Wess–Zumino–Witten anomaly.

The soliton mass M and the pionic and kaonic moments of inertia Θ_π and Θ_K are model-dependent quantities. The latter two are relevant to the baryon spectrum. The soliton mass M, subject to large quantum corrections, enters the absolute masses only. With the right and left angular momenta

$$R_a = -\frac{\partial L^0}{\partial \Omega_a^R}, \quad L_a = \sum_{b=1}^{8} D_{ab} R_b, \tag{2}$$

which transform according to Wigner functions D_{ab} depending on the soliton orientation, the Hamiltonian obtained by the Legendre transformation

$$H^S = M + \frac{1}{2\Theta_\pi}\mathbf{R}^2 + \frac{1}{2\Theta_K}\left(C_2(SU(3)) - \mathbf{R}^2 - \frac{N_C^2 B^2}{12}\right) \tag{3}$$

can be expressed by the second-order Casimir operators of the $SU(3)$ group and its nonstrange $SU(2)$ subgroup,

$$C_2(SU(3)) = \sum_{a=1}^{8} R_a^2, \quad \mathbf{R}^2 = \sum_{a=1}^{3} R_a^2. \tag{4}$$

For a given $SU(3)$ irrep (p, q) with the dimension

$$N = (p+1)(q+1)(p+q+2)/2,$$

the eigenvalues of these operators are given by

$$\begin{aligned} &C_2(SU(3))|\{N\}(p,q),(Y_R J J_3)\rangle \\ &= \left[\frac{p^2+q^2+pq}{3} + p + q\right]|\{N\}(p,q),(Y_R J J_3)\rangle, \\ &\mathbf{R}^2|\{N\}(p,q),(Y_R J J_3)\rangle \\ &= J(J+1)\{N\}(p,q),(Y_R J J_3), \end{aligned} \tag{5}$$

where $(Y_R J J_3)$ denote the right hypercharge and the baryon spin. The latter relation is due to the hedgehog ansatz that connects the spin to the right isospin. The states are still degenerate with respect to the left (flavor) quantum numbers (YTT_3) suppressed here. The constraint

$$R_8 = N_C B/2\sqrt{3}$$

fixes

$$Y_R = N_C B/3$$

(see [11]) and is written as the triality condition [5]

$$Y_{\max} = \frac{p+2q}{3} = B + m, \tag{6}$$

with $Y_{\max}$ representing the maximal hypercharge of the (p, q) multiplet. Thus, baryons belong to irreps of $SU(3)/Z_3$. With the octet being the lowest $B = 1$ multiplet, the number of colors must be $N_C = 3$. We also obtain the spin-statistics baryon number relation

$$(-1)^{2J+B} = 1,$$

which allows only half-integer spins for $B = 1$ [5].

From the standpoint of the quark model, the integer m in combination with (p, q) is related to the number of additional $q\bar{q}$ pairs present in the baryon state [9]. When $B = 1$, we obtain the minimal multiplets {8} and {10} for $m = 0$; the family of penta- and septuquark multiplets $\{\overline{10}\}$, {27}, {35}, and {28} for $m = 1$; and $\{\overline{35}\}$, {64}, and {81}, for $m = 2$ (Fig. 1). For the masses of the multiplets {8} $J = 1/2$, {10} $J = 3/2$,

$\{\overline{10}\}$ $J = 1/2$, $\{27\}$ $J = 3/2$, and $\{\overline{35}\}$ $J = 3/2$, simple relations

$$\begin{aligned} M_{\{10\}} - M_{\{8\}} &= 3/2\Theta_\pi, \\ M_{\{\overline{10}\}} - M_{\{8\}} &= 3/2\Theta_K, \\ M_{\{27\}} - M_{\{10\}} &= 1/\Theta_K, \\ M_{\{\overline{35}\}} - M_{\{10\}} &= 15/4\Theta_K \end{aligned} \tag{7}$$

hold. We note that the mass difference of the minimal multiplets depends on Θ_π only,[1] whereas the mass differences between minimal and nonminimal multiplets depend on Θ_K and Θ_π. With the values $\Theta_\pi \approx 5$ GeV^{-1} and $\Theta_K \approx 2.5$ GeV^{-1} from a naive Skyrme model, the estimate $M_{\{\overline{10}\}} - M_{\{8\}} \approx 0.60$ GeV [5] was obtained in agreement with (7). The mass of $\{27\}$ then lies approximately 0.10 GeV higher. In Fig. 1, we show the spectrum of all baryon multiplets with the excitation energy up to 2.5 GeV using these moments of inertia for illustration. The sequence of the lowest baryon multiplets

$$\begin{aligned} \{8\}J = \frac{1}{2}, \quad \{10\}J = \frac{3}{2}, \quad \{\overline{10}\}J = \frac{1}{2}, \\ \{27\}J = \frac{3}{2}, \quad \{35\}J = \frac{5}{2}, \ldots \end{aligned} \tag{8}$$

turns out to be unique within a large range of moments of inertia $\Theta_\pi/3 < \Theta_K < \Theta_\pi/2$, covering the realistic cases. Diagrams for the lowest nonminimal baryon multiplets $\{\overline{10}\}$ and $\{27\}$ that accommodate the interesting $S = +1$ states are shown in Fig. 2.

So far, we have considered the $SU(3)$-symmetric case. To explain the splitting of baryon states within each multiplet, we must take the explicit symmetry breaking into account.

2.2. SU(3) Symmetry Breaking

The predominant standard symmetry breaker comes from the mass and kinetic terms in the effective action that account for different meson masses and decay constants, e.g., $m_K \neq m_\pi$ and $F_K \neq F_\pi$,

$$L^{SB} = -\frac{1}{2}\Gamma(1 - D_{88}) - \Delta\sum_{a=1}^{3} D_{8a}\Omega_a^R + \ldots \tag{9}$$

(the first term). There can be further terms of minor importance that depend on the specific effective action used. As an example, we optionally include such a term

[1] It was shown for arbitrary B [9] that the factor at $1/2\Theta_K$ in (3), $C_2(SU(3)) - \mathbf{R}^2 - 3B^2/4 = 3B/2$ for any minimal multiplet with $p + 2q = 3B$; $N_c = 3$.

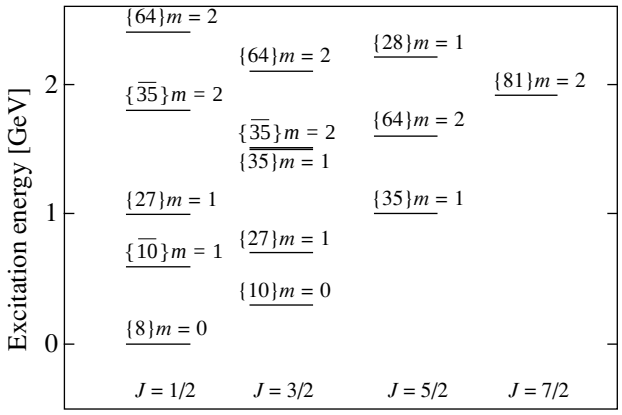

Fig. 1. $B = 1$ baryon multiplets with the excitation energy less than 2.5 GeV for $\Theta_\pi = 5$ GeV^{-1} and $\Theta_K = 2.5$ GeV^{-1}. The number m (see relation (6)) is also given.

arising from $\rho - \omega$ mixing in vector meson Lagrangians (the second term). This can serve as a test for the model dependence of our results. The corresponding Hamiltonian is

$$H^{SB} = \frac{1}{2}\Gamma(1 - D_{88}) - \frac{\Delta}{\Theta_\pi}\sum_{a=1}^{3} D_{8a}R_a + \ldots \tag{10}$$

The quantities Γ and Δ are again model-dependent quantities; they determine the strength of symmetry breaking. We first consider only the standard symmetry breaker Γ.

It was noticed earlier that a perturbative treatment of this symmetry breaker leads to the splitting

$$(M_\Lambda - M_N) : (M_\Sigma - M_\Lambda) : (M_\Xi - M_\Sigma) = 2 : 2 : 1$$

for the $\{8\}$ baryons [4, 11], at variance with the observation. Because symmetry breaking is strong, Eq. (10) must be diagonalized in the basis of the unperturbed eigenstates of H^S. Under this procedure, the states of a certain multiplet acquire components of higher representations. We nevertheless call the mixed states $\{8\}$ states, $\{10\}$ states, etc., according to their predominant contribution.

The best values for the moments of inertia Θ_π and Θ_K and the symmetry breaker Γ are listed in Table 1 (fit *A*). Optionally, the symmetry breaker Δ is also included (fit *B*). In Fig. 3, we show the dependence of the Z and Z^* energies on the kaonic moment of inertia Θ_K with the other parameters kept fixed. The sensitive dependence expected from Eq. (7) persists when symmetry breaking is included. If the experimental datum for Z proves correct, a relatively large kaonic moment of inertia (Table 1) is required.

We now compare this with the implications of the standard Skyrme model [2, 3] with only one adjustable

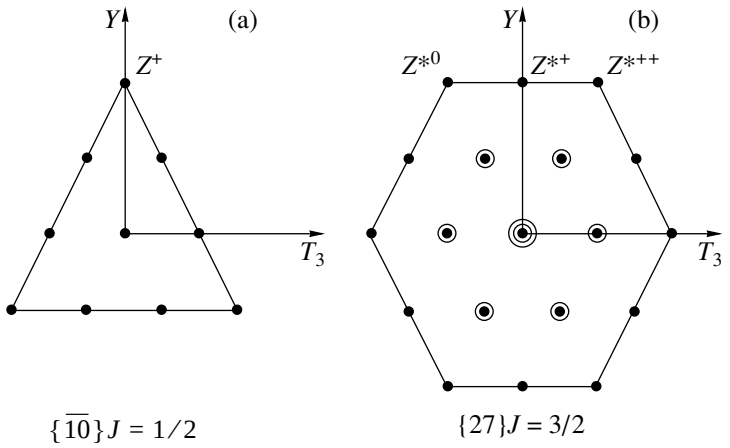

Fig. 2. The T_3–Y diagrams for the baryon multiplets $\{\overline{10}\}$ and $\{27\}$ that include the lowest $S = +1$ states.

parameter $e = 4.05$. The mass and kinetic symmetry breakers are included with mesonic parameters. The kinetic symmetry breaker involves time derivatives that were neglected in [6] (in the adiabatic approximation) in accordance with the argument that they are suppressed by two orders in the $1/N_C$ expansion and this order must include many other symmetry breaking terms that are not taken into account either. This leads to $\Theta_\pi = 5.88\ \text{GeV}^{-1}$, $\Gamma = 1.32$ GeV, and a relatively small kaonic moment of inertia $\Theta_K = 2.19\ \text{GeV}^{-1}$ (connected with larger Z and Z^* masses, Fig. 3). However, the nonadiabatic terms in the kinetic symmetry breaker are not actually small, and they make a sizeable contribution to the kaonic moment of inertia $\Theta_K = 2.80\ \text{GeV}^{-1}$ together with symmetry breaking terms and even terms that are nondiagonal in the angular momenta. Because the latter were never properly treated, these numbers should be compared with reservation to those in Table 1. Nevertheless, it seems that the standard Skyrme model can potentially provide values close to fit *B*. Relative to fit *A*, the standard symmetry breaker from the Skyrme model appears too weak, indicating that an important symmetry breaking piece is missing in this model. Concluding this discussion, we emphasize that the nonadiabatic terms in the kinetic symmetry breaker are of course not the only possibility of obtaining larger kaonic moments of inertia. Inclusion of other degrees of freedom or of additional terms in the effective action influences this quantity sensitively. In this respect, the position of the exotic Z baryon proves an important constraint on soliton models.

Table 1. Moments of inertia and symmetry breakers obtained from a fit to the baryon spectrum including the novel Z datum

	Θ_π, GeV^{-1}	Θ_K, GeV^{-1}	Γ, GeV	Δ
fit *A*	5.61	2.84	1.45	–
fit *B*	5.87	2.74	1.34	0.40

The resulting baryon spectrum is shown in Fig. 4. It can be seen that for fit *A*, with the standard symmetry breaker alone, the $\Sigma - \Lambda$ mass difference is too large, the splitting in the $J = 1/2$ multiplets relative to that in the $J = 3/2$ multiplets is overestimated, and the corresponding $SU(2)$ symmetry breaker can account for only half of the neutron–proton split (not shown here; see, e.g., [6]). All three deficiencies can be remedied by including the second symmetry breaker, fit *B*. Of course, this does not mean that the additional symmetry breaker must be exactly of form (10); other operator structures are possible. As mentioned, we include fit *B* mainly to illustrate the model dependence of our results. It seems that the levels of the $\{\overline{10}\}$ multiplet are relatively stable in contrast to the $\{27\}$ multiplet, whose states depend sensitively on the specific form of the symmetry breakers such that even the ordering of the levels is changed.

All states of the $\{\overline{10}\}$ and $\{27\}$ multiplets are listed in Tables 2 and 3. We distinguish states with exotic quantum numbers from those with nonexotic quantum numbers $-2 \le Y \le 1$ and $T \le 1 + Y/2$. Generally, the former are "cleaner," because they cannot mix with vibrational excitations (apart from their own radial excitations). Because additional vibrations on top of these states can only enhance the energy, these turn out to be the lowest states with exotic quantum numbers starting with the $S = 1$ baryon states Z and Z^*. The latter are experimentally accessible via the reactions

$$\gamma N \longrightarrow \bar{K}Z \longrightarrow \bar{K}KN,$$

$$\pi N \longrightarrow \bar{K}Z \longrightarrow \bar{K}KN,$$

$$NN \longrightarrow (\Lambda, \Sigma)Z \longrightarrow (\Lambda, \Sigma)KN,$$

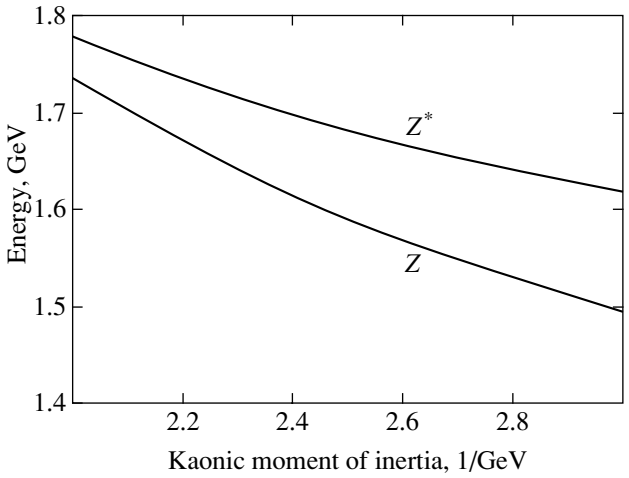

Fig. 3. The masses of the $S = 1$ baryons Z and Z^* depending on the kaonic moment of inertia. $\Theta_\pi = 5.87$ GeV^{-1} and $\Gamma = 1.34$ GeV are kept fixed.

and in KN scattering. The novel measurement [1] was a photoproduction experiment of the first type. The $S \neq 1$ exotic states are more difficult to measure, e.g., X in Table 2 via the reactions

$$\pi N \longrightarrow \pi X \longrightarrow \pi\pi\Delta,$$

$$NN \longrightarrow \Delta X \longrightarrow \pi\Delta\Delta.$$

We also included the lowest exotic states with strangeness $S = +2$ and $S = -4$ with the main components in the respective multiplets $\{\overline{35}\}$ and $\{35\}$. The $S = +2$ state Z^{**} can still be produced in binary reactions, e.g.,

$$K^0 p \longrightarrow K^- Z^{**,++},$$

but the energy of this state is already quite considerable, approximately 2.4 GeV. On the other hand, the $S = -4$ state is more difficult to produce, but detection seems to be simpler because the final Ω^- and K^- are easy to see.

In contrast, the states with nonexotic quantum numbers in Table 3 mix strongly with vibrational excitations of the {8} and {10} baryons. For example, the N^* rotational state, identified with the nucleon resonance $P11(1.71)$ in [7], mixes strongly with a $2\hbar\omega$ radial excitation, which can even lead to a doubling of states as found in [8]. This situation renders an easy interpretation difficult. Probably the cleanest of these states with nonexotic quantum numbers is the one called Λ^*, which predominantly couples with the nonresonant magnetic dipole mode. But even here, it is not excluded that the good agreement with the position of the experimental Λ resonance $P03(1.89)$ is accidental. Also, there is not even a candidate for the rotational state called Δ^* listed by the PDG in the required energy region with the empirical Δ resonance $P33(1.92)$ lying approximately 0.1 GeV too high. On the other hand, do we have candidates close to the estimated energies in five cases. There is certainly some evidence that the numbers presented are not unreasonable.

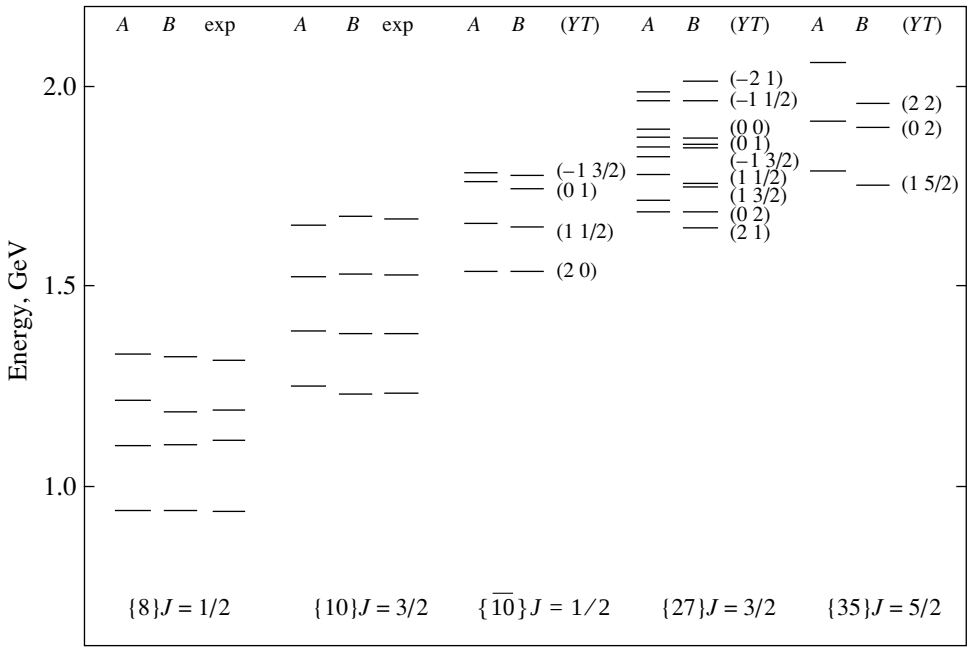

Fig. 4. The lowest rotational states in the $SU(3)$ soliton model for fits A and B. The experimental masses of the {8} and {10} baryons are depicted for comparison. Not all states of {35} are shown.

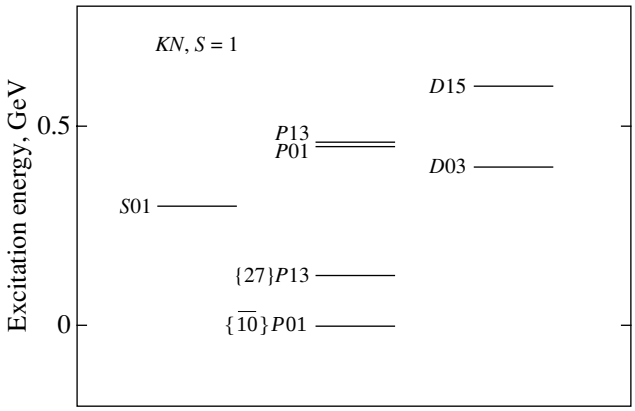

Fig. 5. Tentative baryon spectrum for the $S = 1$ sector.

It should be added that the energies for the $\{\overline{10}\}$ baryons presented here differ substantially from those obtained in [7] using simple perturbation theory. Their $\{\overline{10}\}$ total splitting is overestimated by more than a factor of 2.

3. THE $S = 1$ BARYON SPECTRUM

So far, we have considered only the rotational states. The real situation is complicated because there is an entire tower of vibrational excitations connected with each of these rotational states. We briefly address this issue on a qualitative level, particularly for the $S = 1$ sector. This may possibly be of help to experimentalists in search of more exotic baryons.

The lowest states in the $S = 1$ sector are the rotational states Z and Z^* discussed in the previous section. As mentioned, we believe that the energies of these two states must be close to each other and the energy of Z^* somewhat larger (about 0.10–0.15 GeV). Such rotational states appear as sharp resonances with small widths relative to the broader vibrational states. The width of Z was given in [1] to be smaller than 25 MeV, and that of Z^* must be somewhat larger due to phase space arguments. The Z^* state will probably be the next exotic state detected.

In soliton models, radial excitations (breathing modes) certainly exist for each rotational state. For most of the {8} and {10} baryons, such excitations correspond to the well-known resonances; e.g., the Roper resonance for the nucleon. The breathing mode excitation energy of Z calculated in [8] is approximately 0.45 GeV, and that of Z^* should be considerably smaller because the latter object is more extended due to centrifugal forces related to a larger spin (similarly to the Roper resonance and the Δ resonance $P33(1.60)$). We can therefore expect the excited $P01$ and $P13$ states to occur close to each other, as indicated in Fig. 5 (the order may be reversed!).

In addition, there must be strong quadrupole excitations, as those obtained in soliton models [12] and seen empirically in the well-studied $S = 0$ and $S = -1$ sectors (with roughly 0.4 and 0.6 GeV excitation energy). In

Table 2. Rotational states of nonminimal multiplets with exotic quantum numbers below 2 GeV including all members of $\{\overline{10}\}$ and {27}. The experimental Z datum enters the fits. The lowest exotic $Y = \pm 3$ baryon states are also included

	J	Y	T	Decay modes	Estimated energy, GeV	
					A	B
$Z\{\overline{10}\}$	1/2	2	0	KN	1.54	1.54
$Z^*\{27\}$	3/2	2	1	KN	1.69	1.65
$\{27\}$	3/2	0	2	$\pi\Sigma$, $\pi\Sigma^*$, $\pi\pi\Lambda$	1.72	1.69
X{35}	5/2	1	5/2	$\pi\Delta$, $\pi\pi N$	1.79	1.76
$\{\overline{10}\}$	1/2	–1	3/2	$\pi\Xi$, $\pi\Xi^*$, $\bar{K}\Sigma$	1.79	1.78
$\{27\}$	3/2	–1	3/2	$\pi\Xi$, $\pi\Xi^*$, $\bar{K}\Sigma$	1.85	1.85
$\{35\}$	5/2	0	2	$\pi\Sigma$, $\pi\Sigma^*$	1.92	1.90
$\{35\}$	5/2	2	2	$K\Delta$, $K\pi N$	2.06	1.96
$\{27\}$	3/2	–2	1	$\pi\Omega$, $\bar{K}\Xi$, $\bar{K}\Xi^*$	1.99	2.02
$\{35\}$	5/2	–3	1/2	$\bar{K}\Omega$, $\bar{K}\bar{K}\Xi$	2.31	2.36
$Z^{**}\{\overline{35}\}$	3/2	3	1/2	KKN, $KK\Delta$	2.41	2.38

Table 3. Rotational states of higher multiplets with nonexotic quantum numbers below 2 GeV including all members of the $\{\overline{10}\}$ and {27} multiplets

	J	Y	T	Candidate	Estimated energy, GeV	
					A	B
$N^*\{\overline{10}\}$	1/2	1	1/2	$NP11(1.71)$***	1.66	1.65
$\Sigma^*\{\overline{10}\}$	1/2	0	1	$\Sigma P11(1.77)$*	1.77	1.75
Σ^* {27}	3/2	1	3/2		1.83	1.75
{27}	3/2	1	1/2	$NP13(1.72)$****	1.78	1.76
{27}	3/2	0	1	$\Sigma P13(1.84)$*	1.90	1.86
Σ^* {27}	3/2	0	0	$\Delta P03(1.89)$****	1.88	1.87
{27}	3/2	–1	1/2	Ξ ??(1.95)***	1.97	1.97

these sectors, a number of S-wave resonances appear via $\bar{K}N$, $K\Lambda$, $K\Sigma$, and $K\Xi$ bound states just below the corresponding thresholds [12]. Although such an interpretation seems less clear in the $S = 1$ sector, a low-lying $S01$ resonance is nevertheless expected, simply by inspection of the other sectors.

Tentatively, this leads to an $S = +1$ baryon spectrum, depicted in Fig. 5. The T-matrix poles $P01(1.83)$, $P13(1.81)$, $D03(1.79)$, and $D15(2.07)$ extracted from early KN scattering experiments [13] qualitatively fit such a scheme, but the spacings are considerably smaller than in Fig. 5. Therefore, if these T-matrix poles prove correct, a strong quenching of the spectrum shown in Fig. 5 has to be understood. The existence of such poles, particularly in the D-waves, would likewise favor the location of Z considerably below these resonances, compatible with the datum 1.54 GeV.

4. CONCLUSIONS

We have shown that a low position of the exotic $S = +1$ baryon Z with quantum numbers $J = 1/2$ and $T = 0$ at the reported energy 1.54 GeV is compatible with soliton models and the known baryon spectrum. For all members of the $\{\overline{10}\}$ and {27} multiplets with nonexotic quantum numbers, we find candidates close to the estimated energies, with one exception: the empirical Δ resonance $P33(1.92)$ lies approximately 0.1 GeV too high. A strong mixing of these states with vibrational modes of the {8} and {10} baryons may lead to considerable energy shifts and even to a doubling of states. Also, the T-matrix poles of early KN scattering experiments favor a low Z baryon sufficiently below these resonances, with the caveat that when these poles are correct, a strong quenching of the $S = 1$ baryon spectrum compared to other sectors has to be explained.

However, the soliton model by itself does not exclude a Z baryon at higher energies. Confirmation of this datum, which proves a stringent constraint on these models, is therefore extremely important.

Under the assumption that the exotic Z baryon is actually located at the reported position, we have estimated the energies of other exotic baryons. First of all, there is a further $S = +1$ baryon Z^* with quantum numbers of $J = 3/2$ and $T = 1$, some 0.10–0.15 GeV above Z. This will probably be the next state to be discovered in similar experiments, also as a sharp resonance with a somewhat larger width. Moreover, there will be a tower of vibrational excitations built on these two exotic states, which should appear as broader resonances several 0.1 GeV above these energies.

There are also several low-lying $S \neq 1$ baryons with exotic isospin, starting with a $J = 3/2$ state with quantum numbers $Y = 0$ and $T = 2$ at approximately 1.7 GeV. These states are more difficult to access experimentally. The lowest $S = +2$ and $S = -4$ baryon states may also be of some interest, although they are already expected at high energies of about 2.3–2.4 GeV.

ACKNOWLEDGMENTS

One of the authors (V. B. K.) is indebted to B.O. Kerbikov, A.E. Kudryavtsev, L.B. Okun', and other participants of the ITEP Thursday seminar for useful questions and discussions; his work was supported by the Russian Foundation for Basic Research (project no. 01-02-16615).

Note added in proof (September 2, 2003). In a recent paper by V.V. Barmin *et al.* [hep-ex/0304040], the Z^+ hyperon is observed in K^+ collisions with Xe nuclei. The spectrum of the $K_S^0 p$ effective mass shows an enhancement at $M = 1539 \pm 2$ MeV with a width of

$\Gamma \leq 9$ MeV, in agreement with the results of T. Nakano *et al.* [1].

Further confirmation of the existence of Z^+ comes from CLAS collaboration [hep-ex/0307018] and from SAPHIR collaboration [hep-ex/0307083]. Discussion of the Skyrme model predictions for the baryon antidecuplat spectrum was presented recently by M. Praszalovicz [hep-ph/0308114].

REFERENCES

1. T. Nakano *et al.*, hep-ex/0301020.
2. T. H. R. Skyrme, Proc. R. Soc. London, Ser. A **260**, 127 (1961); Nucl. Phys. **31**, 556 (1962).
3. E. Witten, Nucl. Phys. B **223**, 422 (1983); Nucl. Phys. B **223**, 433 (1983).
4. M. Chemtob, Nucl. Phys. B **256**, 600 (1985).
5. L. C. Biedenharn and Y. Dothan, *From SU(3) to Gravity*, Ne'eman Festschrift (Cambridge Univ. Press, Cambridge, 1986).
6. H. Walliser, Nucl. Phys. A **548**, 649 (1992); in *Baryons as Skyrme Solitons*, Ed. by G. Holzwarth (World Sci., Singapore, 1992), p. 247.
7. D. Diakonov, V. Petrov, and M. Polyakov, Z. Phys. A **359**, 305 (1997).
8. H. Weigel, Eur. Phys. J. A **2**, 391 (1998).
9. V. B. Kopeliovich, Phys. Lett. B **259**, 234 (1991).
10. V. B. Kopeliovich, B. Schwesinger, and B. E. Stern, Nucl. Phys. A **549**, 485 (1992).
11. E. Guadagnini, Nucl. Phys. B **236**, 35 (1984).
12. B. Schwesinger, Nucl. Phys. A **537**, 253 (1992).
13. J. S. Hyslop, R. A. Arndt, L. D. Roper, and R. L. Workman, Phys. Rev. D **46**, 961 (1992).

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 441–465.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 491–515.

NUCLEI, PARTICLES, AND THEIR INTERACTION

Nonlinear $k_\perp$-Factorization for Forward Dijets in Deep Inelastic Scattering off Nuclei in the Saturation Regime¶

N. N. Nikolaev[a,b,*], W. Schäfer[a], B. G. Zakharov[b], and V. R. Zoller[c]

[a]*Institut für Kernphysik, Forschungszentrum Jülich D-52425, Jülich, Germany*
[b]*Landau Institute for Theoretical Physics, Chernogolovka, Moscow oblast, 142432 Russia*
**e-mail: N.Nikolaev@fz-juelich.de*
[c]*Institute for Theoretical and Experimental Physics, Moscow, 117259 Russia*

Received March 11, 2003

Abstract—We develop a QCD description of the breakup of photons into forward dijets in small-x deep inelastic scattering off nuclei in the saturation regime. Based on the color dipole approach, we derive a multiple scattering expansion for intranuclear distortions of the jet–jet transverse momentum spectrum. Special attention is paid to the non-Abelian aspects of the propagation of color dipoles in the nuclear medium. We report a nonlinear $k_\perp$-factorization formula for the breakup of photons into dijets in terms of the collective Weizsäcker–Williams glue of nuclei defined in [5, 6]. For hard dijets with the transverse momenta above the saturation scale, the azimuthal decorrelation (acoplanarity) momentum is on the order of the nuclear saturation momentum Q_A. For minijets with the transverse momentum below the saturation scale, the nonlinear $k_\perp$-factorization predicts a complete disappearance of the jet–jet correlation. We comment on possible relevance of the nuclear decorrelation of jets to the experimental data from the STAR-RHIC Collaboration. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

From the parton model point of view, the opacity of heavy nuclei to high-energy projectiles entails a highly nonlinear relation between the parton densities of free nucleons and nuclei. The trademark of the conventional pQCD factorization theorems for hard interactions of leptons and hadrons is that the hard scattering observables are linear functionals of the appropriate parton densities in the projectile and target [1]. The parton model interpretation of hard phenomena in ultrarelativistic heavy ion collisions calls upon the understanding of factorization properties in the nonlinear regime. A priori, it is not obvious that nuclear parton densities can be defined such that they enter different observables in a universal manner. Indeed, opacity of nuclei brings in a new large scale Q_A that separates the regimes of opaque nuclei and weak attenuation [2–5]. Furthermore, for parton momenta below the saturation scale Q_A, the evolution of sea from gluons was shown to be dominated by the anticollinear, anti-DGLAP splitting [5]. In our early studies [5, 6], we have demonstrated that such observables as the amplitude of the coherent hard diffractive breakup of a projectile on a nucleus or the transverse momentum distribution of forward quark and antiquark jets in deep inelastic scattering (DIS) off the nucleus and/or the sea parton density of nuclei can be cast in precisely the same $k_\perp$-factorization form as for a free nucleon target. Specifically, this only requires replacing the unintegrated gluon structure function (SF) of the free nucleon with the collective nuclear Weizsäcker–Williams (WW) unintegrated nuclear glue, which is the expansion over the collective gluon SF of spatially overlapping nucleons of a Lorentz-contracted ultrarelativistic nucleus. This exact correspondence between the BFKL unintegrated glue of the free nucleon [7] and the nonlinear collective WW glue of the nucleus in the calculation of these observables is a heartening finding. It persists despite the sea quarks and antiquarks with the transverse momenta below Q_A being generated by the anticollinear, anti-DGLAP splitting of gluons into sea, when the transverse momentum of the parent gluons is larger than the momentum of the produced sea quarks.

In [5], we noticed that less inclusive quantities like the spectrum of leading quarks from the truly inelastic DIS or coherent diffractive breakup off nuclei are nonlinear functionals of the collective nuclear WW glue. Consequently, in the quest for factorization properties of nuclear interactions, we must go beyond the linear observables such as the amplitude of coherent diffractive breakup of pions or photons into back-to-back dijets, single-jet inclusive cross section, and/or nuclear sea parton density. In this paper, we discuss the truly inelastic hard interaction with nuclei followed by a

¶This article was submitted by the authors in English.

breakup of the projectile into forward hard dijets.[1] We illustrate our major point in the example of DIS at small x with a breakup of the (virtual) photon into a hard, approximately back-to-back dijet with a small separation in rapidity, such that the so-called lightcone plus-components of the jet momenta sum up to the lightcone plus-component of the photon momentum; i.e., the so-called $x_\gamma = 1$ criterion is fulfilled (see, e.g., [10] and references therein). In the familiar collinear approximation, such a dijet originates from the photon–gluon fusion $\gamma^* g \longrightarrow q\bar{q}$, often referred to as the interaction of the unresolved or direct photon. Allowing a transverse momentum of gluons leads to a disparity of the momenta and to an azimuthal decorrelation of the quark and antiquark jets, which can be quantified in DIS off free protons within the $k_\perp$-factorization in terms of the unintegrated gluon SF of the target (see [11, 12] and references therein). A substantial nuclear broadening of the unintegrated gluon SF of nuclei at small x and of the nuclear sea parton distributions [2, 5] points to a stronger azimuthal decorrelation of jets produced in DIS off nuclei. Furthermore, our finding of anticollinear, anti-DGLAP splitting of gluons into sea strongly suggests the complete azimuthal decorrelation of forward quark and antiquark jets with the transverse momenta below the saturation scale, $p_\pm \lesssim Q_A$. In this paper, we quantify these expectations and formulate a nonlinear generalization of the $k_\perp$-factorization for the inclusive dijet spectrum.

The technical basis of our approach is the color-dipole multiple-scattering theory of small-x DIS off nuclei [13, 14]. We derive a consistent $k_\perp$-factorization description of the azimuthal decorrelation of jets in terms of the collective WW unintegrated gluon SF of the nucleus. In this derivation, we closely follow our early work [5] on the color-dipole approach to saturation of nuclear partons. We focus on DIS at $x \lesssim x_A = 1/R_A m_N \ll 1$, which is dominated by interactions of $q\bar{q}$ Fock states of the photon. Here, m_N is the nucleon mass and R_A is the radius of the target nucleus of the mass number A. Nuclear attenuation of these $q\bar{q}$ color dipoles [13, 15] quantifies the fusion of gluons and sea quarks from spatially overlapping nucleons of the Lorentz contracted nucleus ([16], see also [3, 4]). Here, we also report some of the technical details, especially on the non-Abelian aspects of propagation of color dipoles in nuclear matter, which were omitted in the letter publication [5].

We focus on genuinely inelastic DIS followed by color excitation of the target nucleus. For heavy nuclei, equally important is the coherent diffractive DIS, in which the target nucleus does not break up and is retained in the ground state. Coherent diffractive DIS makes up 50% of the total DIS events at small x [14]; in these coherent diffractive events, quark and antiquark jets are produced exactly back-to-back with a negligibly small transverse decorrelation momentum

$$|\mathbf{\Delta}| = |\mathbf{p}_+ + \mathbf{p}_-| \lesssim 1/R_A \sim m_\pi/A^{1/3}.$$

This paper is organized as follows. We work at the parton level and discuss the transverse momentum distribution of the final state quark and antiquark in interactions of $q\bar{q}$ Fock states of the photon with heavy nuclei. In Section 2, we set up the formalism with a brief discussion of the decorrelation of jets in DIS off free nucleons. In Section 3, we report the derivation of the general formula for the two-body transverse momentum distribution. Color exchange between the initially color-neutral $q\bar{q}$ dipole and the nucleons of the target nucleus leads to intranuclear propagation of the color-octet $q\bar{q}$-states. Our formalism, based on the technique described in [17, 18], consistently includes the diffractive attenuation of octet dipoles and effects of transitions between color-singlet and color-octet $q\bar{q}$ pairs, as well as between different color-octet states of the $q\bar{q}$ pair. The hard jet–jet inclusive cross section is discussed in Section 4. For hard dijets, diffractive attenuation effects are weak and we obtain a nuclear $k_\perp$-factorization formula for the broadening of azimuthal correlations between the quark and antiquark jets, which is reminiscent of that for a free nucleon target and is still a linear functional of the collective WW gluon SF of the nucleus. We relate the decorrelation (acoplanarity) momentum to the nuclear saturation scale Q_A. In Section 5, working in the large-N_c approximation, we derive our central result, a nonlinear nuclear $k_\perp$-factorization formula for the inclusive dijet cross section, and prove the complete disappearance of the jet–jet correlation for minijets with the transverse momentum below the saturation scale Q_A. In Section 6, we present numerical estimates for the acoplanarity momentum distribution based on the unintegrated glue of the proton determined in [19]. We point out a strong enhancement of decorrelations from the average to central DIS and comment on the possible relevance of our mechanism of azimuthal decorrelations to the recent observation of the dissolution of the away jets in central nuclear collisions at RHIC [20]. The next-to-leading order $1/N_c^2$-corrections to the large-N_c results in Section 5 are discussed in Section 7. Here, we derive a nonlinear $k_\perp$-factorization representation for the $1/N_c^2$ corrections and establish a close connection between the $1/N_c$ and higher twist expansions. In Section 8, we summarize our principal findings.

Some of the technical details are presented in the Appendices. In Appendix A, we present the calculation of the matrix of 4-body cross sections that enters the evolution operator for the intranuclear propagation of color dipoles. In Appendix B, we revisit the single-jet

[1] Preliminary results of this study have been reported elsewhere [8, 9].

spectrum and total cross section of DIS off nuclei and demonstrate how the color-dipole extension [13, 14] of the Glauber–Gribov results [21, 22] is recovered despite a nontrivial spectrum of eigen–cross sections for the non-Abelian propagation of color dipoles in the nuclear matter. The properties of the collective unintegrated gluon SF for overlapping nucleons of a Lorentz-contracted ultrarelativistic nucleus are discussed in Appendix C.

2. $k_\perp$-FACTORIZATION FOR BREAKUP OF PHOTONS INTO FORWARD DIJETS IN DIS OFF FREE NUCLEONS

We briefly recall the color dipole formulation of DIS [13, 14, 23–25] and set up the formalism in the example of jet–jet decorrelation in DIS off free nucleons at moderately small x, which is dominated by interactions of $q\bar{q}$ states of the photon. The total cross section for the interaction of the color dipole $\mathbf{r}$ with the target nucleon is given by [26, 27]

$$\begin{aligned}\sigma(r) &= \alpha_S(r)\sigma_0\int d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa})(1-e^{i\boldsymbol{\kappa}\cdot\mathbf{r}})\\ &= \frac{1}{2}\alpha_S(r)\sigma_0\int d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa})(1-e^{i\boldsymbol{\kappa}\cdot\mathbf{r}})(1-e^{-i\boldsymbol{\kappa}\cdot\mathbf{r}}),\end{aligned} \tag{1}$$

where σ_0 is an auxiliary soft parameter and α_S is the running coupling constant for the gauge group $SU(N_c)$. The function $f(\boldsymbol{\kappa})$ is normalized, as $\int d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa}) = 1$ and is related to the BFKL unintegrated gluon SF of the target nucleon $\mathcal{F}(x, \kappa^2) = \partial G(x, \kappa^2)/\partial\ln\kappa^2$ ([7], see also [19, 28] for the phenomenology and review) by

$$f(\boldsymbol{\kappa}) = \frac{4\pi}{N_c\sigma_0}\frac{1}{\kappa^4}\mathcal{F}(x, \kappa^2). \tag{2}$$

For DIS off a free nucleon target (see Figs. 1a–1d), the total photoabsorption cross section is given by [13]

$$\sigma_N(Q^2, x) = \int d\mathbf{r}dz\left|\Psi(Q^2, z, \mathbf{r})\right|^2\sigma(x, \mathbf{r}), \tag{3}$$

where $\Psi(Q^2, z, \mathbf{r})$ is the wave function of the $q\bar{q}$ Fock state of the photon and Q^2 and x are the standard DIS variables. In the momentum representation,

$$\begin{aligned}\frac{d\sigma_N}{d\mathbf{p}_+dz} &= \frac{\sigma_0}{2}\frac{\alpha_S(\mathbf{p}_+^2)}{(2\pi)^2}\\ &\times\int d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa})\left|\langle\gamma^*|z, \mathbf{p}_+\rangle - \langle\gamma^*|z, \mathbf{p}_+ - \boldsymbol{\kappa}\rangle\right|^2,\end{aligned} \tag{4}$$

where $\mathbf{p}_+$ is the transverse momentum of the quark, the antiquark has the transverse momentum $\mathbf{p}_- = -\mathbf{p}_+ + \boldsymbol{\kappa}$, and $z_+ = z$ and $z_- = 1 - z$ are the fractions of the photon lightcone momentum carried by the quark and antiquark, respectively. The variables $z_\pm$ for the observed jets add up to unity, $x_\gamma = z_+ + z_- = 1$, which in the realm of DIS is said to be the unresolved or direct photon interaction.

Summing over the helicities λ and $\bar{\lambda}$ of the final state quark and antiquark, we obtain

$$\begin{aligned}\left|\langle\gamma^*|z, \mathbf{p}\rangle - \langle\gamma^*|z, \mathbf{p}-\boldsymbol{\kappa}\rangle\right|^2_{\lambda_\gamma=\pm1} &= 2N_ce_f^2\alpha_{em}\\ \times\left\{[z^2+(1-z)^2]\left(\frac{\mathbf{p}}{\mathbf{p}^2+\varepsilon^2} - \frac{\mathbf{p}-\boldsymbol{\kappa}}{(\mathbf{p}-\boldsymbol{\kappa})^2+\varepsilon^2}\right)^2_{\lambda+\bar{\lambda}=0}\right.&\\ \left. + m_f^2\left(\frac{1}{\mathbf{p}^2+\varepsilon^2} - \frac{1}{(\mathbf{p}-\boldsymbol{\kappa})^2+\varepsilon^2}\right)^2_{\lambda+\bar{\lambda}=\lambda_\gamma}\right\}&\end{aligned} \tag{5}$$

for transverse photons and quarks of flavor f and

$$\begin{aligned}\left|\langle\gamma^*|z, \mathbf{p}\rangle - \langle\gamma^*|z, \mathbf{p}-\boldsymbol{\kappa}\rangle\right|^2_{\lambda_\gamma=0}&\\ = 8N_ce_f^2\alpha_{em}Q^2z^2(1-z)^2&\\ \times\left(\frac{1}{\mathbf{p}^2+\varepsilon^2} - \frac{1}{(\mathbf{p}-\boldsymbol{\kappa})^2+\varepsilon^2}\right)^2_{\lambda+\bar{\lambda}=\lambda_\gamma}&\end{aligned} \tag{6}$$

for longitudinal photons, where $\varepsilon^2 = z(1-z)Q^2 + m_f^2$.

We now note that the transverse momentum of the gluon is precisely the decorrelation momentum $\boldsymbol{\Delta} = \mathbf{p}_+ + \mathbf{p}_-$, and in the differential form, we have

$$\begin{aligned}\frac{d\sigma_N}{dzd\mathbf{p}_+d\boldsymbol{\Delta}} &= \frac{\sigma_0}{2}\frac{\alpha_S(\mathbf{p}_+^2)}{(2\pi)^2}f(\boldsymbol{\Delta})\\ &\times\left|\langle\gamma^*|z, \mathbf{p}_+\rangle - \langle\gamma^*|z, \mathbf{p}_+ - \boldsymbol{\Delta}\rangle\right|^2\\ &= \frac{\alpha_S(\mathbf{p}_+^2)}{2\pi N_c}\frac{\mathcal{F}(x, \boldsymbol{\Delta}^2)}{\Delta^4}\left|\langle\gamma^*|z, \mathbf{p}_+\rangle - \langle\gamma^*|z, \mathbf{p}_+ - \boldsymbol{\Delta}\rangle\right|^2.\end{aligned} \tag{7}$$

The small-x result in Eq. (7) shows that in DIS, forward dijets acquire their large transverse momentum from the intrinsic momentum of the quark and antiquark in the wave function of the projectile photon; hence, it is appropriate to call this process the breakup of the photon into forward hard dijets. In addition to the criterion $x_\gamma = 1$, the experimental signature of photon breakup is a small rapidity separation of forward jets, $z_+ \sim z_-$. The perturbative hard scale for our process is set by $\mathbf{p}^2 = \mathbf{p}_+^2 + Q^2z(1-z)$, and the gluon SF of the proton enters Eq. (7) at the Bjorken variable $x = (4\mathbf{p}_+^2 + Q^2)/W^2$, where W is the γ^*p center-of-mass energy. The purpose of our study is an extension of Eq. (7) to the breakup of photons into dijets in truly inelastic DIS on nuclear targets.

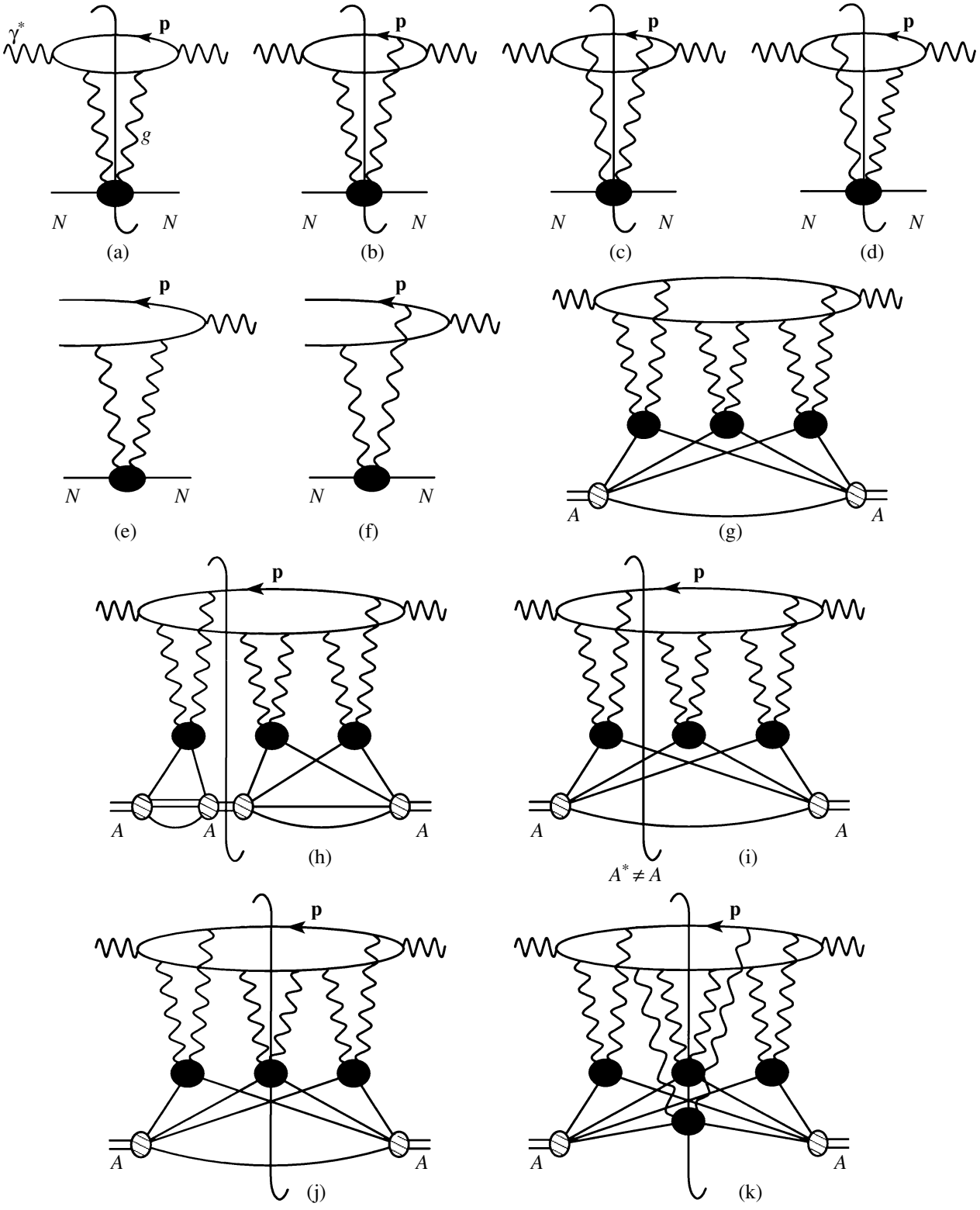

Fig. 1. pQCD diagrams for the cross section of inclusive DIS off nucleons (a–d) and nuclei (g–k) and the amplitude of diffractive DIS off protons (e, f). Diagrams a–d show the unitarity cuts with color excitation of the target nucleon; g is a generic multiple scattering diagram for the Compton scattering amplitude off nucleus; h is the unitarity cut for a coherent diffractive DIS with retention of the ground state nucleus A in the final state; i is the unitarity cut for quasielastic diffractive DIS with excitation and breakup of the nucleus A^*; and j and k are the unitarity cuts for truly inelastic DIS with single (j) and multiple (k) color excitation of nucleons of the nucleus.

3. BREAKUP OF PHOTONS INTO DIJETS ON NUCLEAR TARGETS

We focus on DIS at $x \lesssim x_A = 1/R_A m_N \ll 1$, which is dominated by interactions of $q\bar{q}$ states of the photon. This is a starting term of the leading $\ln(1/x)$ expansion; extension to interactions of higher Fock states of the photon and the corresponding $\ln(1/x)$ evolution to smaller x will be discussed elsewhere. For $x \lesssim x_A$, the propagation of the $q\bar{q}$ pair inside the nucleus can be treated in the straight-path approximation.

We work in the conventional approximation of two t-channel gluons in DIS off free nucleons. The relevant unitarity cuts of the forward Compton scattering amplitude shown in Figs. 1a–1d describe the transition from the color-neutral $q\bar{q}$ dipole to the color-octet $q\bar{q}$ pair.[2] The two-gluon exchange approximation amounts to neglecting unitarity constraints in DIS off free nucleons. As a quantitative measure of unitarity corrections, one can take the diffractive DIS off free nucleons, whose amplitude is described by higher order diagrams in Figs. 1e and 1f [23, 24, 27] and which is only a small fraction of the total DIS, $\eta_D \ll 1$ [29–31]. The unitarity cuts of the nuclear Compton scattering amplitude that correspond to the genuine inelastic DIS with color excitation of the nucleus are shown in Figs. 1j and 1k. The diagram in Fig. 1k describes a consecutive color excitation of the target nucleus accompanied by the color-space rotation of the color-octet $q\bar{q}$.

Let $\mathbf{b}_+$ and $\mathbf{b}_-$ be the impact parameters of the quark and antiquark, respectively, and $S_A(\mathbf{b}_+, \mathbf{b}_-)$ be the S-matrix for the interaction of the $q\bar{q}$ pair with the nucleus. We are interested in the truly inelastic inclusive cross section summed over all excitations of the target nucleus when one or several nucleons are color excited. A convenient way to sum such cross sections is offered by the closure relation [21]. Regarding the color states c_{km} of the $q_k\bar{q}_m$ pair, we sum over all octet and singlet states. Then the 2-jet inclusive spectrum is calculated in terms of the 2-body density matrix as

$$\frac{d\sigma_{\rm in}}{dz d\mathbf{p}_+ d\mathbf{p}_-} = \frac{1}{(2\pi)^4}\int d\mathbf{b}'_+ d\mathbf{b}'_- d\mathbf{b}_+ d\mathbf{b}_-$$

$$\times \exp[-i\mathbf{p}_+\cdot(\mathbf{b}_+ - \mathbf{b}'_+) - i\mathbf{p}_-\cdot(\mathbf{b}_- - \mathbf{b}'_-)]$$

$$\times \Psi^*(Q^2, z, \mathbf{b}'_+ - \mathbf{b}'_-)\Psi(Q^2, z, \mathbf{b}_+ - \mathbf{b}_-)$$

$$\times \left\{\sum_{A^*}\sum_{km}\langle 1; A|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|A^*; c_{km}\rangle\right. \tag{8}$$

$$\times \langle c_{km}; A^*|S_A(\mathbf{b}_+, \mathbf{b}_-)|A; 1\rangle$$

$$\left. - \langle 1; A|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|A; 1\rangle\langle 1; A|S_A(\mathbf{b}_+, \mathbf{b}_-)|A; 1\rangle\right\}.$$

In the integrand in Eq. (8), we subtracted the coherent diffractive component of the final state. We note that four straight-path trajectories $\mathbf{b}_\pm$ and $\mathbf{b}'_\pm$ enter the calculation of the full-fledged 2-body density matrix and S_A and S_A^* describe the propagation of two quark-antiquark pairs, $q\bar{q}$ and $q'\bar{q}'$, inside a nucleus.

The further analysis of the integrand in Eq. (8) is a non-Abelian generalization of the formalism developed by one of the authors (B. G. Z.) for the in-medium evolution of ultrarelativistic positronium [32]. Upon application of the closure relation to sum over nuclear final states A^*, the integrand in Eq. (8) can be considered as an intranuclear evolution operator for the 2-body density matrix

$$\sum_{A^*}\sum_{km}\langle A|\{\langle 1|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|c_{km}\rangle\}|A^*\rangle$$

$$\times \langle A^*|\{\langle c_{km}|S_A(\mathbf{b}_+, \mathbf{b}_-)|1\rangle\}|A\rangle \tag{9}$$

$$= \langle A|\left\{\sum_{km}\langle 1|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|c_{km}\rangle\langle c_{km}|S_A(\mathbf{b}_+, \mathbf{b}_-)|1\rangle\right\}|A\rangle$$

(for a related discussion, also see [33]). Let the eikonal for the quark–nucleon and antiquark–nucleon QCD gluon exchange interaction be $T_+^a\chi(\mathbf{b})$ and $T_-^a\chi(\mathbf{b})$, where T_+^a and T_-^a are the $SU(N_c)$ generators for the quark and antiquark states, respectively. The vertex V_a for excitation of the nucleon, $g^aN \longrightarrow N_a^*$, into the color octet state is normalized such that after application of the closure relation, the vertex g^ag^bNN in the diagrams in Figs. 1a–1d becomes δ_{ab}. In the two-gluon exchange approximation, the S-matrix of the $(q\bar{q})$–nucleon interaction is then given by

$$S_N(\mathbf{b}_+, \mathbf{b}_-) = 1 + i[T_+^a\chi(\mathbf{b}_+) + T_-^a\chi(\mathbf{b}_-)]V_a$$

$$-\frac{1}{2}[T_+^a\chi(\mathbf{b}_+) + T_-^a\chi(\mathbf{b}_-)]^2. \tag{10}$$

The profile function for the interaction of the $q\bar{q}$ dipole with the nucleon is $\Gamma(\mathbf{b}_+, \mathbf{b}_-) = 1 - S_N(\mathbf{b}_+, \mathbf{b}_-)$. For a color-singlet dipole, $(T_+^a + T_-^a)^2 = 0$ and the dipole

[2] To be more precise, for arbitrary N_c, the color-excited $q\bar{q}$ pair is in the adjoint representation and quarks are in the fundamental representation of $SU(N_c)$; our reference to the color octet and triplet must not cause any confusion.

cross section for the interaction of the color-singlet $q\bar{q}$ dipole with the nucleon equals

$$\sigma(\mathbf{b}_+ - \mathbf{b}_-) = 2\int d\mathbf{b}_+ \langle N|\Gamma(\mathbf{b}_+, \mathbf{b}_-)|N\rangle = \frac{N_c^2 - 1}{2N_c}\int d\mathbf{b}_+ [\chi(\mathbf{b}_+) - \chi(\mathbf{b}_-)]^2. \quad (11)$$

The nuclear S-matrix of the straight-path approximation is

$$S_A(\mathbf{b}_+, \mathbf{b}_-) = \prod_{j=1}^{A} S_N(\mathbf{b}_+ - \mathbf{b}_j, \mathbf{b}_- - \mathbf{b}_j),$$

where the ordering along the longitudinal path is understood. We evaluate the nuclear expectation value in (9) in the standard dilute gas approximation. In the two-gluon exchange approximation, for each and every nucleon N_j, only the terms quadratic in $\chi(\mathbf{b}_j)$ must be kept in the single-nucleon matrix element

$$\langle N_j|S_N^*(\mathbf{b}'_+ - \mathbf{b}_j, \mathbf{b}'_- - \mathbf{b}_j)S_N(\mathbf{b}_+ - \mathbf{b}_j, \mathbf{b}_- - \mathbf{b}_j)|N_j\rangle$$

that enters the calculation of $S_A^* S_A$. Following the technique developed in [17, 18], we can reduce the calculation of the evolution operator for the 2-body density matrix (9) to the evaluation of the S-matrix $S_{4A}(\mathbf{b}_+, \mathbf{b}_-, \mathbf{b}'_+, \mathbf{b}'_-)$ for the scattering of a fictitious 4-parton state composed of the two quark–antiquark pairs in the overall color-singlet state. Because $(T_+^a)^* = -T_-^a$, the quarks entering the complex-conjugate S_A^* in (9) can be viewed as antiquarks within the two-gluon exchange approximation, and therefore,

$$\sum_{km}\langle 1|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|c_{km}\rangle\langle c_{km}|S_A(\mathbf{b}_+, \mathbf{b}_-)|1\rangle = \sum_{kmjl}\delta_{kl}\delta_{mj}\langle c_{km}c_{jl}|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle, \quad (12)$$

where $S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)$ is the S-matrix for the propagation of two quark-antiquark pairs in the overall singlet state. While the first $q\bar{q}$ pair is formed by the initial quark q and antiquark $\bar{q}$ at the respective impact parameters $\mathbf{b}_+$ and $\mathbf{b}_-$, the quark q' in the second $q'\bar{q}'$ pair propagates at the impact parameter $\mathbf{b}'_-$ and the antiquark $\bar{q}'$ at the impact parameter $\mathbf{b}'_+$. In the initial state, both quark–antiquark pairs are in color-singlet states, $|\mathrm{in}\rangle = |11\rangle$.

We introduce the normalized singlet–singlet and octet–octet states

$$|11\rangle = \frac{1}{N_c}(\bar{q}q)(\bar{q}'q'), \quad |88\rangle = \frac{2}{\sqrt{N_c^2 - 1}}(\bar{q}T^a q)(\bar{q}'T^a q'), \quad (13)$$

where N_c is the number of colors and T^a are the generators of $SU(N_c)$ in the fundamental representation. Using the color Fiertz identity,

$$\delta_j^k\delta_l^m = \frac{1}{N_c}\delta_l^k\delta_j^m + 2\sum_a (T^a)_l^k (T^a)_j^m, \quad (14)$$

we can represent the sum (12) over color states of the produced quark–antiquark pair as

$$\sum_{km}\langle c_{km}c_{km}|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle = \langle 11|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle + \sqrt{N_c^2 - 1}\langle 88|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle. \quad (15)$$

If $\sigma_4(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)$ is the color-dipole cross section operator for the 4-body state, evaluation of the nuclear expectation value for a dilute gas nucleus in the standard approximation of neglecting the size of color dipoles compared to the radius of a heavy nucleus gives [21]

$$S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-) = \exp\left\{-\frac{1}{2}\sigma_4(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)T(\mathbf{b})\right\}, \quad (16)$$

where $T(\mathbf{b}) = \int db_z n_A(b_z, \mathbf{b})$ is the optical thickness of the nucleus at the impact parameter[3]

$$\mathbf{b} = \frac{1}{4}(\mathbf{b}_+ + \mathbf{b}'_+ + \mathbf{b}_- + \mathbf{b}'_-)$$

and $n_A(b_z, \mathbf{b})$ is the nuclear matter density with the normalization $\int d\mathbf{b}T(\mathbf{b}) = A$.

[3] One should not confuse $\mathbf{b}$ with the center of gravity of color dipoles, where the impact parameters $\mathbf{b}_\pm$ and $\mathbf{b}'_\pm$ must be weighted with $z_\pm$; the difference between the two quantities is irrelevant here.

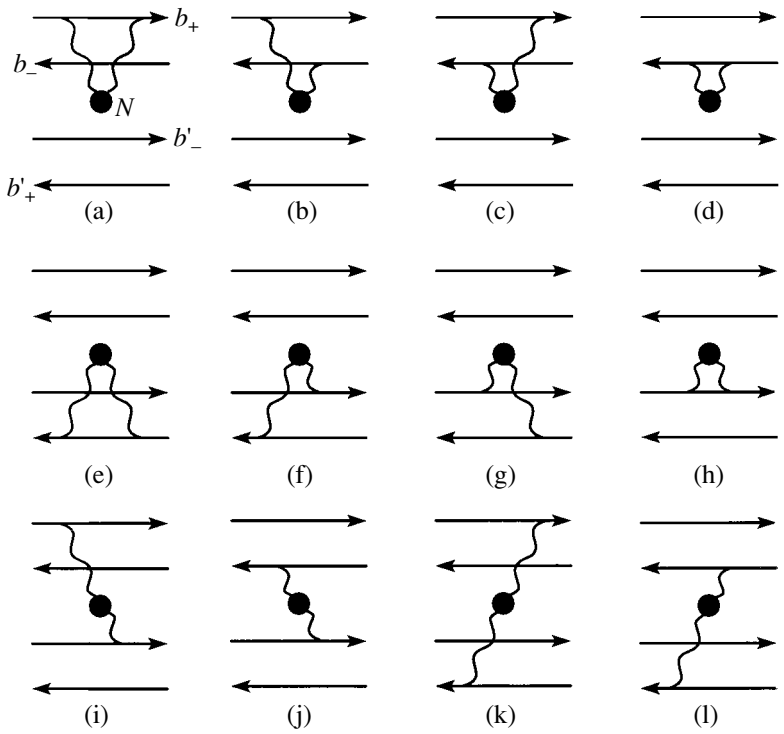

Fig. 2. pQCD diagrams for the matrix of color dipole cross section for the 4-body $(q\bar{q})(q'\bar{q}')$ state. Sets a–d and e–h show the diagrams for the scattering without changing the color state of the $q\bar{q}$ and $q'\bar{q}'$ dipoles, and sets i–l show only half of the diagrams for scattering with rotation of the color state of dipoles.

Single-nucleon S-matrix (10) contains transitions from the color-singlet to both color-singlet and color-octet $q\bar{q}$ pairs. However, only color-singlet operators contribute to

$$\langle N_j|S_N^*(\mathbf{b}'_+ - \mathbf{b}_j, \mathbf{b}'_- - \mathbf{b}_j)S_N(\mathbf{b}_+ - \mathbf{b}_j, \mathbf{b}_- - \mathbf{b}_j)|N_j\rangle,$$

and hence the matrix $\sigma_4(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)$ only includes transitions between the $|11\rangle$ and $|88\rangle$ color-singlet 4-parton states; the $|18\rangle$ states are not allowed.

The pQCD diagrams for the 4-body cross section are shown in Fig. 2. It is convenient to introduce

$$\mathbf{s} = \mathbf{b}_+ - \mathbf{b}'_+ \tag{17}$$

for the variable conjugate to the decorrelation momentum, and $\mathbf{r} = \mathbf{b}_+ - \mathbf{b}_-$, $\mathbf{r}' = \mathbf{b}'_+ - \mathbf{b}'_-$, in terms of which

$$\mathbf{b}_+ - \mathbf{b}'_- = \mathbf{s} + \mathbf{r}', \quad \mathbf{b}_- - \mathbf{b}'_+ = \mathbf{s} - \mathbf{r}, \quad \mathbf{b}_- - \mathbf{b}'_- = \mathbf{s} - \mathbf{r} + \mathbf{r}'. \tag{18}$$

Performing the relevant color algebra, we find (some details of the derivation are presented in Appendix A)

$$\sigma_{11} = \langle 11|\sigma_4|11\rangle = \sigma(\mathbf{r}) + \sigma(\mathbf{r}'), \tag{19}$$

$$\sigma_{18} = \langle 11|\sigma_4|88\rangle = \frac{\sigma(\mathbf{s}) + \sigma(\mathbf{s} - \mathbf{r} + \mathbf{r}') - \sigma(\mathbf{s} + \mathbf{r}') - \sigma(\mathbf{s} - \mathbf{r})}{\sqrt{N_c^2 - 1}} = -\frac{\Sigma_{18}(\mathbf{s}, \mathbf{r}, \mathbf{r}')}{\sqrt{N_c^2 - 1}}, \tag{20}$$

$$\sigma_{88} = \langle 88|\sigma_4|88\rangle = \frac{N_c^2 - 2}{N_c^2 - 1}[\sigma(\mathbf{s}) + \sigma(\mathbf{s} - \mathbf{r} + \mathbf{r}')] + \frac{2}{N_c^2 - 1}[\sigma(\mathbf{s} + \mathbf{r}') + \sigma(\mathbf{s} - \mathbf{r})] - \frac{1}{N_c^2 - 1}[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]. \tag{21}$$

The term in (8) that subtracts the contribution from diffractive processes without color excitation of the target nucleus is given by

$$\langle 1; A|S_A^*(\mathbf{b}'_+, \mathbf{b}'_-)|A; 1\rangle\langle 1; A|S_A(\mathbf{b}_+, \mathbf{b}_-)|A; 1\rangle = \exp\left\{-\frac{1}{2}[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\} = \exp\left\{-\frac{1}{2}\sigma_{11}T(\mathbf{b})\right\}. \tag{22}$$

In the discussion of nuclear effects, it is convenient to use the Sylvester expansion

$$\exp\left\{-\frac{1}{2}\sigma_4 T(\mathbf{b})\right\} = \exp\left\{-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right\}\frac{\sigma_4-\Sigma_2}{\Sigma_1-\Sigma_2} + \exp\left\{-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right\}\frac{\sigma_4-\Sigma_1}{\Sigma_2-\Sigma_1}, \tag{23}$$

where $\Sigma_{1,2}$ are the two eigenvalues of the operator σ_4,

$$\Sigma_{1,2} = \frac{1}{2}(\sigma_{11}+\sigma_{88}) \mp \frac{1}{2}(\sigma_{11}-\sigma_{88})\sqrt{1+\frac{4\sigma_{18}^2}{(\sigma_{11}-\sigma_{88})^2}}. \tag{24}$$

For the integrand in (8), application of the Sylvester expansion to (15) gives

$$\begin{aligned}&\sum_{A^*}\sum_{km}\langle 1;\ A|S_A^*(\mathbf{b}'_+,\mathbf{b}'_-)|A^*;\ c_{km}\rangle \times\langle c_{km};\ A^*|S_A(\mathbf{b}_+,\mathbf{b}_-)|A;\ 1\rangle \\ &-\langle 1;\ A|S_A^*(\mathbf{b}'_+,\mathbf{b}'_-)|A;\ 1\rangle\langle 1;\ A|S_A(\mathbf{b}_+,\mathbf{b}_-)|A;\ 1\rangle \\ &= (\langle 11| + \sqrt{N_c^2-1}\langle 88|)\exp\left\{-\frac{1}{2}\sigma_4 T(\mathbf{b})\right\}|11\rangle - \exp\left\{-\frac{1}{2}\sigma_{11}T(\mathbf{b})\right\} \\ &= \exp\left\{-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right\} - \exp\left\{-\frac{1}{2}\sigma_{11}T(\mathbf{b})\right\} \\ &+\frac{\sigma_{11}-\Sigma_2}{\Sigma_1-\Sigma_2}\left\{\exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right]-\exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right]\right\} \\ &+\frac{\sqrt{N_c^2-1}\,\sigma_{18}}{\Sigma_1-\Sigma_2}\left\{\exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right]-\exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right]\right\}.\end{aligned} \tag{25}$$

4. BREAKING OF PHOTONS INTO HARD DIJETS: A STILL LINEAR NUCLEAR $k_\perp$-FACTORIZATION

Diagonalization of the 2×2 matrix σ_4 is a straightforward task, and therefore, technically, Eqs. (8) and (25) allow a direct calculation of the jet–jet inclusive cross section in terms of the color dipole cross section $\sigma(\mathbf{r})$. But evaluation of the 6-fold Fourier transform is a nontrivial task.

We first note that the difference between Σ_2 and $\sigma_{11} = \sigma(\mathbf{r}) + \sigma(\mathbf{r}')$ is of the second or higher order in the off-diagonal σ_{18}, see Eq. (24). Consequently, the first two lines in Sylvester expansion (25) start with terms proportional to σ_{18}^2, whereas the last line starts with terms proportional to σ_{18}. It is then convenient to represent (25) as the impulse approximation (IA) term times the nuclear distortion factor $D_A(\mathbf{s}, \mathbf{r}, \mathbf{r}', \mathbf{b})$,

$$\begin{aligned}&\sum_{A^*}\sum_{km}\langle 1;\ A|S_A^*(\mathbf{b}'_+,\mathbf{b}'_-)|A^*;\ c_{km}\rangle \times\langle c_{km};\ A^*|S_A(\mathbf{b}_+,\mathbf{b}_-)|A;\ 1\rangle \\ &-\langle 1;\ A|S_A^*(\mathbf{b}'_+,\mathbf{b}'_-)|A;\ 1\rangle\langle 1;\ A|S_A(\mathbf{b}_+,\mathbf{b}_-)|A;\ 1\rangle \\ &= T(\mathbf{b})\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')D_A(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}),\end{aligned} \tag{26}$$

whence

$$\begin{aligned}\frac{d\sigma_{\rm in}}{d\mathbf{b}\,dz\,d\mathbf{p}_+\,d\mathbf{p}_-} &= \frac{1}{2(2\pi)^4}\int d\mathbf{s}\,d\mathbf{r}\,d\mathbf{r}' \\ &\times\exp[-i(\mathbf{p}_+ + \mathbf{p}_-)\mathbf{s} + i\mathbf{p}_-(\mathbf{r}'-\mathbf{r})]\Psi^*(Q^2,z,\mathbf{r}') \\ &\times\Psi(Q^2,z,\mathbf{r})T(\mathbf{b})\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')D_A(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}).\end{aligned} \tag{27}$$

As an introduction to nuclear $k_\perp$-factorization, we start with forward hard jets with the momenta $\mathbf{p}_\pm^2 \gtrsim Q_A^2$, which are produced from interactions with the target nucleus of small color dipoles in the incident photon such that diffractive nuclear attenuation effects can be neglected. We proceed with the formulation of the Fourier representations for each factor in (26). The application of integral representation (1) gives

$$\begin{aligned}&\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}') \\ &= [\sigma(\mathbf{s}) - \sigma(\mathbf{s}+\mathbf{r}') - \sigma(\mathbf{s}-\mathbf{r}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')] \\ &= \alpha_S\sigma_0\int d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa})e^{i\boldsymbol{\kappa}\cdot\mathbf{s}}(1-e^{i\boldsymbol{\kappa}\cdot\mathbf{r}'})(1-e^{-i\boldsymbol{\kappa}\cdot\mathbf{r}}).\end{aligned} \tag{28}$$

Hard jets correspond to $|\mathbf{r}|, |\mathbf{r}'| \ll |\mathbf{s}|$. Then, the two eigenvalues are $\Sigma_2 \approx \sigma_{11}$ and $\Sigma_1 \approx \sigma_{88} \approx 2\lambda_c\sigma(\mathbf{s})$ with $\lambda_c = N_c^2/(N_c^2-1) = C_A/2C_F$, where C_F and C_A are the Casimir operators for the fundamental and adjoint representations of $SU(N_c)$. Because $\Sigma_2 \approx \sigma_{11} \approx 0$, only the last term, proportional to σ_{18}, must be kept in the

Sylvester expansion (25), and the nuclear distortion factor takes the simple form

$$D_A(\mathbf{s}, \mathbf{r}, \mathbf{r}', \mathbf{b}) = \frac{2}{(\Sigma_2 - \Sigma_1)T(\mathbf{b})} \times \left\{ \exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right] - \exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right] \right\} = \frac{1 - \exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right]}{\frac{1}{2}\Sigma_1 T(\mathbf{b})}. \quad (29)$$

The Fourier representation for the nuclear distortion factor $D_A(\mathbf{s}, \mathbf{r}, \mathbf{r}')$ is readily obtained from the NSS representation [5, 6] for the nuclear attenuation factor,

$$\begin{aligned} &\exp\left[-\frac{1}{2}\sigma(\mathbf{s})T(\mathbf{b})\right] \\ &= \exp[-\nu_A(\mathbf{b})]\exp[\nu_A(\mathbf{b})\int d\boldsymbol{\kappa} f(\boldsymbol{\kappa})e^{i\boldsymbol{\kappa}\cdot\mathbf{s}}] \\ &= \exp[-\nu_A(\mathbf{b})]\sum_{j=0}^{\infty}\frac{\nu_A^j(\mathbf{b})}{j!}\int d\boldsymbol{\kappa} f^{(j)}(\boldsymbol{\kappa})e^{i\boldsymbol{\kappa}\cdot\mathbf{s}} \\ &= \int d\boldsymbol{\kappa}\Phi(\nu_A(\mathbf{b}), \boldsymbol{\kappa})e^{i\boldsymbol{\kappa}\cdot\mathbf{s}}, \end{aligned} \quad (30)$$

in terms of the nuclear WW glue per unit area in the impact parameter plane, $\phi_{WW}(\nu_A(\mathbf{b}), \boldsymbol{\kappa})$, defined in [5],

$$\begin{aligned} \Phi(\nu_A(\mathbf{b}), \boldsymbol{\kappa}) &= \sum_{j=0} w_j(\nu_A(\mathbf{b}))\, f^{(j)}(\boldsymbol{\kappa}) \\ &= \exp[-\nu_A(\mathbf{b})]\, f^{(0)}(\boldsymbol{\kappa}) + \phi_{WW}(\nu_A(\mathbf{b}), \boldsymbol{\kappa}). \end{aligned} \quad (31)$$

Here,

$$\nu_A(\mathbf{b}) = \frac{1}{2}\alpha_S(r)\sigma_0 T(\mathbf{b}) \quad (32)$$

and

$$w_j(\nu_A(\mathbf{b})) = \frac{\nu_A^j(\mathbf{b})}{j!}\exp[-\nu_A(\mathbf{b})] \quad (33)$$

is the probability of finding j spatially overlapping nucleons in a Lorentz-contracted nucleus; and

$$f^{(j)}(\boldsymbol{\kappa}) = \int\prod_{i=1}^{j} d\boldsymbol{\kappa}_i f(\boldsymbol{\kappa}_i)\delta\left(\boldsymbol{\kappa} - \sum_{i=1}^{j}\boldsymbol{\kappa}_i\right), \quad f^{(0)}(\boldsymbol{\kappa}) = \delta(\boldsymbol{\kappa}) \quad (34)$$

is a collective gluon field of j overlapping nucleons. As usual, the strong coupling in (32) must be taken on the hardest relevant scale [34].

The denominator Σ_1 in (29) is problematic from the point of view of the Fourier transform but can be eliminated by the integral representation,

$$\begin{aligned} D_A(\mathbf{s}) &= \int_0^1 d\beta \exp\left[-\frac{1}{2}\beta\Sigma_1 T(\mathbf{b})\right] \\ &= \int_0^1 d\beta\int d\boldsymbol{\kappa}\Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\kappa})e^{i\boldsymbol{\kappa}\cdot\mathbf{s}}. \end{aligned} \quad (35)$$

Here, β has the meaning of the fraction of the nuclear thickness that the $q\bar{q}$ pair propagates in the color octet state. The introduction of this distortion factor in (27) is straightforward and gives our central result for the hard jet–jet inclusive cross section:

$$\begin{aligned} \frac{d\sigma_{\text{in}}}{d\mathbf{b}dzd\mathbf{p}_+d\boldsymbol{\Delta}} &= T(\mathbf{b})\int d\boldsymbol{\kappa} \\ &\times \int_0^1 d\beta\Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\Delta} - \boldsymbol{\kappa})\frac{d\sigma_N}{dzd\mathbf{p}_+d\boldsymbol{\kappa}}. \end{aligned} \quad (36)$$

Because $\mathbf{r}^2 \sim 1/\mathbf{p}_+^2$ for hard jets, we must use $\alpha_S(\mathbf{p}_+^2)$ in the evaluation of $\nu_A(\mathbf{b})$. For a thin nucleus with $\nu_A(\mathbf{b}) \ll 1$, we have $\Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\Delta} - \boldsymbol{\kappa}) = \delta(\boldsymbol{\Delta} - \boldsymbol{\kappa})$ [see Eq. (31)] and recover the IA result

$$\frac{d\sigma_{\text{in}}}{d\mathbf{b}dzd\mathbf{p}_+d\boldsymbol{\Delta}} = T(\mathbf{b})\frac{d\sigma_N}{dzd\mathbf{p}_+d\boldsymbol{\Delta}}. \quad (37)$$

Our result (36) for nuclear broadening of the acoplanarity momentum distribution of hard dijets can be regarded as a nuclear counterpart of the $k_\perp$-factorization result (7) for a free nucleon target.

The probabilistic form of convolution (36) for the differential cross section on a free nucleon target with the manifestly positive-definite distribution $\Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\kappa})$ can be understood as follows. Hard jets originate from small color dipoles. Their interaction with gluons of the target nucleus is suppressed by the mutual neutralization of color charges of the quark and antiquark in the small-size color-singlet $q\bar{q}$ state, which manifests itself from the small cross section for a free nucleon target, see Eq. (7). The first inelastic interaction inside the nucleus converts the $q\bar{q}$ pair into the color-octet state in which color charges of the quark and antiquark do not neutralize each other, rescatterings of the quark and antiquark in the collective color field of intranuclear nucleons become uncorrelated, and the

broadening of the momentum distribution with nuclear thickness follows a probabilistic picture.

5. NONLINEAR NUCLEAR $k_\perp$-FACTORIZATION FOR BREAKUP OF PHOTONS INTO SEMIHARD DIJETS: LARGE-N_c APPROXIMATION

We can now relax the hardness restriction and consider semihard dijets, $|\mathbf{p}_\pm| \sim Q_A$. In this section, we give a consistent treatment of this case in the venerable large-N_c approximation. Our formulation can be called a nonlinear nuclear generalization of the $k_\perp$-factorization.

The crucial point is that in the large-N_c approximation, $\Sigma_2 = \sigma_{11} = \sigma(\mathbf{r}) + \sigma(\mathbf{r}')$, and therefore only the last term in Sylvester expansion (25) contributes to the jet–jet inclusive cross section. The nuclear distortion factor is still given by Eq. (29), but for finite Σ_2. Slightly generalizing (35) and using

$$\Sigma_1 = \sigma(\mathbf{s}) + \sigma(\mathbf{s} + \mathbf{r}' - \mathbf{r}), \tag{38}$$

we can recast the distortion factor in the form

$$\begin{aligned} &D_A(\mathbf{s}, \mathbf{r}, \mathbf{r}', \mathbf{b}) \\ &= \int_0^1 d\beta \exp\left\{-\frac{1}{2}[\beta\Sigma_1 + (1-\beta)\Sigma_2]T(\mathbf{b})\right\} \\ &= \int_0^1 d\beta \exp\left\{-\frac{1}{2}(1-\beta)[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\} \\ &\times \exp\left\{-\frac{1}{2}\beta[\sigma(\mathbf{s}) + \sigma(\mathbf{s} + \mathbf{r}' - \mathbf{r})]T(\mathbf{b})\right\}, \end{aligned} \tag{39}$$

where the different exponential factors admit a simple interpretation. The first and the second describe the intranuclear distortion of the incoming color-singlet $q\bar{q}$ and $q'\bar{q}'$ dipole state, whereas the last two factors describe the distortion of the outgoing color-octet ($q\bar{q}$) and ($q'\bar{q}'$) states. Application of the NSS representation [6] to the attenuation factors in (39) yields

$$\begin{aligned} &D_A(\mathbf{s}, \mathbf{r}, \mathbf{r}', \mathbf{b}) \\ &= \int_0^1 d\beta \int d\boldsymbol{\kappa}_1 \Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1) \exp(-i\boldsymbol{\kappa}_1 \cdot \mathbf{r}) \\ &\times \int d\boldsymbol{\kappa}_2 \Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_2) \exp(i\boldsymbol{\kappa}_2 \cdot \mathbf{r}) \\ &+ \int d\boldsymbol{\kappa}_3 \Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\kappa}_3) \exp[i\boldsymbol{\kappa}_3 \cdot (\mathbf{s} + \mathbf{r}' - \mathbf{r})] \\ &\times \int d\boldsymbol{\kappa}_4 \Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\kappa}_4) \exp(i\boldsymbol{\kappa}_4 \cdot \mathbf{r}). \end{aligned} \tag{40}$$

The integral representation in (39) furnishes two important tasks: it removes $\Sigma_1 - \Sigma_2$ from the denominator in (25) and gives the Fourier transform (40) of the nuclear distortion factor as a product of manifestly positive-definite nuclear WW gluon distributions. Finally, the jet–jet inclusive cross section takes the form

$$\begin{aligned} \frac{d\sigma_{\text{in}}}{d\mathbf{b}dzd\mathbf{p}_- d\boldsymbol{\Delta}} &= \frac{1}{2(2\pi)^2}\alpha_S\sigma_0 T(\mathbf{b}) \\ &\times \int_0^1 d\beta \int d\boldsymbol{\kappa}_1 d\boldsymbol{\kappa}_2 d\boldsymbol{\kappa}_3 d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa}) \\ &\times \Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\Delta} - \boldsymbol{\kappa}_3 - \boldsymbol{\kappa})\Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\kappa}_3) \\ &\times \Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1)\Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_2) \\ &\times \{\langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_2 + \boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_2 + \boldsymbol{\kappa}_3 + \boldsymbol{\kappa}\rangle\} \\ &\times \{\langle z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3|\gamma^*\rangle - \langle z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3 + \boldsymbol{\kappa}|\gamma^*\rangle\} \\ &= \frac{1}{2(2\pi)^2}\alpha_S\sigma_0 T(\mathbf{b}) \int_0^1 d\beta \int d\boldsymbol{\kappa}_3 d\boldsymbol{\kappa}\, f(\boldsymbol{\kappa}) \\ &\times \Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\Delta} - \boldsymbol{\kappa}_3 - \boldsymbol{\kappa})\Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\kappa}_3) \\ &\times \left|\int d\boldsymbol{\kappa}_1 \Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1)\right. \\ &\left.\times \{\langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3 + \boldsymbol{\kappa}\rangle\}\right|^2. \end{aligned} \tag{41}$$

This is our central result for the inclusive cross section of the photon breakup into dijets on nuclei. It demonstrates how the broadening of the transverse momentum distribution of dijets is uniquely calculable in terms of the collective WW glue of a nucleus and as such must be regarded as a nonlinear $k_\perp$-factorization for the inclusive dijet cross section.

The last form of (41) shows clearly that the integrand is manifestly positive-valued. Returning to (39) and (40), we can identify the convolution of the collective nuclear WW glue $\Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1)$ with the photon wave functions in the last form in (41) as an effect of distortions of the photon wave function when the $q\bar{q}$ pair propagates in the state that is still color-singlet.

We finally consider the limiting case where $|\mathbf{p}_-|$, $|\boldsymbol{\Delta}| \lesssim Q_A$. In our analysis [5] of the single particle spectrum, we discovered that the transverse momentum distribution of sea quarks is dominated by anticollinear, anti-DGLAP splitting of gluons into sea when the transverse momentum of the parent gluons is larger than the momentum of sea quarks. As stated in the

Introduction, this strongly suggests a complete azimuthal decorrelation of forward minijets with the transverse momenta below the saturation scale, $p_\pm \lesssim Q_A$. Our analysis of $f^{(j)}(\boldsymbol{\kappa})$ in Appendix C shows that for the average DIS on realistic nuclei, Q_A^2 does not exceed several $(\mathrm{Gev}/c)^2$, and hence this regime is a somewhat academic one (see Section 6, however). We nevertheless assume that Q_A is so large that jets with $p_\pm \lesssim Q_A$ are measurable.

We note that $|\boldsymbol{\kappa}_i| \sim Q_A$, and we can therefore neglect $\mathbf{p}_-$ in the photon wave functions and the decorrelation momentum $\boldsymbol{\Delta}$ in the argument of $\Phi(\beta\nu_A(\mathbf{b}), \boldsymbol{\Delta} - \boldsymbol{\kappa}_3 - \boldsymbol{\kappa})$. The approximation

$$\Big|\int d\boldsymbol{\kappa}_1 \Phi((1-\beta)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1)\ \{\langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1 + \boldsymbol{\kappa}_3 + \boldsymbol{\kappa}\rangle\}\Big|^2 \approx \big|\langle\gamma^*|z, \boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \boldsymbol{\kappa}_3 + \boldsymbol{\kappa}\rangle\big|^2 \tag{42}$$

is then justified in (41). The principal point is that the minijet–minijet inclusive cross section is independent of either the minijet or the decorrelation momentum, which proves the disappearance of the azimuthal decorrelation of minijets with the transverse momentum below the saturation scale.

6. AZIMUTHAL DECORRELATION OF DIJETS IN DIS OFF NUCLEI: NUMERICAL ESTIMATES

The azimuthal decorrelation of two jets is quantified by the mean transverse acoplanarity momentum squared $\langle\boldsymbol{\Delta}_\perp^2(\mathbf{b})\rangle$, where $\boldsymbol{\Delta}_\perp$ is transverse to the axis of the jet with the higher momentum (Fig. 3). Here, we present numerical estimates for hard dijets, $|\mathbf{p}_+| \gg Q_A$. The convolution property of hard dijet cross section (35) suggests that

$$\langle\boldsymbol{\Delta}_\perp^2(\mathbf{b})\rangle_A = \left\{\int_{\mathscr{C}} d\mathbf{p}_- \boldsymbol{\Delta}_\perp^2 \frac{d\sigma_N}{dz\,d\mathbf{p}_+\,d\mathbf{p}_-}\right\} \times \left\{\int_{\mathscr{C}} d\mathbf{p}_- \frac{d\sigma_N}{dz\,d\mathbf{p}_+\,d\mathbf{p}_-}\right\}^{-1} \approx \langle\boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A + \langle\boldsymbol{\Delta}_\perp^2\rangle_N, \tag{43}$$

where $\langle\boldsymbol{\Delta}_\perp^2\rangle_N$ refers to DIS on a free nucleon and

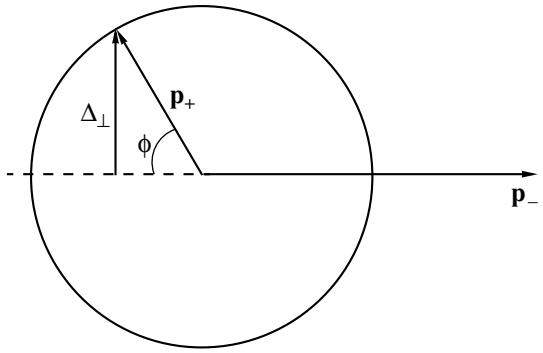

Fig. 3. The definition of the dijet configurations considered and of the transverse component of the acoplanarity momentum $\boldsymbol{\Delta}_\perp$.

$\langle\boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A$ is the nuclear broadening term:

$$\langle\boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A = \left\{\int_{\mathscr{C}} d\boldsymbol{\kappa}\boldsymbol{\kappa}_\perp^2 \Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\kappa})\right\} \times \left\{\int_{\mathscr{C}} d\boldsymbol{\kappa}\Phi(2\beta\lambda_c\nu_A(\mathbf{b}), \boldsymbol{\kappa})\right\}^{-1}. \tag{44}$$

The sign "≈" in (43) reflects the kinematical limitations $\mathscr{C}$ on $\mathbf{p}_-$ and $\boldsymbol{\kappa}$ in the practical evaluation of the acoplanarity distribution. In a typical final state shown in Fig. 3, it is the harder jet with the larger transverse momentum that defines the jet axis and the acoplanarity momentum $\boldsymbol{\Delta}$ is defined in terms of components of the momentum of the softer jet with respect to that axis (see, e.g., [20]). For definiteness, we present numerical estimates for the Gedanken experiment in which we classify an event as a dijet if the quark and antiquark are produced in different hemispheres; i.e., if the azimuthal angle π–ϕ between the two jets is below $\pi/2$, the quark jet has fixed $|\mathbf{p}_+|$, and the antiquark jet has a higher transverse momentum, $|\mathbf{p}_+| \lesssim |\mathbf{p}_-| \lesssim 10|\mathbf{p}_+|$ (in the discussion of the experimental data, one often refers to the higher momentum jet as the trigger jet and the softer jet as the away jet [20]).

The free-nucleon quantity $\langle\boldsymbol{\Delta}_\perp^2\rangle_N$ is evaluated from Eq. (43) with free nucleon cross section (7). For evaluation purposes, we can start with the small-$\boldsymbol{\Delta}$ expansion for excitation of hard $\mathbf{p}_+^2 \gg \varepsilon^2 = z(1-z)Q^2$, light flavor dijets from transverse photons,

$$\frac{d\sigma_N}{dz\,d\mathbf{p}_+\,d\boldsymbol{\Delta}} \approx \frac{1}{\pi} e_f^2 \alpha_{em}\alpha_S(\mathbf{p}_+^2)[z^2 + (1-z)^2] \times \frac{1}{\Delta^4}\frac{\partial G(x, \Delta^2)}{\partial \ln(\Delta^2)}\frac{\Delta^2}{(\varepsilon^2 + \mathbf{p}_+^2)(\varepsilon^2 + \mathbf{p}_+^2 + \Delta^2)}. \tag{45}$$

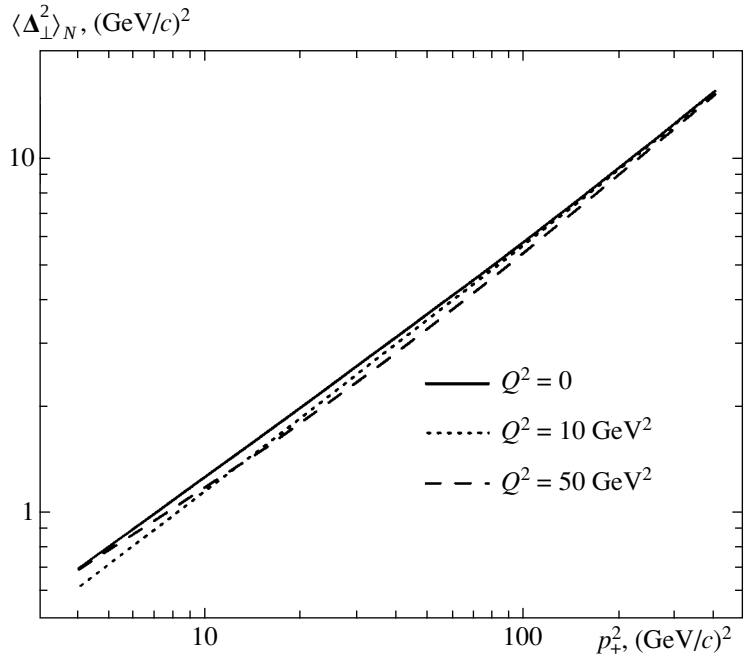

Fig. 4. The mean acoplanarity momentum squared $\langle \Delta_\perp^2 \rangle_N$ for DIS off a free nucleon target with production of trigger jets with a transverse momentum higher than $\mathbf{p}_+$ for several values of Q^2. The numerical results are for $x = 0.01$, and the input unintegrated gluon structure of the proton is taken from [19].

The form of the last factor in (45) only mimics its leveling off at $\Delta^2 \gtrsim \mathbf{p}_+^2$, see Eq. (7). In the denominator of (43), we then find the typical logarithmic integral

$$\frac{1}{\pi}\int_0^\pi d\phi \int^{\mathbf{p}_+^2} \frac{d\Delta^2}{\Delta^2}\frac{\partial G(x, \Delta^2)}{\partial \ln \Delta^2} = G(x, \mathbf{p}_+^2), \qquad (46)$$

to be compared with the numerator of the form

$$\frac{1}{\pi}\int_0^\pi d\phi \sin^2\phi \int^{\mathbf{p}_+^2} d\Delta^2 \frac{\partial G(x, \Delta^2)}{\partial \ln \Delta^2} \sim \frac{1}{2}\mathbf{p}_+^2 \mathcal{F}(x, \mathbf{p}_+^2). \qquad (47)$$

More accurate numerical estimates for the selection criteria of our Gedanken experiment suggest a numerical factor of ≈0.7 in (47); the expression

$$\langle \Delta_\perp^2 \rangle_N = \left\{ \int_{p_+} d\mathbf{p}_- \Delta_\perp^2 \frac{d\sigma_N}{dz d\mathbf{p}_+ d\mathbf{p}_-} \right\} \times \left\{ \int_{p_+} d\mathbf{p}_- \frac{d\sigma_N}{dz d\mathbf{p}_+ d\mathbf{p}_-} \right\}^{-1} \approx 0.7 \frac{\mathcal{F}(x, \mathbf{p}_+^2)}{G(x, \mathbf{p}_+^2)} \mathbf{p}_+^2 \qquad (48)$$

correctly describes the numerical results shown in Fig. 4. As far as the dijets are hard, $\mathbf{p}_+^2 \gtrsim z(1-z)Q^2 \sim \frac{1}{4}Q^2$, the acoplanarity momentum distribution is independent of Q^2, which holds even better if we consider $\sigma_T + \sigma_L$. This point is illustrated in Fig. 4, where we show $\langle \Delta_\perp^2 \rangle_N$ at $z = 1/2$ for several values of Q^2. Because of this weak dependence on Q^2, we make no distinction between DIS and real photoproduction, $Q^2 = 0$, in what follows.

In practical evaluations of the nuclear contribution $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$, we can use the explicit expansion

$$\int_0^1 d\beta \Phi(2\beta\lambda_c \nu_A(\mathbf{b}), \boldsymbol{\kappa}) = \sum_{j=0}^\infty w_A(\mathbf{b}, j) f^{(j)}(\boldsymbol{\kappa}) = \sum_{j=0}^\infty \frac{1}{j!}\frac{\gamma(j+1, 2\lambda_c \nu_A(\mathbf{b}))}{2\lambda_c \nu_A(\mathbf{b})} f^{(j)}(\boldsymbol{\kappa}), \qquad (49)$$

where

$$\gamma(j, x) = \int_0^x dy y^{j-1} e^{-y}$$

is the incomplete gamma-function. The properties of the collective glue for j overlapping nucleons, $f^{(j)}(\boldsymbol{\kappa})$, are presented in Appendix C. For a heavy nucleus, Eq. (49) can be approximated by its integrand at $\beta \approx$

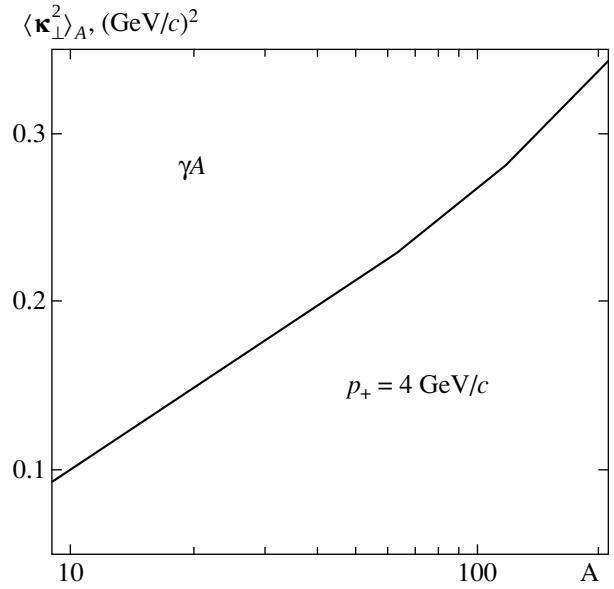

Fig. 5. The atomic mass number dependence of the nuclear broadening contribution, $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$, to the mean acoplanarity momentum squared for real photoproduction off nuclei at $x = 0.01$. The input unintegrated gluon SF of the proton is taken from [19].

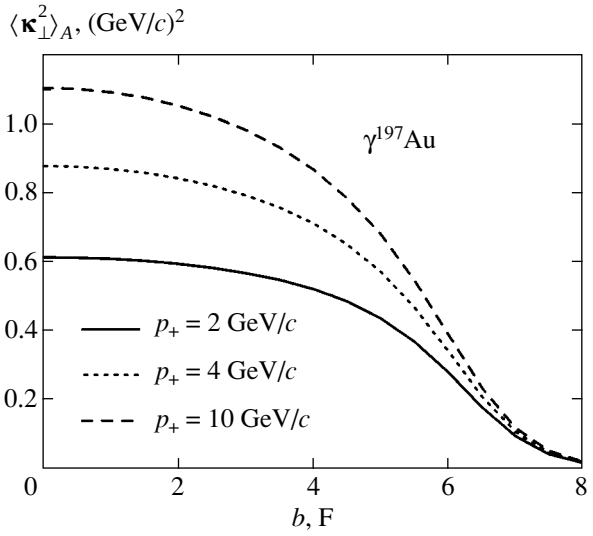

Fig. 6. The impact parameter dependence of the nuclear broadening contribution, $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$, to the mean acoplanarity momentum squared from peripheral DIS at a large impact parameter to the central DIS at $\mathbf{b} = 0$ for several values of the away jet momentum p_+. The numerical results are for $x = 0.01$, and the input unintegrated gluon SF of the proton is taken from [19]

1/2, i.e., by $\Phi(\lambda_c \nu_A(\mathbf{b}), \boldsymbol{\kappa})$. A slightly more accurate evaluation of the numerically important no-broadening contribution from $j = 0$ gives

$$\int_0^1 d\beta \Phi(2\beta\lambda_c \nu_A(\mathbf{b}), \boldsymbol{\kappa}) \approx w_A(\mathbf{b}, 0)\delta(\boldsymbol{\kappa}) + (1 - w_A(\mathbf{b}, 0))\frac{1}{\pi}\frac{\lambda_c Q_A^2(\mathbf{b})}{(\boldsymbol{\kappa}^2 + \lambda_c Q_A^2(\mathbf{b}))^2}, \quad (50)$$

where Q_A^2 is given by Eq. (108) and

$$w_A(\mathbf{b}, 0) = \frac{1 - \exp[-\nu_A(\mathbf{b})]}{\nu_A(\mathbf{b})} \quad (51)$$

is the probability of the no-broadening contribution, which is still substantial for realistic nuclei. In our Gedanken experiment, $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$ must be evaluated over the constrained phase space $\mathcal{C}$, $\kappa_\perp \le |\mathbf{p}_+|$ and $\kappa_L > 0$, and analytic parameterization (50) gives

$$\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A \approx \lambda_c Q_A^2(\mathbf{b}) \times \left[\ln \tan\left(\frac{\pi}{4} + \frac{1}{2}\arctan\frac{p_+}{\sqrt{\lambda_c} Q_A(\mathbf{b})}\right) - \frac{p_+}{\sqrt{\lambda_c Q_A^2(\mathbf{b}) + p_+^2}}\right] \times \frac{(1 - w_A(\mathbf{b}, 0))\sqrt{\lambda_c Q_A^2(\mathbf{b}) + p_+^2}}{w_A(\mathbf{b}, 0)\sqrt{\lambda_c Q_A^2(\mathbf{b}) + p_+^2} + (1 - w_A(\mathbf{b}, 0))p_+}. \quad (52)$$

We recall that (43) and (52) must only be used for $|\mathbf{p}_+| \gg Q_A(\mathbf{b})$.

For the average DIS off heavy nuclei, the reference value is $\langle Q_{\mathrm{Au}}^2(\mathbf{b}) \rangle = 0.9$ (GeV/c)^2 (see Appendix C). The atomic mass number dependence of the nuclear broadening $\langle \boldsymbol{\kappa}_\perp^2 \rangle_A$ for jets with $p_+ = 4$ GeV/c in the average DIS off nucleus is shown in Fig. 5. The principal reason for $\langle \boldsymbol{\kappa}_\perp^2 \rangle_A$ being numerically small compared to $\langle Q_{\mathrm{Au}}^2(\mathbf{b}) \rangle$ is that even for such a heavy nucleus as ^{197}Au, the no-broadening probability in the average DIS is large, $\langle w_{\mathrm{Au}}(\mathbf{b}, 0) \rangle \approx 0.5$. Comparison of the free nucleon broadening $\langle \boldsymbol{\Delta}_\perp^2 \rangle_N$ in Fig. 4 with the nuclear contribution $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$ in Fig. 5 shows that the nuclear mass number dependence of the azimuthal decorrelation of dijets in the average DIS off nuclei is relatively weak.

However, nuclear broadening is substantially stronger for a subsample of central DIS events at $\mathbf{b} \sim 0$. In Fig. 6, we show the dependence of the β-averaged nuclear broadening $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$ on the impact parameter at several values of p_+ for the gold—^{197}Au—target. There are two related sources of the p_+ dependence of $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b}) \rangle_A$. First, because r, $r' \sim 1/p_+$ for hard dijets, the strong coupling enters Eqs. (33) and (108) as $\alpha_S(\mathbf{p}_+^2)$. Then, for hard jets, $\nu_A(\mathbf{b}) \propto \alpha_S(\mathbf{p}_+^2)$ and $w_A(\mathbf{b} = 0, 0)$

rises substantially with p_+ in the region p_+ of practical interest, $1 \lesssim p_+ \lesssim$ 5–10 GeV/c, where the strong coupling varies rapidly. For a nucleus with a mass number of $A = 200$, it rises from $w_A(\mathbf{b} = 0, 0) \approx 0.12$ at $p_+ =$ 2 GeV/c to ≈0.20 at 4 GeV/c, and to ≈0.25 at $p_+ =$ 10 GeV/c (see [39] for the nuclear density parameterization). Second, for the same reason that $\nu_A(\mathbf{b}) \propto \alpha_S(\mathbf{p}_+^2)$, the contribution from large j in (49), and hence $Q_A^2(\mathbf{b})$, diminishes gradually with rising p_+, proportionally to $\alpha_S(\mathbf{p}_+^2)/\alpha_S(Q_A^2)$. In the region $p_+ \lesssim 10$ GeV/c of practical interest, we find that $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A \sim Q_A^2(\mathbf{b})$.

We now compare the numerical results in Figs. 5 and 6 for $p_+ = 4$ GeV/c and the ^{197}Au target. According to Eq. (109) in Appendix C,

$$Q_A^2(0) = \left(\frac{4}{3} - 2\right)\langle Q_A^2(\mathbf{b})\rangle. \tag{53}$$

The no-broadening probability $w_A(\mathbf{b} = 0, 0) \approx 0.20$ for central DIS is substantially smaller than $\langle w_{\mathrm{Au}}(\mathbf{b}, 0)\rangle \approx 0.5$ for average DIS. In conjunction with (53), this entails an enhancement of $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A$ by a factor 2.5–3 from the average to central DIS. The same point is illustrated by the expectation value of j in (49) for the Au target: for jets with $p_+ = 4$ GeV/c, it decreases by a factor of about 3 from $\langle j(\mathbf{b} = 0)\rangle = 2.86$ to $\langle j\rangle_A = 0.87$ from the central to average DIS.

One can enhance Q_A^2 and the nuclear contribution $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A$ even further by selecting the DIS events where the photon breaks up into a $q\bar{q}$ pair on the front face of the nucleus, which in the language of (36) corresponds to the contribution from $\beta \longrightarrow 1$ [see the discussion of (49)]. Experimentally, precisely such events are isolated by selecting very large multiplicities or very high transverse energies of the secondary particles produced (see [20] and references therein). Equation (36) then shows (also see the discussion of the $\beta \approx 1/2$ approximation in (49)) that for a very high multiplicity central DIS off the Au nucleus, $Q_A^2 \approx 2.5$ GeV^2 is quite feasible. Equation (52) shows that for such a large $Q_A^2 \approx 2.5$ GeV^2 and $p_+ =$ 5–10 GeV/c of practical interest, $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b} = 0)\rangle$ grows slower rather than proportionally to Q_A^2 and therefore the value of $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b} = 0)\rangle$ for a high-multiplicity central DIS off Au nucleus is enhanced by a factor of 4–5 from $\langle \boldsymbol{\kappa}_\perp^2\rangle_{\mathrm{Au}}$ for an average DIS.

We have an overall good understanding of the gross features of nuclear azimuthal decorrelations in DIS off nuclei. We now comment on the recent finding by the STAR Collaboration of the disappearance of a back-to-back high-$p_\perp$ hadron correlation occurring in passing from peripheral to central gold–gold collisions at RHIC [20]. Our experience with application of the color dipole formalism to hard hadron–nucleus interactions [17] suggests that our analysis of acoplanarity of forward hard jets can be generalized to mid-rapidity jets. This only requires choosing an appropriate system of dipoles, for instance, the open heavy flavor production can be treated in terms of the intranuclear propagation of the gluon–quark–antiquark system in the overall color-singlet state. At RHIC energies, jets with moderately large $p_\perp$ are mostly due to gluon–gluon collisions. In our language, this can be treated as a breakup of gluons into dijets, and azimuthal decorrelation of hard jets must be discussed in terms of intranuclear propagation of color-octet gluon–gluon dipoles. For such gluon–gluon dipoles, the relevant saturation scale Q_{8A}^2 is larger than that for the quark–antiquark dipoles by a factor of $2\lambda_c = C_A/C_F = 9/4$ [24]. Arguably, distortions in the target and projectile nuclei add up in central nucleus–nucleus collisions and the effective thickness of nuclear matter is about twice that in DIS. The results shown in Fig. 5 then suggest that for central gold–gold collisions, the nuclear broadening of gluon–gluon dijets could be quite substantial, $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b} = 0)\rangle_{\mathrm{AuAu}} \sim$ 3–4 (GeV/c)^2 for average central Au–Au collisions and even twice as large if collisions occur at the front surface of colliding nuclei.

The principal effect of nuclear broadening is a reduction of the probability of observing back-to-back jets,

$$P(b) \propto \frac{\langle \boldsymbol{\Delta}_\perp^2\rangle_N}{\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A + \langle \boldsymbol{\Delta}_\perp^2\rangle_N}, \tag{54}$$

where $\langle \boldsymbol{\Delta}_\perp^2\rangle_N$ is to be compared to $\langle \boldsymbol{\kappa}_\perp^2(\mathbf{b})\rangle_A$. Equation (48) for the free nucleon case also holds for gluon–gluon collisions. The results shown in Fig. 3 then entail that $\langle \boldsymbol{\Delta}_\perp^2\rangle_N \approx \langle \boldsymbol{\kappa}_\perp^2(0)\rangle_{\mathrm{AuAu}} \sim$ 3–4 (GeV/c)^2 at the jet momentum $p_+ = p_J =$ 6–8 GeV/c and our nuclear broadening becomes substantial for all jets with p_+ below the decorrelation threshold momentum p_J. In practice, the STAR Collaboration studied the azimuthal correlation of two high-$p_\perp$ hadrons; for a quantitative correspondence between the STAR observable and the azimuthal decorrelation in the parent dijet, one must model the fragmentation of jets into hadrons (see [35] for the modern fragmentation schemes). We note here that the cutoff p_+ in our Gedanken experiment is related to the momentum cutoff $p_{T,\min}$ of the associated tracks from the away jet, whereas our jet with the momentum $\mathbf{p}_-$ can be regarded as a counterpart of the STAR trigger jet.

The STAR cutoff $p_T = 2$ GeV/c corresponds to parent jets with transverse momentum $p_+ \sim (2\text{–}3)p_T = 4\text{–}6$ GeV/c, which is comparable to, or even smaller than the decorrelation threshold momentum $p_J = 6\text{–}8$ GeV/c. Equation (54) then suggests that in the STAR kinematics, the probability of observing back-to-back away and trigger jets approximately reduces to half, and perhaps even more strongly, from peripheral-to-central Au–Au collisions, and our azimuthal decorrelation may therefore substantially contribute to the STAR effect.

In practical consideration of azimuthal decorrelations in central heavy ion collisions, the above distortions of the jet–jet inclusive spectrum produced due to interactions with the nucleons of the target and projectile ions must be complemented by rescatterings of the parent high-$p_\perp$ partons on the abundantly produced secondary hadrons. Our nuclear decorrelation effect must be predominant and reinteractions with secondary particles must be marginal in pA collisions, where we expect $\langle \boldsymbol{\kappa}_\perp^2(0)\rangle_{p\mathrm{Au}} \approx 1.5$ (GeV/c)^2 for central collisions and even $\langle \boldsymbol{\kappa}_\perp^2(0)\rangle_{p\mathrm{Au}} \approx 3$ (GeV/c)^2 is feasible for central collisions in the regime of $\beta \longrightarrow 1$, i.e., with the limiting high multiplicity.

7. NUCLEAR $k_\perp$-FACTORIZATION FOR $1/N_c^2$ CORRECTIONS TO THE PHOTON BREAKUP

Having established nuclear $k_\perp$-factorization properties of the dijet cross section to the leading order of the large-N_c approximation, we turn to the $1/N_c^2$-corrections and demonstrate that with one simple exception, the $1/N_c^2$-expansion can be regarded as the higher twist expansion. The two sources of the $1/N_c^2$-corrections to the nuclear distortion factor are higher-order terms in the off-diagonal σ_{18} matrix element and the terms proportional to $1/(N_c^2-1)$ in σ_{88}, Eq. (21). We note that σ_{88} can be decomposed as

$$\begin{aligned}\sigma_{88} &= \sigma(\mathbf{s}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}') + \frac{2\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')}{N_c^2-1}\\ &+ \frac{\sigma(\mathbf{s}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}') - \sigma(\mathbf{r}) - \sigma(\mathbf{r}')}{N_c^2-1}\\ &= \frac{N_c^2}{N_c^2-1}[\sigma(\mathbf{s}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]\\ &+ \frac{2\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')}{N_c^2-1} - \frac{\Delta\Sigma_{88}(\mathbf{r},\mathbf{r}')}{N_c^2-1},\end{aligned} \tag{55}$$

where

$$\Delta\Sigma_{88}(\mathbf{r},\mathbf{r}') = \sigma(\mathbf{r}) + \sigma(\mathbf{r}'), \tag{56}$$

and we have exactly reabsorbed one part of the $1/N_c^2$-correction into the leading large-N_c term of σ_{88} by scaling it with the color factor λ_c.

After some algebra, we find

$$\begin{aligned}&\langle 11|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle\\ &= \exp\left\{-\frac{1}{2}[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\}\\ &+ \frac{\Sigma_{18}^2(\mathbf{s},\mathbf{r},\mathbf{r}')T^2(\mathbf{b})}{4(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1\\ &\times \exp\left\{-\frac{1}{2}(1-\beta+\beta_1)[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\}\\ &\times \exp\left\{-\frac{1}{2}(\beta-\beta_1)[\sigma(\mathbf{s}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}.\end{aligned} \tag{57}$$

The first term in (57) is canceled by the subtraction of coherent diffractive term (22) in (8) and (25), and therefore only the subleading term in (57), proportional to $1/(N_c^2-1)$, contributes to the dijet cross section. Evaluation of corrections to the leading term of the Sylvester expansion is somewhat more complicated,

$$\begin{aligned}&\sqrt{N_c^2-1}\langle 88|S_{4A}(\mathbf{b}'_+, \mathbf{b}'_-, \mathbf{b}_+, \mathbf{b}_-)|11\rangle\\ &= \frac{1}{2}\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')T(\mathbf{b})\\ &\times\left[\int_0^1 d\beta \exp\left\{-\frac{1}{2}\beta[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\}\right.\\ &\times \exp\left\{-\frac{1}{2}(1-\beta)\sigma_{88}T(\mathbf{b})\right\}\\ &+ \frac{\Sigma_{18}^2(\mathbf{s},\mathbf{r},\mathbf{r}')T^2(\mathbf{b})}{4(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1\int_0^{\beta_1} d\beta_2\\ &\times \exp\left\{-\frac{1}{2}(\beta-\beta_1+\beta_2)[\sigma(\mathbf{r}) + \sigma(\mathbf{r}')]T(\mathbf{b})\right\}\\ &\left.\times \exp\left\{-\frac{1}{2}(1-\beta+\beta_1-\beta_2)[\sigma(\mathbf{s}) + \sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}\right].\end{aligned} \tag{58}$$

The first term in (58) contains the attenuation factor, where σ_{88} is still the exact diagonal matrix element, and we must isolate the leading term and the $1/(N_c^2-1)$-correction,

$$\exp\left\{-\frac{1}{2}(1-\beta)\sigma_{88}T(\mathbf{b})\right\}$$

$$= \exp\left\{-\frac{1}{2}(1-\beta)\lambda_c[\sigma(\mathbf{s})+\sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\} \times\left\{1-\frac{(1-\beta)\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')T(\mathbf{b})}{N_c^2-1} + \frac{(1-\beta)\Delta\Sigma_{88}(\mathbf{r},\mathbf{r}')T(\mathbf{b})}{2(N_c^2-1)}\right\}. \quad (59)$$

The fundamental reason for which the different components of the second term, proportional to $1/(N_c^2-1)$, in Eq. (21) are treated differently is that the NSS representation [6] with a positive-valued Fourier transform holds only for attenuating exponentials of the dipole cross section. The related expansion for the rising exponential $\exp\left[\frac{1}{2}\sigma(\mathbf{r})T(\mathbf{b})\right]$ can easily be written, but its Fourier transform is a sign-oscillating expansion,

$$\exp\left[\frac{1}{2}\sigma(\mathbf{s})T(\mathbf{b})\right] = \exp[2\nu_A(\mathbf{b})] \times\sum_{j=0}^{\infty}(-1)^j w_j(\nu_A(\mathbf{b}))\int d\boldsymbol{\kappa} f^{(j)}(\boldsymbol{\kappa})\exp(i\boldsymbol{\kappa}\cdot\mathbf{s}). \quad (60)$$

Therefore, combining the two exponentials with similar exponents proportional to $[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]$ in the first term of (57) is not guaranteed, because the sign of the exponent changes from attenuation to growth in the course of the β integration,

$$\beta[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]-\frac{1}{N_c^2-1}(1-\beta)\Delta\Sigma_{88}(\mathbf{r},\mathbf{r}') = \frac{N_c^2\beta-1}{N_c^2-1}[\sigma(\mathbf{r})+\sigma(\mathbf{r}')], \quad (61)$$

and it is advisable to work with the perturbative expansion in (59).

The final result for the nuclear absorption factor to an accuracy of $1/(N_c^2-1)$ is given by

$$D_A(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) = D_A^{(1)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) + D_A^{(2)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) + D_A^{(3)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) + D_A^{(4)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) + D_A^{(5)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}),$$

where

$$D_A^{(1)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) = \frac{\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')T(\mathbf{b})}{2(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1 \times\exp\left\{-\frac{1}{2}(1-\beta+\beta_1)[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right\} \times\exp\left\{-\frac{1}{2}(\beta-\beta_1)[\sigma(\mathbf{s})+\sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}, \quad (62)$$

$$D_A^{(2)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) = \frac{\Sigma_{18}^2(\mathbf{s},\mathbf{r},\mathbf{r}')T^2(\mathbf{b})}{4(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1 \times\int_0^{\beta_1} d\beta_2\exp\left\{-\frac{1}{2}(\beta-\beta_1+\beta_2)[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right\} \times\exp\left\{-\frac{1}{2}(1-\beta+\beta_1-\beta_2)[\sigma(\mathbf{s})+\sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}, \quad (63)$$

$$D_A^{(3)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) = -\frac{\Sigma_{18}(\mathbf{s},\mathbf{r},\mathbf{r}')T(\mathbf{b})}{N_c^2-1}\int_0^1 d\beta(1-\beta) \times\exp\left\{-\frac{1}{2}\beta[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right\} \times\exp\left\{-\frac{1}{2}(1-\beta)\lambda_c[\sigma(\mathbf{s})+\sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}, \quad (64)$$

$$D_A^{(4)}(\mathbf{s},\mathbf{r},\mathbf{r}',\mathbf{b}) = \frac{\Delta\Sigma_{18}(\mathbf{r},\mathbf{r}')T(\mathbf{b})}{2(N_c^2-1)}\int_0^1 d\beta(1-\beta) \times\exp\left\{-\frac{1}{2}\beta[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right\} \times\exp\left\{-\frac{1}{2}(1-\beta)\lambda_c[\sigma(\mathbf{s})+\sigma(\mathbf{s}-\mathbf{r}+\mathbf{r}')]T(\mathbf{b})\right\}, \quad (65)$$

$$D_A^{(5)}(\mathbf{s}, \mathbf{r}, \mathbf{r}', \mathbf{b}) = \int_0^1 d\beta \exp\left\{-\frac{1}{2}(1-\beta)[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right\} \times \exp\left\{-\frac{1}{2}\beta[\sigma(\mathbf{s})+\sigma(\mathbf{s}+\mathbf{r}'-\mathbf{r})]T(\mathbf{b})\right\}. \quad (66)$$

Equation (66) is the leading large-N_c result, Eqs. (62) and (63) describe contributions to the dijet cross section of the second and third order in the off-diagonal matrix element σ_{18}, and Eqs. (64) and (65) come from expansion (59).

As an illustration of salient features of the $1/(N_c^2-1)$-corrections, we expose the contribution from the first term (62) in detail. Following the considerations in Sections 4 and 5, we readily obtain

$$\frac{d\Delta\sigma_{\rm in}^{(1)}}{d\mathbf{b}dzd\mathbf{p}_-d\mathbf{\Delta}} = \frac{\alpha_S^2\sigma_0^2T^2(\mathbf{b})}{4(2\pi)^2(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1 \times\int d\mathbf{q}_1 d\mathbf{q}_2 d\boldsymbol{\kappa}_3 f(\mathbf{q}_1) f(\mathbf{q}_2) \times\Phi((\beta-\beta_1)\nu_A(\mathbf{b}), \mathbf{\Delta}-\boldsymbol{\kappa}_3-\mathbf{q}_1-\mathbf{q}_2) \times\Phi((\beta-\beta_1)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_3) \times\left|\int d\boldsymbol{\kappa}_1\Phi((1-\beta+\beta_1)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1)\right. \times\{\langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_1\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_2\rangle + \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2\rangle\}\left.\right|^2. \quad (67)$$

Of particular interest is the large-$|\mathbf{p}_-|$ behavior of (67). We note that for $\mathbf{p}_-^2 \gg Q_A^2(\mathbf{b})$, we can neglect $\kappa_{1,3}$ in the argument of the photon wave function, and hence

$$\int d\boldsymbol{\kappa}_1\Phi((1-\beta+\beta_1)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1) \times\{\langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_1\rangle - \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_2\rangle + \langle\gamma^*|z, \mathbf{p}_- + \boldsymbol{\kappa}_1+\boldsymbol{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2\rangle\} \approx \{\langle\gamma^*|z, \mathbf{p}_-\rangle - \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_1\rangle - \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_2\rangle + \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_1+\mathbf{q}_2\rangle\}, \quad (68)$$

where we used the normalization property

$$\int d\boldsymbol{\kappa}_1\Phi((1-\beta+\beta_1)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_1) = 1.$$

Next, we can readily verify that

$$\int d\boldsymbol{\kappa}_3\Phi((\beta-\beta_1)\nu_A(\mathbf{b}), \mathbf{\Delta}-\boldsymbol{\kappa}_3-\mathbf{q}_1-\mathbf{q}_2) \times\Phi((\beta-\beta_1)\nu_A(\mathbf{b}), \boldsymbol{\kappa}_3) = \Phi((\beta-\beta_1)\nu_A(\mathbf{b}), \mathbf{\Delta}-\mathbf{q}_1-\mathbf{q}_2). \quad (69)$$

Incidentally, by a similar analysis of the onset of the high-$\mathbf{p}_+$ limit, one would obtain the linear nuclear $k_\perp$-factorization (36) for hard dijets from the nonlinear nuclear $k_\perp$-factorization (41).

The combination of the photon wave functions in (68) corresponds to the second finite difference in $\mathbf{q}_1$ and $\mathbf{q}_2$, and therefore, for jets with $\mathbf{p}_-^2 \gg \varepsilon^2$, we have the estimate

$$\left|\langle\gamma^*|z, \mathbf{p}_-\rangle - \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_1\rangle + \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_2\rangle - \langle\gamma^*|z, \mathbf{p}_- + \mathbf{q}_1+\mathbf{q}_2\rangle\right|^2 \approx \left|\langle\gamma^*|z, \mathbf{p}_-\rangle\right|^2\frac{\mathbf{q}_1^2\mathbf{q}_2^2}{(\mathbf{p}_-^2)^2}, \quad (70)$$

which shows that the contribution to the dijet cross section from terms of the second order in σ_{18}^2 is the higher twist correction. Compared to the leading large-N_c cross section, it contains extra $\int d\mathbf{q}_2\mathbf{q}_2^2 f(\mathbf{q}_2)$ and an extra power of $\alpha_S\sigma_0 T(\mathbf{b})$, which combine to precisely the dimensional nuclear saturation scale $Q_A^2(\mathbf{b})$, see Eq. (52), such that the resulting suppression factor is

$$\frac{d\Delta\sigma_{\rm in}^{(1)}}{d\sigma_{\rm in}} \sim \frac{1}{(N_c^2-1)}\frac{Q_A^2(\mathbf{b})}{\mathbf{p}_-^2}. \quad (71)$$

As far as the expansion in higher inverse powers of the hard scale $\mathbf{p}_-^2$ is concerned, $\Delta\sigma_{\rm in}^{(1)}$ has the form of a higher twist correction. In the retrospect, we observe that the principal approximation (68) in the above derivation for hard dijets amounts to setting $|\mathbf{r}|, |\mathbf{r}'| \ll |\mathbf{s}|$ in the attenuation factors in the β, β_1 integrand in (62). However, the exact $\mathbf{r}, \mathbf{r}'$-dependence must be retained in the prefactor $\Sigma_{18}(\mathbf{s}, \mathbf{r}, \mathbf{r}')$, because it vanishes if either $\mathbf{r} = 0$ or $\mathbf{r}' = 0$. It is precisely the latter property that provides the finite-difference structure of the combination of the photon wave functions in (67) and (68) and is behind higher twist property (71) of the $1/(N_c^2-1)$-correction.

The second term, Eq. (63), gives the correction

$$\frac{d\Delta\sigma_{\rm in}^{(2)}}{d\mathbf{b}dzd\mathbf{p}_-d\mathbf{\Delta}} = \frac{\alpha_S^3\sigma_0^3T^3(\mathbf{b})}{8(2\pi)^2(N_c^2-1)}\int_0^1 d\beta\int_0^\beta d\beta_1$$

$$\times\int_0^{\beta_1} d\beta_2\int d\mathbf{q}_1 d\mathbf{q}_2 d\mathbf{q}_3 d\mathbf{\kappa}_3 f(\mathbf{q}_1) f(\mathbf{q}_2) f(\mathbf{q}_3)$$

$$\times\Phi((1-\beta+\beta_1-\beta_2)\nu_A(\mathbf{b}), \mathbf{\Delta}-\mathbf{\kappa}_3-\mathbf{q}_1-\mathbf{q}_2-\mathbf{q}_3)$$

$$\times\Phi((1-\beta+\beta_1-\beta_2)\nu_A(\mathbf{b}), \mathbf{\kappa}_3)$$

$$\times\left|\int d\mathbf{\kappa}_1\Phi((\beta-\beta_1+\beta_2)\nu_A(\mathbf{b}), \mathbf{\kappa}_1)\right|$$

$$\times\{\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_2\rangle$$

$$+\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_3\rangle$$

$$+\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_3+\mathbf{q}_1\rangle$$

$$+\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_3+\mathbf{q}_2\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3|\gamma^*\rangle\}\Big|^2. \quad (72)$$

The combination of the photon wave functions in (72) corresponds to the third finite derivative in $\mathbf{q}_{1,2,3}$. Starting from (72), we can readily repeat the analysis that leads to an estimate (71). Alternatively, we can take the simplified form of the attenuation factors, as explained below Eq. (71). Either way, we find that the contribution from third-order terms in σ_{18} is of an even higher twist and has the smallness

$$\frac{d\Delta\sigma_{\rm in}^{(2)}}{d\sigma_{\rm in}} \sim \frac{1}{N_c^2-1}\left(\frac{Q_A^2(\mathbf{b})}{\mathbf{p}_-^2}\right)^2. \quad (73)$$

Apart from a slight difference in the structure of the β integrations, correction (64) does not differ from $d\Delta\sigma^{(1)}$ in Eq. (68),

$$\frac{d\Delta\sigma_{\rm in}^{(3)}}{d\mathbf{b}dzd\mathbf{p}_-d\mathbf{\Delta}} = -\frac{\alpha_S^2\sigma_0^2T^2(\mathbf{b})}{4(2\pi)^2(N_c^2-1)}$$

$$\times\int_0^1 d\beta(1-\beta)\int d\mathbf{q}_1 d\mathbf{q}_2 d\mathbf{\kappa}_3 f(\mathbf{q}_1) f(\mathbf{q}_2)$$

$$\times\Phi((1-\beta)\lambda_c\nu_A(\mathbf{b}), \mathbf{\Delta}-\mathbf{\kappa}_3-\mathbf{q}_1-\mathbf{q}_2)$$

$$\times\Phi((1-\beta)\lambda_c\nu_A(\mathbf{b}), \mathbf{\kappa}_3)\left|\int d\mathbf{\kappa}_1\Phi(\beta\nu_A(\mathbf{b}), \mathbf{\kappa}_1)\right| \quad (74)$$

$$\times\{\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_2\rangle$$

$$+\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2\rangle\}\Big|^2.$$

Consequently, the same estimate (71) is also valid for $d\Delta\sigma^{(3)}$.

The correction $d\Delta\sigma^{(4)}$ requires a bit more scrutiny. It contains a product of the first and second finite derivatives of the photon wave function,

$$\frac{d\Delta\sigma_{\rm in}^{(4)}}{d\mathbf{b}dzd\mathbf{p}_-d\mathbf{\Delta}} = \frac{\alpha_S^2\sigma_0^2T^2(\mathbf{b})}{2(2\pi)^2(N_c^2-1)}\int_0^1 d\beta(1-\beta)$$

$$\times\int d\mathbf{q}_1 d\mathbf{q}_2 d\mathbf{\kappa}_1 d\mathbf{\kappa}_2 d\mathbf{\kappa}_3 f(\mathbf{q}_1) f(\mathbf{q}_2)$$

$$\times\Phi(\beta\nu_A(\mathbf{b}), \mathbf{\kappa}_1)\Phi(\beta\nu_A(\mathbf{b}), \mathbf{\kappa}_2)$$

$$\times\Phi((1-\beta)\lambda_c\nu_A(\mathbf{b}), \mathbf{\Delta}-\mathbf{\kappa}_3-\mathbf{q}_1-\mathbf{q}_2)$$

$$\times\Phi((1-\beta)\lambda_c\nu_A(\mathbf{b}), \mathbf{\kappa}_3)$$

$$\times\{\langle\gamma^*|z, \gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3\rangle$$

$$-\langle\gamma^*|z, \gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1\rangle\}$$

$$\times\{\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3\rangle - \langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_2\rangle$$

$$+\langle\gamma^*|z, \mathbf{p}_-+\mathbf{\kappa}_1+\mathbf{\kappa}_3+\mathbf{q}_1+\mathbf{q}_2\rangle\}, \quad (75)$$

and in the interesting case of hard dijets,

$$\frac{d\Delta\sigma_{\rm in}^{(4)}}{d\mathbf{b}dzd\mathbf{p}_-d\mathbf{\Delta}} = \frac{\alpha_S^2\sigma_0^2T^2(\mathbf{b})}{2(2\pi)^2(N_c^2-1)}\int_0^1 d\beta(1-\beta)$$

$$\times\int d\mathbf{q}_1 d\mathbf{q}_2 f(\mathbf{q}_1) f(\mathbf{q}_2)\Phi(2(1-\beta)\lambda_c\nu_A(\mathbf{b}), \mathbf{\Delta}-\mathbf{q}_1-\mathbf{q}_2)$$

$$\times\{\langle\gamma^*|z, \gamma^*|z, \mathbf{p}_-\rangle - \langle\gamma^*|z, \gamma^*|z, \mathbf{p}_-+\mathbf{q}_1\rangle\} \quad (76)$$

$$\times\{\langle\gamma^*|z, \mathbf{p}_-\rangle - \langle\gamma^*|z, \mathbf{p}_-+\mathbf{q}_1\rangle$$

$$-\langle\gamma^*|z, \mathbf{p}_-+\mathbf{q}_2\rangle + \langle\gamma^*|z, \mathbf{p}_-+\mathbf{q}_1+\mathbf{q}_2\rangle\}.$$

The leading term of the small-$\mathbf{q}_{1,2}$ expansion of the product of the photon wave functions in (74) is a quadratic function of $\mathbf{q}_1$ and a linear function of $\mathbf{q}_2$ of the form

$$\left|\langle\gamma^*|z,\mathbf{p}_-\rangle\right|^2\frac{(\mathbf{p}_-\cdot\mathbf{q}_1)^2(\mathbf{p}_-\cdot\mathbf{q}_2)}{p_-^6}. \tag{77}$$

The leading nonvanishing term comes from the expansion of the nuclear WW glue,

$$\begin{aligned}&\Phi(2(1-\beta)\lambda_c\nu_A(\mathbf{b}),\Delta-\mathbf{q}_1-\mathbf{q}_2)\\ &-\Phi(2(1-\beta)\lambda_c\nu_A(\mathbf{b}),\Delta)\sim\Phi(2(1-\beta)\lambda_c\nu_A(\mathbf{b}),\Delta)\\ &\times\frac{\Delta\cdot(\mathbf{q}_1+\mathbf{q}_2)}{2(1-\beta)\lambda_cQ_A^2(\mathbf{b})+\Delta^2}.\end{aligned} \tag{78}$$

Namely, upon azimuthal averaging of (77) in conjunction with (78), we find the leading nonvanishing term of the form $2(\mathbf{p}_-\mathbf{q}_2)(\Delta\mathbf{q}_2)\longrightarrow(\mathbf{p}_-\Delta)\mathbf{q}_2^2$ and therefore

$$\frac{d\Delta\sigma_{\rm in}^{(4)}}{d\sigma_{\rm in}}\sim\frac{1}{N_c^2-1}\frac{\mathbf{p}_-\cdot\Delta}{\mathbf{p}_-^2}, \tag{79}$$

which is reminiscent of a higher twist-3 correction.

To summarize, nonlinear nuclear $k_\perp$-factorization allows a consistent evaluation of the $1/N_c^2$-corrections. We demonstrated how the expansion in $1/(N_c^2-1)$ accompanies a higher twist expansion. One exception is the reabsorption of one of the terms proportional to $1/(N_c^2-1)$ in σ_{88} into the renormalization of the leading term in σ_{88} by the N_c-dependent factor λ_c. We conclude this discussion by a comment that all the arguments in Section 5 regarding the disappearance of azimuthal correlations of minijets hold for the $1/N_c^2$-corrections as well.

8. CONCLUSIONS

We formulated the theory of the breakup of photons into dijets in DIS off nuclear targets based on the consistent treatment of propagation of color dipoles in a nuclear medium. The non-Abelian intranuclear evolution of color dipoles gives rise to a nontrivial spectrum of the attenuation eigenvalues, but the familiar Glauber–Gribov multiple-scattering results are recovered for the nuclear total cross sections. However, in more special cases like DIS, in which the photon breaks up into color-singlet dijets, the cross section depends on the complete spectrum of the attenuation eigenstates.

We derived the nuclear broadening of the acoplanarity momentum distribution in the breakup of photons into dijets, see Eqs. (35) and (41). Our principal finding is that all nuclear DIS observables—the amplitude of coherent diffractive breakup into dijets [6], nuclear sea quark SF and its decomposition into equally important genuine inelastic and diffractive components performed in [5], and the jet–jet inclusive cross section derived in the present paper—are uniquely calculable in terms of the NSS-defined collective nuclear WW glue. This property can be regarded as a nuclear $k_\perp$-factorization theorem that connects DIS in the regimes of low and high density of partons. For the generic dijet cross section, nuclear $k_\perp$-factorization is of a highly nonlinear form, which must be contrasted to the linear hard factorization for the free nucleon target. This result is derived to the leading order in large N_c; further evaluation of the $1/N_c^2$-corrections shows a close relationship between the $1/N_c^2$ and high-twist expansions. Furthermore, the $1/N_c^2$-corrections themselves admit the nonlinear nuclear $k_\perp$-factorization representation.

We demonstrated the disappearance of azimuthal jet–jet correlations of minijets with momenta below the saturation scale. Based on the ideas on generalization of the dipole picture to hadron–nucleus collisions [17, 18], we presented qualitative estimates of the broadening effect for mid-rapidity jets produced in central nucleus–nucleus collisions and argued that our azimuthal decorrelation may contribute substantially to the disappearance of back-to-back high-$p_\perp$ hadron correlation in central gold–gold collisions observed by the STAR Collaboration at RHIC [20].

We conclude by the comment that all the results for hard single-jet and jet–jet inclusive cross sections can be readily extended from DIS to the breakup of projectile hadrons into forward jets. Indeed, as argued in [6], the final state interaction between the final state quark and antiquark can be neglected and the plane-wave approximation becomes applicable as soon as the invariant mass of the forward jet system exceeds a typical mass scale of prominent meson and baryon resonances. Here, we confine ourselves to the statement that although our principal point about a nonlinear nuclear $k_\perp$-factorization is fully retained, we find important distinctions between the breakup of pointlike photons and nonpointlike hadrons.

This work has been partly supported by INTAS (grant nos. 97-30494 and 00-00366) and the DFG (grant no. 436RUS17/119/02).

APPENDIX A

Calculation of the 4-Body Color Dipole Cross Section

The Feynman diagrams for the matrix of 4-parton dipole cross section $\sigma_4(\mathbf{s},\mathbf{r},\mathbf{r}')$, Eqs. (19)–(21), are

shown in Fig. 2. The profile function for the color-singlet $q\bar{q}$ pair is given by the diagrams in Figs. 2a–2d,

$$2\Gamma(\text{Figs. 2a–2d};\ (q\bar{q})_1 N;\ \mathbf{b}_+, \mathbf{b}_-) = \frac{1}{N_c}\delta_{ab}\{[\chi^2(\mathbf{b}_+) + \chi^2(\mathbf{b}_-)]\mathrm{Tr}(T^a T^b) - 2\chi(\mathbf{b}_+)\chi(\mathbf{b}_-)\mathrm{Tr}(T^a T^b)\} = \frac{N_c^2-1}{2N_c}[\chi(\mathbf{b}_+) - \chi(\mathbf{b}_-)]^2, \quad (80)$$

which has already been cited in the main text, Eq. (11). Upon adding the contribution from diagrams in Figs. 2e–2h, we obtain the obvious result in Eq. (19).

The color-diagonal contribution of the same diagrams to the interaction of the color-octet $q\bar{q}$ pair with the nucleon is given by

$$2\Gamma(\text{Figs. 2a–2d};\ (q\bar{q})_8 N;\ \mathbf{b}_+, \mathbf{b}_-) = \frac{2}{N_c^2-1}\delta_{ab}\{[\chi^2(\mathbf{b}_+) + \chi^2(\mathbf{b}_-)]\mathrm{Tr}(T^c T^a T^b T^c) - 2\chi(\mathbf{b}_+)\chi(\mathbf{b}_-)\mathrm{Tr}(T^c T^a T^c T^b)\} = \frac{N_c^2-1}{2N_c} \times \left\{[\chi^2(\mathbf{b}_+) + \chi^2(\mathbf{b}_-)] + \frac{2}{N_c^2-1}\chi(\mathbf{b}_+)\chi(\mathbf{b}_-)\right\}. \quad (81)$$

The contribution to the matrix element $\langle 88|\sigma_4|88\rangle$ from color-diagonal interactions of the $q'\bar{q}'$ pair is obtained from (81) by the substitution $\mathbf{b}_\pm \longrightarrow \mathbf{b}'_\pm$,

$$\Gamma_4(\text{Figs. 2a–2d + Figs. 2e–2h};\ (88)N;\ \mathbf{b}_+, \mathbf{b}_-, \mathbf{b}'_+, \mathbf{b}'_-) = \Gamma(\text{Figs. 2a–2d};\ (q\bar{q})_8 N;\ \mathbf{b}_+, \mathbf{b}_-) + \Gamma(\text{Figs. 2a–2d};\ (q\bar{q})_8 N;\ \mathbf{b}'_+, \mathbf{b}'_-). \quad (82)$$

The diagrams in Figs. 2i–2l describe processes with color-space rotation of the $q\bar{q}$ pair,

$$2\Gamma_4(\text{Figs. 2i–2l};\ (88)N \longrightarrow (88)N;\ \mathbf{b}_+, \mathbf{b}_-, \mathbf{b}'_+, \mathbf{b}'_-) = \frac{8}{N_c^2-1}\delta_{ab}\{[\chi(\mathbf{b}_+)\chi(\mathbf{b}'_-) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_+)] \times \mathrm{Tr}(T^c T^a T^d)\mathrm{Tr}(T^c T^b T^d) - [\chi(\mathbf{b}_+)\chi(\mathbf{b}'_+) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_-)] \times \mathrm{Tr}(T^c T^a T^d)\mathrm{Tr}(T^d T^b T^c)\} = -\frac{N_c^2-1}{N_c}\left\{\frac{2}{N_c^2-1}[\chi(\mathbf{b}_+)\chi(\mathbf{b}'_-) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_+)] + \frac{N_c^2-2}{N_c^2-1}[\chi(\mathbf{b}_+)\chi(\mathbf{b}'_+) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_-)]\right\}. \quad (83)$$

The $(11)N \longrightarrow (88)N$ transition matrix element comes from the diagrams in Figs. 2i–2l,

$$2\Gamma_4(\text{Figs. 2i–2l};\ (11)N \longrightarrow (88)N;\ \mathbf{b}_+, \mathbf{b}_-, \mathbf{b}'_+, \mathbf{b}'_-) = \frac{4}{N_c\sqrt{N_c^2-1}}\delta_{ab}\{[\chi(\mathbf{b}_+)\chi(\mathbf{b}'_-) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_+)] \times \mathrm{Tr}(T^c T^a)\mathrm{Tr}(T^c T^b) - [\chi(\mathbf{b}_+)\chi(\mathbf{b}'_+) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_-)] \times \mathrm{Tr}(T^c T^a)\mathrm{Tr}(T^c T^b)\} = \frac{N_c^2-1}{N_c}\frac{1}{\sqrt{N_c^2-1}} \times \{[\chi(\mathbf{b}_+)\chi(\mathbf{b}'_-) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_+)] - [\chi(\mathbf{b}_+)\chi(\mathbf{b}'_+) + \chi(\mathbf{b}_-)\chi(\mathbf{b}'_-)]\}. \quad (84)$$

Upon the rearrangement

$$-2\chi(\mathbf{b}_i)\chi(\mathbf{b}_j) = [\chi(\mathbf{b}_i) - \chi(\mathbf{b}_j)]^2 - \chi^2(\mathbf{b}_i) - \chi^2(\mathbf{b}_j),$$

we can readily verify that the terms proportional to $\chi^2(\mathbf{b}_i)$ cancel each other and the 4-body cross section matrix contains only linear combinations of $\sigma(\mathbf{b}_i - \mathbf{b}_j)$, recalling the discussion in [24].

APPENDIX B

Non-Abelian vs. Abelian Aspects of Intranuclear Propagation of Color Dipoles and the Glauber–Gribov Formalism

The intranuclear propagation of color-octet $q\bar{q}$ pairs is part and parcel of the complete formalism for DIS off nucleus. It is interesting to recover the quasi-Abelian color-dipole results for the nuclear cross sections [13, 14] that are of the Glauber–Gribov form [21, 22]. We first consider the total inelastic cross section obtained from (8) upon the integration over the transverse momenta $\mathbf{p}_\pm$ of the quark and antiquark, which amounts to putting $\mathbf{b}_+ = \mathbf{b}'_+$ and $\mathbf{b}_- = \mathbf{b}'_-$. Then we are left with the system of two color dipoles of the same size $\mathbf{r} = \mathbf{b}_+ - \mathbf{b}_- = \mathbf{r}' = \mathbf{b}'_+ - \mathbf{b}'_-$, and the matrix of the

4-body cross section has the eigenvalues

$$\Sigma_1 = 0, \tag{85}$$

$$\Sigma_2 = \frac{2N_c^2}{N_c^2-1}\sigma(\mathbf{r}) \tag{86}$$

with the eigenstates

$$|f_1\rangle = \frac{1}{N_c}(|11\rangle + \sqrt{N_c^2-1}|88\rangle), \tag{87}$$

$$|f_2\rangle = \frac{1}{N_c}(\sqrt{N_c^2-1}|11\rangle - |88\rangle). \tag{88}$$

The existence of the nonattenuating 4-quark state with $\Sigma_1 = 0$ is quite obvious and corresponds to an overlap of two $q\bar{q}$ dipoles of the same size with neutralization of color charges. The existence of such a nonattenuating state is shared by an Abelian and non-Abelian quark–gluon interaction. The intranuclear attenuation eigen–cross section (86) differs from $\sigma(\mathbf{r})$ for the color-singlet $q\bar{q}$ pair by the nontrivial color factor

$$2\lambda_c = 2N_c^2/(N_c^2-1) = C_A/C_F,$$

which occurs because the relevant 4-parton state is in the color octet–(anti)octet configuration.

The crucial point is that the final state that enters the calculation of the genuine inelastic DIS off a nucleus, see Eq. (15), is precisely the eigenstate $|f_1\rangle$. Then, even without invoking the Sylvester expansion (23) and (25), the straightforward result for the inelastic cross section is

$$\begin{aligned}\sigma_{\text{in}} &= \int d\mathbf{r}dz\left|\Psi(Q^2,z,\mathbf{r})\right|^2\int d\mathbf{b}\\ &\times\left\{N_c\langle f_1|\exp\left[-\frac{1}{2}\sigma_4 T(\mathbf{b})\right]|11\rangle - \exp[-\sigma(\mathbf{r})T(\mathbf{b})]\right\}\\ &= \int d\mathbf{b}\langle\gamma^*|\left\{\exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right] - \exp[-\sigma(\mathbf{r})T(\mathbf{b})]\right\}|\gamma^*\rangle\\ &= \int d\mathbf{b}\langle\gamma^*|\{1-\exp[-\sigma(\mathbf{r})T(\mathbf{b})]\}|\gamma^*\rangle,\end{aligned} \tag{89}$$

which is precisely the color-dipole generalization [14] of the Glauber–Gribov formula [21, 22] in which no trace of a non-Abelian intranuclear evolution with the nontrivial attenuation eigenstate (88) with eigen–cross section (86) is left.

When the photon breaks into a color-singlet $q\bar{q}$ dijet, the net flow of color between the $q\bar{q}$ pair and color-excited debris of the target nucleus is zero. This suggests that a rapidity gap can survive hadronization, although whether the rapidity gap in genuine inelastic events with the color-singlet $q\bar{q}$ production is stable against higher order corrections remains an interesting open question. Although the debris of the target nucleus have zero net color charge, the debris of color-excited nucleons are spatially separated by a distance on the order of the nuclear radius, which suggests the total excitation energy of the order of 1 GeV times $A^{1/3}$, such that such rapidity-gap events look like a double diffraction with multiple production of mesons in the nucleus fragmentation region (see [36] for a theoretical discussion of conventional mechanisms of diffraction excitation of nuclei in proton–nucleus collisions; experimental observation has been reported in [37]). As such, inelastic excitation of color-singlet dijets is distinguishable from quasielastic diffractive DIS followed by citation and breakup of the target nucleus without production of secondary particles.

Using the Sylvester expansion (23)–(25) and eigenstates (87) and (88), we readily obtain

$$\begin{aligned}\sigma_{\text{in}}(A^*(q\bar{q})_1) &= \int(d\mathbf{b})\langle\gamma^*|\left\{1-\exp[-\sigma(\mathbf{r})T(\mathbf{b})]\right.\\ &\left.-\frac{N_c^2-1}{N_c^2}\left(1-\exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right]\right)\right\}|\gamma^*\rangle,\end{aligned} \tag{90}$$

$$\begin{aligned}\sigma_{\text{in}}(A^*(q\bar{q})_8) &= \frac{N_c^2-1}{N_c^2}\\ &\times\int d\mathbf{b}\langle\gamma^*|\left\{1-\exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right]\right\}|\gamma^*\rangle.\end{aligned} \tag{91}$$

These expressions depend on the entire non-Abelian spectrum of attenuation eigenstates.

Several features of the result in (90) are noteworthy. First, the color neutralization of the $q\bar{q}$ pair after the first inelastic interaction requires at least one more secondary inelastic interaction, and the expansion of the integrand of $\sigma_{\text{in}}(A^*(q\bar{q})_1)$ begins with the term quadratic in the optical thickness,

$$\begin{aligned}&\left\{1-\exp[-\sigma(\mathbf{r})T(\mathbf{b})] - \frac{N_c^2-1}{N_c^2}\right.\\ &\left.\times\left(1-\exp\left[-\frac{1}{2}\Sigma_2 T(\mathbf{b})\right]\right)\right\}\\ &= \frac{1}{2(N_c^2-1)}\sigma^2(\mathbf{r})T^2(\mathbf{b}) + \ldots\end{aligned} \tag{92}$$

Second, in the large-N_c limit, the color-octet state tends to oscillate in color remaining in the octet state. This is clearly seen from (92). Third, in the limit of an opaque nucleus,

$$\sigma_{\rm in}(A^*(q\bar{q})_1) = \frac{1}{N_c}\int d\mathbf{b}\langle\gamma^*| \times\{1-\exp[-\sigma(\mathbf{r})T(\mathbf{b})]\}|\gamma^*\rangle = \frac{1}{N_c^2}\sigma_{\rm in}. \quad (93)$$

This remains a constant fraction of DIS in contrast to the quasielastic diffractive DIS or inelastic diffractive excitation of a nucleus, whose cross sections vanish for an opaque nucleus [14, 36].

The analysis of the single-parton, alias single-jet, inclusive cross section is quite similar. In this case, we integrate over the momentum $\mathbf{p}_-$ of the antiquark jet such that $\mathbf{b}'_- = \mathbf{b}_-$. The corresponding matrix σ_4 has the eigenvalues

$$\Sigma_1 = \sigma(\mathbf{r}-\mathbf{r}'), \quad (94)$$

$$\Sigma_2 = \frac{N_c^2}{N_c^2-1}[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]-\frac{1}{N_c^2-1}\sigma(\mathbf{r}-\mathbf{r}') \quad (95)$$

with exactly the same eigenstates $|f_1\rangle$ and $|f_2\rangle$ as given by Eqs. (87) and (88). Again, the cross section of the genuine inelastic DIS corresponds to the projection on the eigenstate $|f_1\rangle$, and hence,

$$\frac{d\sigma_{\rm in}}{d\mathbf{b}d\mathbf{p}dz} = \frac{1}{(2\pi)^2}\int d\mathbf{r}'d\mathbf{r}\exp[i\mathbf{p}\cdot(\mathbf{r}'-\mathbf{r})] \times\Psi^*(Q^2,z,\mathbf{r}')\Psi(Q^2,z,\mathbf{r})\left\{\exp\left[-\frac{1}{2}\Sigma_1 T(\mathbf{b})\right]\right. - \left.\exp\left[-\frac{1}{2}[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right]\right\} = \frac{1}{(2\pi)^2} \times\int d\mathbf{r}'d\mathbf{r}\exp[i\mathbf{p}\cdot(\mathbf{r}'-\mathbf{r})]\Psi^*(Q^2,z,\mathbf{r}')\Psi(Q^2,z,\mathbf{r}) \times\left\{\exp\left[-\frac{1}{2}\sigma(\mathbf{r}-\mathbf{r}')T(\mathbf{b})\right]\right. - \left.\exp\left[-\frac{1}{2}[\sigma(\mathbf{r})+\sigma(\mathbf{r}')]T(\mathbf{b})\right]\right\}, \quad (96)$$

which is precisely Eq. (10) in [5].

At this point, we emphasize that, for the fundamental reason that the relevant final state is precisely the eigenstate $|f_1\rangle$, the calculations of the integrated inelastic cross section (89) and of the one-particle inclusive inelastic spectrum (96) are essentially Abelian problems, and the final result in (96) is identical, apart from a very different notation, to that for the propagation of relativistic positronium in dense media derived by one of the authors [32]. As can be seen from inspection of the relevant four-parton states, all contributions from the propagation of color-octet dipoles cancel, and the results can be obtained from studying the propagation of color-singlet dipoles without any reference to the full cross section matrix σ_4. Our formalism makes these cancellations nicely explicit. These quasi-Abelian problems have also been studied in [2, 38].

APPENDIX C

Weizsäcker–Williams Glue of Spatially Overlapping Nucleons

According to [5, 6], the multiple convolutions $f^{(j)}(\boldsymbol{\kappa}^2)$ have the meaning of a collective unintegrated gluon SF of j nucleons at the same impact parameter such that their Weizsäcker–Williams gluon fields overlap spatially in a Lorentz-contracted nucleus. These convolutions can also be viewed as a random walk in which $f(\boldsymbol{\kappa}^2)$ describes the single walk distribution.

To the lowest order in pQCD, the large-$\boldsymbol{\kappa}^2$ behavior is $f(\boldsymbol{\kappa}^2)\propto\alpha_S(\boldsymbol{\kappa}^2)/\kappa^4$. The phenomenological study of the differential glue of the proton in [19] suggests a useful large-$\boldsymbol{\kappa}^2$ approximation $f(\boldsymbol{\kappa}^2)\propto 1/(\boldsymbol{\kappa}^2)^\gamma$ with the exponent $\gamma\approx 2$ (a closer inspection of numerical results in [19] gives $\gamma\approx 2.15$ at $x=10^{-2}$). The QCD evolution effects enhance $f(\boldsymbol{\kappa}^2)$ at large $\boldsymbol{\kappa}^2$: the smaller x, the stronger the enhancement.

Because $f(\boldsymbol{\kappa}^2)$ decreases very slowly, we encounter a manifestly non-Gaussian random walk. For instance, as argued in [6], a j-fold walk to large $\boldsymbol{\kappa}^2$ is realized by one large walk, $\boldsymbol{\kappa}_1^2\sim\boldsymbol{\kappa}^2$, accompanied by $j-1$ small walks. We simply quote the main result in [6],

$$f^{(j)}(\boldsymbol{\kappa}^2) = j f(\boldsymbol{\kappa}^2)\left[1+\frac{4\pi^2(j-1)\gamma^2}{N_c\sigma_0\boldsymbol{\kappa}^2}G(\boldsymbol{\kappa}^2)\right], \quad (97)$$

where $G(\boldsymbol{\kappa}^2)$ is the conventional integrated gluon SF. Then the hard tail of unintegrated nuclear glue per bound nucleon,

$$f_{WW}(\mathbf{b},\boldsymbol{\kappa}^2) = \phi_{WW}(\nu_A(\mathbf{b}),\boldsymbol{\kappa}^2)/\nu_A(\mathbf{b}),$$

can be calculated parameter-free,

$$f_{WW}(\mathbf{b},\boldsymbol{\kappa}^2) = \frac{1}{\nu_A(\mathbf{b})}\sum_{j=1}^{\infty} w_j(\nu_A(\mathbf{b}))\, j f^{(j)}(\boldsymbol{\kappa}^2) \times\left[1+\frac{4\pi^2\gamma^2}{N_c\sigma_0\boldsymbol{\kappa}^2}(j-1)G(\boldsymbol{\kappa}^2)\right] = f(\boldsymbol{\kappa}^2)\left[1+\frac{2\pi^2\gamma^2\alpha_S(r)T(\mathbf{b})}{N_c\boldsymbol{\kappa}^2}G(\boldsymbol{\kappa}^2)\right]. \quad (98)$$

In the hard regime, the differential nuclear glue is not shadowed; furthermore, because of the manifestly positive-valued and model-independent nuclear higher twist correction, it exhibits a nuclear antishadowing property [6].

We now present the arguments in favor of the scaling small-$\boldsymbol{\kappa}^2$ behavior

$$f^{(j)}(\boldsymbol{\kappa}^2)\approx\frac{1}{Q_j^2}\xi\left(\frac{\boldsymbol{\kappa}^2}{Q_j^2}\right)\approx\frac{1}{\pi}\frac{Q_j^2}{(\boldsymbol{\kappa}^2+Q_j^2)^2} \quad (99)$$

with

$$Q_j^2\approx jQ_0^2. \quad (100)$$

In the evolution of $f^{(j)}(\boldsymbol{\kappa}^2)$ with j at moderate $\boldsymbol{\kappa}^2$,

$$f^{(j+1)}(\boldsymbol{\kappa}^2) = \int d\mathbf{k}\, f(\mathbf{k}^2) f^{(j)}((\boldsymbol{\kappa}-\mathbf{k})^2), \quad (101)$$

the function $f(\mathbf{k}^2)$ is steep compared to the smooth and broad function $f^{(j)}((\boldsymbol{\kappa}-\mathbf{k})^2)$, and we can therefore expand

$$f^{(j)}((\boldsymbol{\kappa}-\mathbf{k})^2) = f^{(j)}(\boldsymbol{\kappa}^2)+\frac{df^{(j)}(\boldsymbol{\kappa}^2)}{d\boldsymbol{\kappa}^2}[\mathbf{k}^2-2\boldsymbol{\kappa}\cdot\mathbf{k}] +\frac{1}{2}\frac{df^{(j)}(\boldsymbol{\kappa}^2)}{(d\boldsymbol{\kappa}^2)^2}4(\boldsymbol{\kappa}\cdot\mathbf{k})^2 \longrightarrow f^{(j)}(\boldsymbol{\kappa}^2) +\mathbf{k}^2\left[\frac{df^{(j)}(\boldsymbol{\kappa}^2)}{d\boldsymbol{\kappa}^2}+\boldsymbol{\kappa}^2\frac{d^2f^{(j)}(\boldsymbol{\kappa}^2)}{(d\boldsymbol{\kappa}^2)^2}\right] = f^{(j)}(\boldsymbol{\kappa}^2)+\mathbf{k}^2\frac{d}{d\boldsymbol{\kappa}^2}\left[\boldsymbol{\kappa}^2\frac{df^{(j)}(\boldsymbol{\kappa}^2)}{d\boldsymbol{\kappa}^2}\right], \quad (102)$$

where $\longrightarrow$ indicates azimuthal averaging. The expansion (102) holds for $\mathbf{k}^2\lesssim Q_j^2$, and after the $d\mathbf{k}$ integration in (101), we obtain

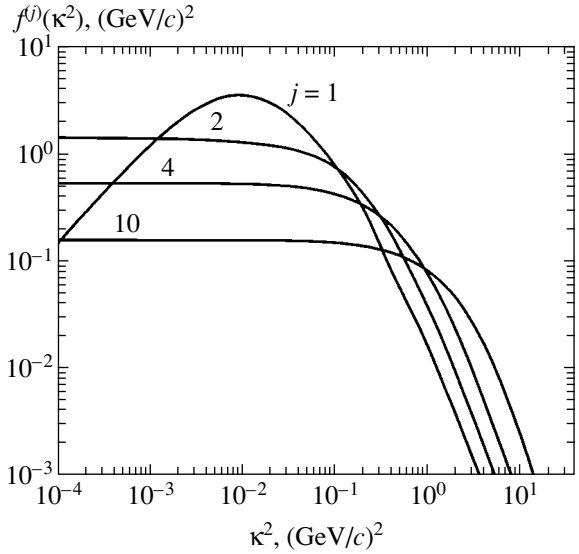

Fig. 7. The nuclear dilution for soft momenta and broadening for hard momenta of the collective glue of j overlapping nucleons, $f^{(j)}(\boldsymbol{\kappa}^2)$. The numerical results are for DIS at $x = 0.01$, and the input unintegrated gluon SF of the proton is taken from [19].

$$f^{(j+1)}(\boldsymbol{\kappa}^2) = f^{(j)}(\boldsymbol{\kappa}^2)+g(j)\frac{d}{d\boldsymbol{\kappa}^2}\left[\boldsymbol{\kappa}^2\frac{df^{(j)}(\boldsymbol{\kappa}^2)}{d\boldsymbol{\kappa}^2}\right], \quad (103)$$

where

$$g(j) = \int^{Q_j^2} d\mathbf{k}\mathbf{k}^2 f(\mathbf{k}^2) = \frac{4\pi^2}{N_c\sigma_0}G(Q_j^2). \quad (104)$$

This is a smooth function of j. It is easy to verify that our approximation preserves the normalization condition $\int d\boldsymbol{\kappa} f^{(j)}(\boldsymbol{\kappa}^2) = 1$.

For small $\boldsymbol{\kappa}^2$ and large j, recurrence relation (104) amounts to the differential equation

$$\frac{Q_{j+1}^2-Q_j^2}{Q_{j+1}^2Q_j^2} = \frac{1}{Q_j^4}\frac{dQ_j^2}{dj} = -\frac{1}{Q_j^4}\frac{\xi'(0)}{\xi(0)}g(j) \quad (105)$$

with the solution

$$Q_j^2 = -\frac{\xi'(0)}{\xi(0)}\int^j dj' g(j') \approx -jg(j)\frac{\xi'(0)}{\xi(0)}. \quad (106)$$

Expansion (102) holds up to the terms proportional to $\boldsymbol{\kappa}^2$ and its differentiation at $\boldsymbol{\kappa}^2 = 0$ gives a similar constraint on the j-dependence of Q_j^2.

We note that expansion of the plateau with j entails a dilution of the differential collective glue $f^{(j)}(\boldsymbol{\kappa}^2)$ in the plateau region,

$$f^{(j)}(\boldsymbol{\kappa}^2 \lesssim Q_j^2) \propto 1/Q_j^2 \propto 1/j.$$

We conclude by the observation that, when extended to $\boldsymbol{\kappa}^2 \gtrsim Q_j^2$, the parameterization in (101) and (100) behaves as $jQ_0^2/(\boldsymbol{\kappa}^2)^2$, which nicely matches the j-dependence of the leading twist term in the hard asymptotic form (99).

For a heavy nucleus, the dominant contribution to the expansion in (31) comes from $j \approx \nu_A(\mathbf{b})$, and hence

$$\phi_{WW}(\nu_A(\mathbf{b}), \boldsymbol{\kappa}^2) \approx \frac{1}{\pi}\frac{Q_A^2(\mathbf{b})}{\left(\boldsymbol{\kappa}^2 + Q_A^2(\mathbf{b})\right)^2}, \qquad (107)$$

where Eq. (106) gives the width of the plateau,

$$\begin{aligned} Q_A^2(\mathbf{b}) &\approx 2\nu_A(\mathbf{b})g(\nu_A(\mathbf{b})) \\ &\approx \frac{4\pi^2}{N_c}\alpha_S(Q_A^2)G(Q_A^2)T(\mathbf{b}). \end{aligned} \qquad (108)$$

The explicit dependence on the soft parameter σ_0 that is manifest in (104) cancels in (108). For DIS within the saturation domain, $Q^2 \lesssim Q_A^2$, the strong coupling in (33) must be taken at $r \sim 1/Q_A$, and the right-hand side of Eq. (108) exhibits only a weak dependence on the infrared parameters via the Q_A^2 dependence of the running strong coupling constant and scaling violations in the gluon SF of the nucleon. For instance, at $x = 10^{-2}$, the numerical results [19] for $G(Q^2)$ correspond to a nearly Q^2-independent $\alpha_S(Q^2)G(Q^2) \approx 1$. For the average DIS on a heavy nucleus,

$$\langle T(\mathbf{b})\rangle \approx \frac{3}{4}T(0) \approx \frac{9}{8\pi r_0^2}A^{1/3}, \qquad (109)$$

where $r_0 \approx 1.1$ fm. For lighter nuclei with the Gaussian density profile, $\langle T(\mathbf{b})\rangle \approx \frac{1}{2}T(0)$. Then for $N_c = 3$ and $A^{1/3} = 6$, Eqs. (108) and (109) give $\langle Q_A^2(\mathbf{b})\rangle \approx 0.8\ (\mathrm{GeV}/c)^2$.

The utility of approximation (99), (100) is illustrated in Fig. 7, where we show the j-dependence of the collective glue of j overlapping nucleons calculated for the unintegrated gluon SF of the proton from |19]. For the interaction of $q\bar{q}$ color dipoles in the average DIS on gold, ^{197}Au target, we find $\langle Q_{3A}^2(\mathbf{b})\rangle \approx 0.9\ (\mathrm{GeV}/c)^2$, in good agreement with the above estimate in Eq. (108).

For the $q\bar{q}g$ Fock states of the photon, the leading $\ln Q^2$ configurations correspond to small $q\bar{q}$ pairs that act as a color-octet gluon [24]; for such gluon–gluon color dipoles, $\langle Q_{8A}^2(\mathbf{b})\rangle \approx 2.1\ (\mathrm{GeV}/c)^2$. We note in passing that the standard collinear splitting sets in, and the DGLAP evolution [34, 39] becomes applicable to the nuclear structure, only at $Q^2 \gg \langle Q_{8A}^2(\mathbf{b})\rangle$.

REFERENCES

1. E. Leader and E. Predazzi, *Introduction to Gauge Theories and Modern Particle Physics* (Cambridge Univ. Press, Cambridge, 1996), Vol. 1; G. Sterman, *An Introduction to Quantum Field Theory* (Cambridge Univ. Press, Cambridge, 1993).
2. A. H. Mueller, Nucl. Phys. B **558**, 285 (1999); in *Lectures at the Cargèse Summer School* (2001); hep-ph/0111244.
3. A. H. Mueller, Nucl. Phys. B **335**, 115 (1990).
4. L. McLerran and R. Venugopalan, Phys. Rev. D **49**, 2233 (1994); J. Jalilian-Marian, A. Kovner, L. D. McLerran, and H. Weigert, Phys. Rev. D **55**, 5414 (1997); E. Iancu, A. Leonidov, and L. McLerran, in *Lectures at the Cargèse Summer School* (2001); hep-ph/0202270.
5. N. N. Nikolaev, W. Schäfer, B. G. Zakharov, and V. R. Zoller, JETP Lett. **76**, 195 (2002).
6. N. N. Nikolaev, W. Schäfer, and G. Schwiete, Pis'ma Zh. Éksp. Teor. Fiz. **72**, 583 (2000) [JETP Lett. **72**, 405 (2000)]; Phys. Rev. D **63**, 014020 (2001).
7. L. N. Lipatov, Yad. Fiz. **23**, 642 (1976) [Sov. J. Nucl. Phys. **23**, 338 (1976)]; E. A. Kuraev, L. N. Lipatov, and V. S. Fadin, Zh. Éksp. Teor. Fiz. **71**, 840 (1976) [Sov. Phys. JETP **44**, 443 (1976)]; Zh. Éksp. Teor. Fiz. **72**, 377 (1977) [Sov. Phys. JETP **45**, 199 (1977)]; Ya. Ya. Balitsky and L. N. Lipatov, Yad. Fiz. **28**, 1597 (1978) [Sov. J. Nucl. Phys. **28**, 822 (1978)].
8. I. P. Ivanov, N. N. Nikolaev, W. Schäfer, *et al.*, in *Proceedings of 36th Annual Winter School on Nuclear and Particle Physics and 8th St. Petersburg School on Theoretical Physics, St. Petersburg, Russia* (2002); hep-ph/0212161.
9. I. P. Ivanov, N. N. Nikolaev, W. Schäfer, *et al.*, Invited talk at the *NATO Advanced Research Workshop on Diffraction, Alushta, Ukraine* (2002), NATO Science Series: Physics and Chemistry, Ed. by R. Fiore, L. L. Jenkovszky, M. I. Kotsky, *et al.*, (Kluwer, 2003), Vol. 101, pp. 209–220; hep-ph/0212176; in *Proceedings of the Workshop on Exclusive Processes at High Momentum Transfer*, Jefferson Lab., 2002, Ed. by A. Radyushkin and P. Stoler (World Sci., 2002), p. 205; in *Proceedings of the Conference on Quark Nuclear Physics, QNP'2002, Jülich, Germany*, Ed. by C. Elster and Th. Walcher, Eur. Phys. J. (2003) (in press); hep-ph/0209298; Plenary talk at the *International Symposium on Multiparticle Dynamics, ISMD'2002, Alushta, Ukraine, 2002*, Ed. by G. Kozlov and A. Sissakian (World Sci., 2003) (in press).
10. T. Ahmed *et al.* (H1 Collaboration), Nucl. Phys. B **445**, 195 (1995).
11. A. Szczurek, N. N. Nikolaev, W. Schäfer, and J. Speth, Phys. Lett. B **500**, 254 (2001).

12. J. R. Forshaw and R. G. Roberts, Phys. Lett. B **335**, 494 (1994); A. J. Askew, D. Graudenz, J. Kwiecinski, and A. D. Martin, Phys. Lett. B **338**, 92 (1994); J. Kwiecinski, A. D. Martin, and A. M. Stasto, Phys. Lett. B **459**, 644 (1999).
13. N. N. Nikolaev and B. G. Zakharov, Z. Phys. C **49**, 607 (1991).
14. N. N. Nikolaev, B. G. Zakharov, and V. R. Zoller, Z. Phys. A **351**, 435 (1995).
15. V. Barone, M. Genovese, N. N. Nikolaev, *et al.*, Z. Phys. C **58**, 541 (1993).
16. N. N. Nikolaev and V. I. Zakharov, Yad. Fiz. **21**, 434 (1975) [Sov. J. Nucl. Phys. **21**, 227 (1975)]; Phys. Lett. B **55**, 397 (1975).
17. N. N. Nikolaev, G. Piller, and B. G. Zakharov, Zh. Éksp. Teor. Fiz. **108**, 1554 (1995) [JETP **81**, 851 (1995)]; Z. Phys. A **354**, 99 (1996).
18. B. G. Zakharov, Pis'ma Zh. Éksp. Teor. Fiz. **63**, 906 (1996) [JETP Lett. **63**, 952 (1996)]; Pis'ma Zh. Éksp. Teor. Fiz. **65**, 585 (1997) [JETP Lett. **65**, 615 (1997)]; Yad. Fiz. **61**, 924 (1998) [Phys. At. Nucl. **61**, 838 (1998)].
19. I. P. Ivanov and N. N. Nikolaev, Yad. Fiz. **64**, 813 (2001) [Phys. At. Nucl. **64**, 753 (2001)]; Phys. Rev. D **65**, 054004 (2002).
20. C. Adler *et al.* (STAR Collaboration), Phys. Rev. Lett. **90**, 082302 (2003).
21. R. J. Glauber, in *Lectures in Theoretical Physics*, Ed. by W. E. Brittin *et al.* (Interscience, New York, 1959), p. 315.
22. V. N. Gribov, Zh. Éksp. Teor. Fiz. **56**, 892 (1969) [Sov. Phys. JETP **29**, 483 (1969)].
23. N. N. Nikolaev and B. G. Zakharov, Z. Phys. C **53**, 331 (1992).
24. N. N. Nikolaev and B. G. Zakharov, Zh. Éksp. Teor. Fiz. **105**, 1498 (1994) [JETP **78**, 806 (1994)]; Z. Phys. C **64**, 631 (1994).
25. N. N. Nikolaev, B. G. Zakharov, and V. R. Zoller, Pis'ma Zh. Éksp. Teor. Fiz. **59**, 8 (1994) [JETP Lett. **59**, 6 (1994)].
26. N. N. Nikolaev and B. G. Zakharov, Phys. Lett. B **332**, 184 (1994).
27. V. Barone, M. Genovese, N. N. Nikolaev, *et al.*, Phys. Lett. B **326**, 161 (1994).
28. B. Andersson *et al.* (Small-x Collaboration), Eur. Phys. J. C **25**, 77 (2002).
29. M. Genovese, N. N. Nikolaev, and B. G. Zakharov, Zh. Éksp. Teor. Fiz. **108**, 1155 (1995) [JETP **81**, 633 (1995)].
30. C. Adloff *et al.* (H1 Collaboration), Z. Phys. C **76**, 613 (1997).
31. J. Breitweg *et al.* (ZEUS Collaboration), Eur. Phys. J. C **6**, 43 (1999).
32. B. G. Zakharov, Yad. Fiz. **46**, 148 (1987) [Sov. J. Nucl. Phys. **46**, 92 (1987)].
33. N. N. Nikolaev, J. Speth, and B. G. Zakharov, Zh. Éksp. Teor. Fiz. **109**, 1948 (1996) [JETP **82**, 1046 (1996)].
34. Yu. L. Dokshitser, Zh. Éksp. Teor. Fiz. **73**, 1216 (1977) [Sov. Phys. JETP **46**, 641 (1977)]; Yu. L. Dokshitzer, D. Diakonov, and S. I. Troian, Phys. Rep. **58**, 269 (1980).
35. T. Sjöstrand, P. Eden, C. Friberg, *et al.*, Comput. Phys. Commun. **135**, 238 (2001).
36. V. R. Zoller, Z. Phys. C **51**, 659 (1991); M. A. Faessler, Z. Phys. C **58**, 567 (1993).
37. T. Akesson *et al.* (HELIOS Collaboration), Z. Phys. C **49**, 355 (1991).
38. U. A. Wiedemann, Nucl. Phys. B **582**, 409 (2000).
39. V. N. Gribov and L. N. Lipatov, Yad. Fiz. **15**, 781 (1972) [Sov. J. Nucl. Phys. **15**, 438 (1972)]; L. N. Lipatov, Yad. Fiz. **20**, 340 (1974) [Sov. J. Nucl. Phys. **20**, 181 (1975)]; G. Altarelli and G. Parisi, Nucl. Phys. B **126**, 298 (1977).

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 466–478.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 516–528.

ATOMS, SPECTRA, RADIATION

Boundary-Value Problems in Near-Field Optical Microscopy and Optical Size Resonances

O. N. Gadomsky* and A. S. Kadochkin
Ul'yanovsk State University, Ul'yanovsk, 432700 Russia
**e-mail: qed_group@mail.ru*
Received December 30, 2002

Abstract—We have solved a boundary-value problem for a ball probe interacting with a flat dielectric surface in an external optical radiation field. This interaction gives rise to the optical size resonance at frequencies significantly different from the natural frequencies of two-level atoms both in the medium and in the probe with allowance for the local field corrections. These resonances depend significantly on the distance from the probe center to the surface, on the ball probe size, on the concentration of two-level atoms in the probe and in the medium, on the spectral line width, and on the atomic inversion. The field strengths inside and outside the ball probe and a semiinfinite dielectric medium have been calculated in the near-field and wave zones. It is shown that the proposed electrodynamic theory of optical near-field microscopy agrees with the results of experimental measurements. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In an interpretation of the near-field optical microscopy results, the probe is approximated by a sphere of radius a = 10–20 nm interacting with a sample surface occurring at a distance of $R \ll \lambda$ from the probe, where λ is wavelength of an external radiation exciting the probe–sample system. A response signal generated in the region of the probe–sample interaction has the same wavelength as that of the external radiation and propagates in the direction opposite to that of the incident radiation [1–4]. The spatial resolution of such near-field optical microscopes is on the order of 10 nm.

Theoretical justification of the experimental data of near-field optical microscopy is based on calculations of the probe polarizability and the probe–sample interaction within the framework of the electrostatic approximation. In some cases, this approach provides for a qualitative agreement with experiment. However, complete agreement of the theory and experiment can be achieved only based on a quantitative electrodynamic theory of the near-field optical microscope.

This study is devoted to solving a typical boundary-value problem for a probe interacting with the flat surface of a semiinfinite dispersive dielectric medium (Fig. 1a). The proposed theory describes the formation of a near-field wave propagating in the direction opposite to that of the incident wave. Based on the obtained solution, we will study the amplitude and phase properties of the radiation field in the near-field and wave zones relative to the probe–sample interaction.

Theoretical description of the method of scanning near-field optical microscopy has been developed in a number of papers [5–14] using both macroscopic and microscopic approaches. However, the optical radiation field in the near-field zone was described in the point dipole approximation. In this approximation (consistent with the electrostatic approximation), the electric field strength varies in the space according to the $1/R^3$ law, where R is the distance from the probe center to the point of observation. The field inside the probe is assumed to be homogeneous.

Below, we will solve a boundary-value problem for a ball probe of finite size with allowance for an inhomogeneous field inside the probe and dispersive properties of the probe. In this description, the radiation field in the near-field and wave zones may vary according to the laws of $1/R^4$, $1/R^3$, $1/R^2$, and $1/R$. In the near-field zone, the predominant role is played by the $1/R^4$ law describing a decrease in the electric field strength with increasing distance from the probe center. Based on this description, we will calculate geometrical factors determining the field inside and outside the probe taking into account the self- consistent probe–sample interaction in the external radiation field.

We believe that it is important to distinguish between the boundary-value problems of three main types in near-field optical microscopy. In the optical scheme of Fig. 1a, a ball probe interacts with the clean surface of a semiinfinite medium. Besides the Fresnel reflection component, this system features a wave that propagates in the direction opposite to that of the incident radiation field. Figure 1b illustrates the boundary-value problem of another type, in which particles on the substrate surface are studied under the conditions of a highly selective interaction, when the influence of other particles can be ignored. Figure 1c shows the case when the role of "third" particles present on the substrate surface is significant. The present study is devoted to solv-

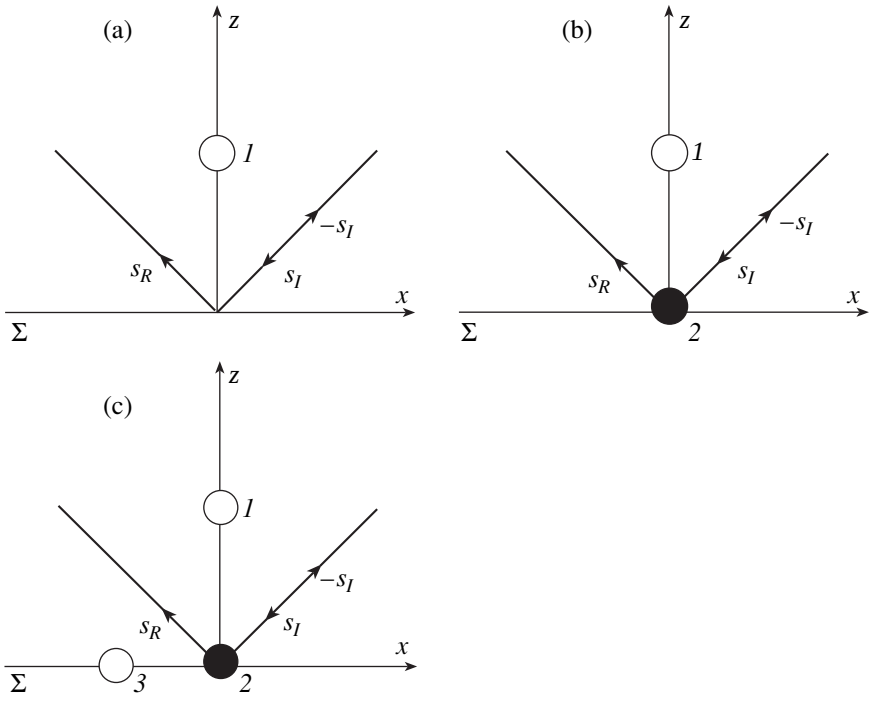

Fig. 1. Three main types of boundary-value problems for near-field optical microscopy: (a) probe *1* interacts with a clean flat sample surface Σ; (b) probe *1* interacts with particle *2* occurring on a flat surface of a semiinfinite dielectric sample surface Σ; (c) probe *1* interacts with particle *2* in the presence of a polarizing action of particle *3* and a semiinfinite dielectric medium (s_I, s_R, and $-s_I$ are the unit vectors in the directions of the incident wave, Fresnel reflection, and backscattered signal, respectively).

ing the boundary-value problem of the type illustrated in Fig. 1a. The solution of this problem determines the analysis of the other typical boundary-value problems of near-field optical microscopy.

Previously [15–17], we reported on the optical size resonances in two-level systems, the frequencies of which were significantly different from the frequencies of transitions in the spectra of interacting atoms. Recently [18], it was reported that the optical size resonances are experimentally observed in the spectra of anisotropic light reflection from arsenic and gallium dimers on the surface of gallium arsenide. We believe that the optical size resonances must also be present in other systems involving small numbers of atoms, such as composites, clusters, fullerenes, self-organized atomic chains on solid surfaces, etc.

As will be shown below, the ball probe–sample surface system features optical size resonances of a new type depending on the probe composition and the optical properties of the substrate. It should be noted that Keller *et al.* [19, 20] mentioned the so-called configuration resonances arising in the interaction of a metal ball probe with the surface of a semiinfinite medium, but the quantum properties of electrons in the microprobe were ignored. In the case of optical size resonances, an important role is played by the quantum properties of interacting atoms. In the ball probe–substrate system, the important role of the quantum properties of two-level atoms is manifested via the inversion.

In this paper, a ball probe represents a system of two-level atoms, which implies that the probe is considered as a dielectric particle. In particular, two-level atoms can be considered as impurity atoms in a glass matrix, whereby the spectral line width varies depending on the concentration of these impurities. As will be shown below, the smaller the spectral line width, the better the spatial resolution and the sensitivity of a near-field optical microscope. However, the two-level approximation can be also used for description of the ball probes of other types (semiconductor and metal probes) employed in near-field optical microscopy. It must be taken into account that ball probes with a radius of 10–30 nm exhibit a dimensional quantization of electrons and holes, so that such probes behave as quantum dots.

2. MAIN EQUATIONS

The radiation field in an arbitrary observation point **r** is determined by the strength of the electric **E**(**r**, *t*) and magnetic **H**(**r**, *t*) fields obeying the following integro-differential equations:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_I(\mathbf{r}, t) + \int \operatorname{curl}\operatorname{curl}\frac{\mathbf{P}_1(t - R_1'/c)}{R_1'}dV_1' + \int \operatorname{curl}\operatorname{curl}\frac{\mathbf{P}(\mathbf{r}', t - R_1'/c)}{R'}dV', \quad (2.1a)$$

$$\mathbf{H}(\mathbf{r}, t) = \mathbf{H}_I(\mathbf{r}, t) + \frac{1}{c}\int \mathrm{curl}\frac{(\partial/\partial t)\mathbf{P}_1(t - R_1'/c)}{R_1'}dV_1' + \frac{1}{c}\int \mathrm{curl}\frac{(\partial/\partial t)\mathbf{P}(\mathbf{r}', t - R_1'/c)}{R'}dV', \quad (2.1b)$$

where $\mathbf{E}_I(\mathbf{r}, t)$ and $\mathbf{H}_I(\mathbf{r}, t)$ are the electric and magnetic field strengths of the incident external wave, respectively; $R_1' = |\mathbf{r} - \mathbf{r}_1'|$; $\mathbf{r}_1'$ is a radius vector in a ball of radius a that serves as the probe (Fig. 1a) in the problem under consideration; $R_1' = |\mathbf{r} - \mathbf{r}'|$; $\mathbf{r}'$ is an arbitrary radius vector inside or on the surface of the dielectric medium studied; $\mathbf{P}_1$ is the induced dipole moment per unit volume of the probe; $\mathbf{P}$ is the polarization vector of the medium; and c is the velocity of light in vacuum.

In Eqs. (2.1), the differentiation is performed with respect to the coordinates of the observation point, while the integrals are taken over all coordinates $\mathbf{r}'$ and $\mathbf{r}_1'$ for all points outside the medium and probe; for a point in the medium, the integrals are taken over the entire volume of the medium bounded by the external surface Σ and a small spherical surface σ_0 with the radius $L_0 \longrightarrow 0$ containing the observation point. The field of discrete dipoles distributed inside the Lorentz sphere σ_0 of radius L_0 is assumed to be zero. This condition is obeyed with good precision for closed Lorentz spheres in the absence of near-field effects [15]. In our analysis, an analogous situation takes place for the observation points inside the probe in the absence of the near-field effect (i.e., with neglect of the contribution from discrete dispersed atoms).

In calculating the polarization vectors $\mathbf{P}$ and $\mathbf{P}_1$ for the boundary-value problem under consideration, we take into account only the electric vector and neglect the relativistic contribution related to the magnetic vector. The magnetic field strength will be taken into account in calculations of the dipole radiation in the wave zone relative to the probe position. The polarization vectors can be presented in the following form:

$$\mathbf{P}_1 = N_1\mathbf{p}_1 = \frac{1}{2}N_1\mathbf{X}_1\exp(-i\omega t) + \text{c.c.},$$
$$\mathbf{P} = N\mathbf{p} = \frac{1}{2}N\mathbf{X}\exp(-i\omega t) + \text{c.c.}, \quad (2.2)$$

where ω is the radiation field frequency and N, N_1 are the coordinate-independent densities of induced dipole moments inside the medium and probe, respectively. The quantum-mechanical means $\mathbf{X}$ and $\mathbf{X}_1$ obey the modified optical Bloch equations [17]. For the boundary-value problem under consideration, these equations are as follows:

$$\frac{d\mathbf{X}_1}{dt} = -i\mathbf{X}_1\Delta_1 - \frac{2i}{\hbar}w_1|\mathbf{d}_{01}|^2\mathbf{E}_{01} - \frac{1}{T_{21}'}\mathbf{X}_1, \quad (2.3a)$$

$$\frac{dw_1}{dt} = \frac{i}{\hbar}\mathbf{X}_1^*\cdot\mathbf{E}_{01} - \frac{i}{\hbar}\mathbf{X}_1\cdot\mathbf{E}_{01}^* - \frac{1}{T_{11}}(w_1 - w_{01}). \quad (2.3b)$$

Here $\Delta = \omega_{01} - \omega$ is the detuning from an isolated resonance at a frequency of ω_{01} in the probe; $\mathbf{d}_{01}$ is the dipole moment of the transition in the probe; w_1 is the inversion; w_{01} is the initial inversion in the probe, which determines the difference of populations of the quantum states involved in the allowed dipole quantum transitions; T_{21}' and T_{11} are the characteristic times of the phase and energy relaxation in the probe for a given quantum transition; and $\mathbf{E}_{01}$ is the effective field strength at the probe center.

By the same token, we can write the modified optical Bloch equations for the induced dipole moments of the medium in the vicinity of an isolated resonance at a frequency of ω_0:

$$\frac{d\mathbf{X}}{dt} = -i\mathbf{X}\Delta - \frac{2i}{\hbar}w|\mathbf{d}_0|^2\mathbf{E}_0 - \frac{1}{T_2'}\mathbf{X}, \quad (2.4a)$$

$$\frac{dw}{dt} = \frac{i}{\hbar}\mathbf{X}^*\cdot\mathbf{E}_0 - \frac{i}{\hbar}(\mathbf{X}\cdot\mathbf{E}_0^*) - \frac{1}{T_1}(w - w_0), \quad (2.4b)$$

where $\Delta = \omega_0 - \omega$ is the detuning from the isolated resonance; the quantities w, w_0, $\mathbf{d}_0$, T_2', and T_1, having the corresponding physical meaning, characterize the allowed dipole quantum transitions in the spectra of atoms (molecules) of the medium studied.

The effective microscopic fields $\mathbf{E}_{01}$ and $\mathbf{E}_0$ entering Eqs. (2.3) and (2.4) can be presented in accordance with Eqs. (2.1) (the time factor $\exp(-i\omega t)$ is omitted):

$$\mathbf{E}_{01}(\mathbf{r}_1) = \mathbf{E}_{0I}\exp(i\mathbf{k}_0\cdot\mathbf{r}_1) + \hat{a}_R(\mathbf{r}_1)\mathbf{X}N + \hat{a}_{T1}(\mathbf{r}_1)\mathbf{X}_1N_1, \quad (2.5a)$$

$$\mathbf{E}_0(\mathbf{r}') = \mathbf{E}_{0I}\exp(i\mathbf{k}_0\cdot\mathbf{r}') + \hat{a}_T(\mathbf{r}')\mathbf{X}N + \hat{a}_{R1}(\mathbf{r}')\mathbf{X}_1N_1, \quad (2.5b)$$

where $\mathbf{E}_{01}$ is the external wave amplitude, $\mathbf{k}_0$ is the corresponding wavevector, $|\mathbf{k}_0| \equiv k_0 = \omega/c$, $\mathbf{r}_1$ is the radius vector of the ball probe center (Fig. 1a), and $\hat{a}_R$, $\hat{a}_{R1}$, $\hat{a}_T$, $\hat{a}_{T1}$ are geometrical factors calculated in the Appendix.

Equations (2.3)–(2.5) form a closed system of equations. The solution of this system describes the self-consistent interaction of the probe with a semiinfinite dielectric medium and allows the fields to be calculated

at various observation points in the near-field and wave zones.

3. STATIONARY SOLUTION OF THE EQUATIONS OF MOTION

In the field of continuous optical radiation, the induced dipole momenta and inversion variations are compensated by relaxation processes. For this reason, Eqs. (2.3) and (2.4) will be solved under the conditions

$$\frac{d\mathbf{X}_1}{dt} = \frac{d\mathbf{X}}{dt} = 0, \quad \frac{dw_1}{dt} = \frac{dw}{dt} = 0, \tag{3.1}$$

which lead, in the general case, to a system of nonlinear algebraic equations.

Let us introduce the quantum polarizabilities

$$\alpha_1 = \frac{e^2}{m} f_1 \frac{1}{\omega_{01}(\Delta_1 - i\gamma_1/2)}, \qquad \alpha = \frac{e^2}{m} f \frac{1}{\omega_0(\Delta - i\gamma/2)}, \tag{3.2}$$

where e and m are the electron charge and mass, respectively;

$$f_1 = \frac{2m\omega_{01}}{\hbar e^2}|\mathbf{d}_{01}|^2, \quad f = \frac{2m\omega_0}{\hbar e^2}|\mathbf{d}_0|^2 \tag{3.3}$$

are the oscillator strengths of the transitions in the probe and the medium, respectively; and γ_1 and γ are the spectral line widths for atoms (molecules) in the probe and the medium, respectively, which obey the relations

$$\frac{\gamma_1}{2} = \frac{1}{T_{21}'}, \quad \frac{\gamma}{2} = \frac{1}{T_2'}. \tag{3.4}$$

Consider the case when the induced dipole moments in the probe and in the medium are linear functions of the field:

$$\mathbf{X}_1 = -w_1\alpha_1\mathbf{E}_{01}, \quad \mathbf{X} = -w\alpha\mathbf{E}_0, \tag{3.5}$$

which is possible provided that $w_1 \approx w_{01}$ and $w \approx w_0$. Indeed, substituting (3.5) into Eqs. (2.3b) and (2.4b), we obtain

$$w_1 = \frac{w_{01}}{T_{11}\left(\frac{2}{\hbar}|\mathbf{E}_{01}|^2 \mathrm{Im}\alpha_1 + \frac{1}{T_{11}}\right)}. \tag{3.6}$$

As can be seen, the inversion w_1 will only slightly differ from the equilibrium value if

$$\frac{2}{\hbar}|\mathbf{E}_{01}|^2 \mathrm{Im}\alpha_1 \ll \frac{1}{T_{11}}, \tag{3.7}$$

where $|\mathbf{E}_{01}|^2$ is determined using (2.5). Thus, in the linear approximation, the induced dipole moments will be calculated using Eqs. (2.3a) and (2.4a) with inversions ω_1 and ω replaced by the corresponding equilibrium values.

Using Eqs. (2.3a) and (2.4a) and taking into account conditions (3.1), we obtain the following expressions,

$$X_1^y = E_{0I}^y \frac{w_{10}\alpha_1(1 + w_0\alpha N a_T^y) - w_{10}\alpha_1 w_0\alpha N a_R^y}{-(1 + w_{10}\alpha_1 N_1 a_{T1})(1 + w_0\alpha N a_T^y) + w_0\alpha N a_R^y w_{10}\alpha_1 N_1 a_{R1}^y} \equiv \alpha_{\mathrm{eff}1}^y E_{0I}^y,$$
$$X^y = E_{0I}^y \frac{w_1\alpha_1(1 + w_{10}\alpha_1 N_1 a_{T1}) - w_{10}\alpha_1 w_0\alpha N_1 a_{R1}^y}{-(1 + w_{10}\alpha_1 N_1 a_{T1})(1 + w_0\alpha N a_T^y) + \alpha w_0 N a_R^y \alpha_1 w_{10} N_1 a_{R1}^y} \equiv \alpha_{\mathrm{eff}}^y E_{0I}^y, \tag{3.8}$$

where $\alpha_{\mathrm{eff}1}^y$ and α_{eff}^y are the effective polarizabilities of the probe and the medium for the induced dipole moments polarized perpendicularly to the plane of incidence xz. For the polarization vectors lying in the incidence plane, we obtain

$$X_1^\beta = E_{0I}^\beta \frac{w_{10}\alpha_1(1 + w_0\alpha N a_T^\beta) - w_{10}\alpha_1 w_0\alpha N a_R^\beta}{-(1 + w_{10}\alpha_1 N_1 a_{T1})(1 + w_0\alpha N a_T^\beta) + w_0\alpha N a_R^\beta w_{10}\alpha_1 N_1 a_{R1}^\beta} \equiv \alpha_{\mathrm{eff}1}^\beta E_{0I}^\beta,$$
$$X^\beta = E_{0I}^\beta \frac{w_1\alpha_1(1 + w_{10}\alpha_1 a_{T1} N_1) - w_0\alpha w_{10}\alpha_1 N_1 a_{R1}^\beta}{-(1 + w_{10}\alpha_1 N_1 a_{T1})(1 + w_0\alpha N a_T^\beta) + w_0\alpha N a_R^\beta w_{10}\alpha_1 N_1 a_{R1}^\beta} \equiv \alpha_{\mathrm{eff}}^\beta E_{0I}^\beta, \tag{3.9}$$

where $\beta = x, z$. The geometrical factors a_{T1} and $\hat{a}_{R1}$ in Eqs. (3.8) and (3.9) are as follows:

$$a_T^y = -\frac{\pi}{n^2 - 1}\frac{\sin(\varphi + \vartheta_T)}{\cos\varphi \sin\vartheta_T} + \frac{2\pi}{3}\frac{n^2 + 2}{n^2 - 1},$$

$$a_T^x = \frac{\pi}{n^2-1}\frac{\sin(\varphi+\vartheta_T)}{\cos\varphi\sin\vartheta_T}(\sin\varphi\cos\varphi-\cos^2\varphi) + \frac{2\pi}{3}\frac{n^2+2}{n^2-1},$$

$$a_T^z = \frac{\pi}{n^2-1}\frac{\sin(\varphi+\vartheta_T)}{\cos\varphi\sin\vartheta_T}(\sin\varphi\cos\varphi-\sin^2\varphi) + \frac{2\pi}{3}\frac{n^2+2}{n^2-1}, \quad (3.10)$$

$$a_R^y = -\frac{\pi}{n^2-1}\frac{\sin(\varphi-\vartheta_T)}{\cos\varphi\sin\vartheta_T},$$

$$a_R^x = \frac{\pi}{n^2-1}\frac{\sin(\varphi-\vartheta_T)}{\cos\varphi\sin\vartheta_T}(\sin\vartheta_R\cos\vartheta_R-\cos^2\vartheta_R),$$

$$a_R^z = \frac{\pi}{n^2-1}\frac{\sin(\varphi-\vartheta_T)}{\cos\varphi\sin\vartheta_T}(\sin\vartheta_R\cos\vartheta_R-\sin^2\vartheta_R).$$

These quantities are determined using the angles φ, ϑ_T, and ϑ_R as described in the Appendix (see Eqs. (A.14) and (A.17), respectively). For the Fresnel reflection and refraction of an external wave incident at an angle ϑ_I onto the surface Σ (Fig. 1a),

$$\varphi = \vartheta_I, \quad n\sin\vartheta_T = \sin\vartheta_I, \quad \vartheta_R = \pi-\vartheta_I. \quad (3.11)$$

Below we will consider the dispersion dependences of the effective polarizabilities (3.8) and (3.9) for various angles φ, ϑ_T, and ϑ_R.

4. OPTICAL FIELD IN THE NEAR-FIELD AND WAVE ZONE RELATIVE TO THE SURFACE OF A SEMIINFINITE MEDIUM

Using induced dipole moments (3.8) and (3.9), it is possible to determine the field at any observation point. Indeed, for any **r** outside the probe and medium, Eq. (2.1a) yields

$$\mathbf{E}_0(\mathbf{r}) = \mathbf{E}_{0I}\exp(i\mathbf{k}_0\cdot\mathbf{r}) + \hat{a}_R(\mathbf{r})\mathbf{X}N + \hat{a}_{R1}(\mathbf{r})\mathbf{X}_1N_1. \quad (4.1)$$

Here

$$\hat{a}_R(\mathbf{r}) = \hat{a}_R(0)\exp(ik_0\mathbf{r}\cdot\mathbf{s}_R),$$

$\hat{a}_R(0)$ is given by formulas (3.10), and the geometrical factor $\hat{a}_{R1}(\mathbf{r})$ is determined from the relation

$$N_1\hat{a}_{R1}(\mathbf{r})\mathbf{X}_1 = \mathrm{curl}\,\mathrm{curl}\int_{\Sigma_1}\left(\mathbf{Q}_1\frac{\partial G_p}{\partial\nu_1'}-G_p\frac{\partial\mathbf{Q}_1}{\partial\nu'}\right)dS_1', \quad (4.2)$$

where $G_p = \exp(ik_0R_1')/R_1'$ is the Green function dependent on the observation point **r**. The surface integral is calculated in the Appendix. In a particular case of $\mathbf{r}\parallel\mathbf{z}$, this integral is determined by formula (A.14), where calculation of the (curl curl) operator reduces to taking the second derivative with respect to z_0 of the surface integral (4.2). For an arbitrary observation point, the geometrical factor $\hat{a}_{R1}(\mathbf{r})$ is determined by formula (4.2) with the (curl curl) operator in the general form. Then, Eq. (4.1) gives expressions for the field in the near-field and wave zones.

Let us consider the optical field in the wave zone for $k_0r \gg 1$. In this case, we may retain only the terms proportional $1/R_1'$ in the geometrical factor $\hat{a}_{R1}(\mathbf{r})$ (R_1' is the distance from the probe center to the observation point **r**), which yields

$$\hat{a}_{R1}\mathbf{X}_1 = \frac{2\pi a_1}{n_1^2-1}\frac{1}{R_1'} \times[\mathbf{n}_1'(\mathbf{n}_1'\cdot\mathbf{X}_1)-\mathbf{X}_1]\exp(ik_0R_1'), \quad (4.3)$$

where $\mathbf{n}_1' = \mathbf{R}_1'/R_1'$.

By the same token, the magnetic field strength in the wave zone for $k_0r \gg 1$ is determined using Eq. (2.1b). According to this, vector **H** has the following form (the time factor $\exp(-i\omega t)$ is omitted):

$$\mathbf{H}_0(\mathbf{r}) = \mathbf{H}_{0I}\exp(i\mathbf{k}_0\cdot\mathbf{r}) + \hat{b}_R(\mathbf{r})\mathbf{X}N + \hat{b}_{R1}(\mathbf{r})\mathbf{X}_1N_1, \quad (4.4)$$

where $\hat{b}_R$ and $\hat{b}_{R1}$ are the geometrical factors given by the formulas

$$\hat{b}_R\mathbf{X} = -\frac{\pi}{n^2-1}\frac{\sin(\varphi-\vartheta_T)}{\cos\varphi\sin\vartheta_T}\mathbf{s}_R\times\mathbf{X}\exp(ik_0\mathbf{r}\cdot\mathbf{s}_R),$$
$$\hat{b}_{R1}\mathbf{X}_1 = -\frac{2\pi a_1}{n^2-1}\frac{1}{R_1'}\mathbf{n}_1'\times\mathbf{X}_1\exp(ik_0R_1'). \quad (4.5)$$

The energy flux via an element of the spherical surface $\Delta\sigma = (R_1')^2\Delta\Omega$ per unit time is determined using

expressions (4.1) and (4.4) for the field vectors:

$$S\Delta\sigma = \frac{c}{4\pi}|\mathbf{E}_0 \times \mathbf{H}_0|(R_1')^2\Delta\Omega. \quad (4.6)$$

In the wave zone ($k_0 R_1' \gg 1$), the fields are determined in terms of geometrical factors (4.3) and (4.5) and effective polarizabilities (3.8) and (3.9), which depend on the near-field geometrical factors (3.10), (A.14), and (A.17).

In the above formulas, the refractive indices n_1 and n can be determined using Eq. (2.1a) by placing the observation point inside and outside the dielectric medium. Separating the terms determined for the corresponding points, we obtain

$$\frac{n_1^2 - 1}{n_1^2 - 2} = \frac{4\pi}{3}N_1\alpha_1, \quad \frac{n^2 - 1}{n^2 - 2} = \frac{4\pi}{3}N\alpha, \quad (4.7)$$

where the polarizabilities are given by formulas (3.2). These expressions correspond to the case of a closed Lorentz sphere containing the observation point. In a strict theory, allowance of the discrete atoms distributed inside a truncated Lorentz sphere leads to the near-field effect [15]. In our study, the role of this effect is negligibly small.

5. LINEAR STATIONARY OPTICAL SIZE RESONANCES IN THE SYSTEM OF A BALL PROBE OVER A SEMIINFINITE OPTICAL MEDIUM

Let us study the dispersion dependences of effective polarizabilities (3.8) and (3.9) in various physical situations. First, consider a probe consisting of a glass ball containing impurity sodium atoms characterized by a resonance absorption at $\lambda = 589$ nm due to the transition from ground state 3*S* to excited state 3*P* (yellow emission line of sodium atom). The natural width of the 3*S*–3*P* transition is 10 MHz, the dipole moment of this transition is $d_{01} = 6.236 \times 10^{-18}$ esu, and the transition frequency is $\omega_{01}/c = 1.066 \times 10^5$ cm^{-1}. The corresponding refractive index n_1 is given by the formula

$$\frac{n_1^2 - 1}{n_1^2 + 2} = \frac{4\pi}{3}(N_{01}\alpha_{01} + N_1\alpha_1), \quad (5.1)$$

where N_{01} and α_{01} are the density and polarizability of atoms (molecules) of the glass. Let us assume that

$$N_{01}\alpha_{01} \approx N_1\alpha_1 \quad (5.2)$$

and the refractive index ($n_1 \approx 1.5$) remains constant within a wide frequency range. We consider a probe with a radius of $a_1 = 20$ nm and an impurity (sodium) concentration of $N_1 = 10^{18}$–10^{21} cm^{-3}. The density of atoms in the glass is $N_{01} = 10^{22}$ cm^{-3}.

Consider a sample representing a semiinfinite, isotropic optical medium in the vicinity of a small resonance at a wavelength of $\lambda_0 = 694.3$ nm (R_1 line of a ruby crystal at $T = 300$ K) with a transition frequency of $\omega/c = 0.905 \times 10^5$ cm^{-1}, a dipole moment of the transition of $d = 1.8 \times 10^{-20}$ esu, and a spectral line width of $(T_2')^{-1} = 3 \times 10^{11}$ Hz. Let the concentration of impurity (Cr^{3+}) ions in the medium be $N = 2 \times 10^{19}$ cm^{-3}, which corresponds to a pink ruby [21].

The atoms of impurities in the probe and in the sample are immersed in the media, the influence of which is manifested as follows. First, surrounding atoms modify the natural frequencies, dipole moments of the transitions, and spectral line widths of the impurity atoms. Second, the media surrounding impurity atoms produce an additional polarizing action upon the probe–sample interaction. In the boundary-value problem under consideration, this action is taken into account by adding vectors $\mathbf{P}' = N_0\alpha_0\mathbf{E}$ and $\mathbf{P}_1' = N_{01}\alpha_{01}\mathbf{E}$ to the polarization vectors $\mathbf{P}$ and $\mathbf{P}_1$ for the observation points inside the sample and probe, respectively. With allowance for the linearity of Eqs. (2.3) and (2.4) and the expressions for the effective fields (2.5), this transformation of the polarization vectors results in that the solution of Eqs. (3.8) and (3.9) will acquire the corresponding corrections to the acting field, while the geometrical factors $\hat{a}_{R1}$ and $\hat{a}_R$ have to be determined using refractive indices given by formulas (5.1) and

$$\frac{n^2 - 1}{n^2 + 2} = \frac{4\pi}{3}(N_0\alpha_0 + N\alpha), \quad (5.3)$$

where N_0 and α_0 are the density and polarizability of atoms (molecules) surrounding the two-level impurity atoms in the sample.

The fields acting at the center of the probe, determined with allowance for the polarizing action of the nonresonant atoms, are as follows:

$$E_{01}^x = \frac{a_{22}^x \exp(i\mathbf{k}_0 \cdot \mathbf{r}_1) - a_{12}^x}{a_{11}^x a_{22}^x - a_{12}^x a_{21}^x}E_{0I}^x,$$
$$E_0^x = \frac{a_{11}^x - a_{21}^x \exp(i\mathbf{k}_0 \cdot \mathbf{r}_1)}{a_{11}^x a_{22}^x - a_{12}^x a_{21}^x}E_{0I}^x, \quad (5.4)$$

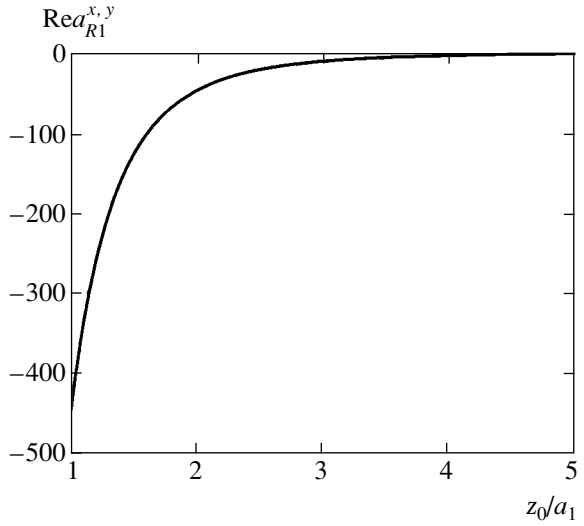

Fig. 2. Plot of the real part of the geometrical factor $a_{R1}^{x,y}$ of a ball probe versus the relative distance z_0/a_1 from the probe center to the sample surface, calculated for $n_1 = 1.5$, $k_0 = 1.066 \times 10^5$ cm^{-1}, and $a_1 = 20$ nm.

where

$$\begin{aligned} a_{11}^x &= 1 - a_{T1}^x N_{01}\alpha_{01} + a_{T1}^x w_1 \alpha_1 N_1, \\ a_{12}^x &= a_R^x w\alpha N - a_R^x N_0 \alpha_0, \\ a_{21}^x &= a_{R1}^x w_1 \alpha_1 N_1 - a_{R1}^x N_{01}\alpha_{01}, \\ a_{22}^x &= 1 - a_T^x N_0 \alpha_0 + a_T^x w\alpha N. \end{aligned} \tag{5.5}$$

In order to determine the other components of the acting fields, it is necessary to determine coefficients (5.5) using the corresponding geometrical factors calculated using expressions (3.10), (A.14), and (A.17). The induced dipole moments of the resonance atoms acquire the following form:

$$\begin{aligned} \mathbf{X}_1 &= -w_1 \alpha_1 \mathbf{E}_{01} \equiv \hat{\alpha}_{\text{eff}1} \mathbf{E}_{0I}, \\ \mathbf{X} &= -w\alpha \mathbf{E}_0 \equiv \hat{\alpha}_{\text{eff}} \mathbf{E}_{0I}, \end{aligned} \tag{5.6}$$

which, in contrast to expressions (3.8) and (3.9), takes into account the polarizing action of nonresonant sub-

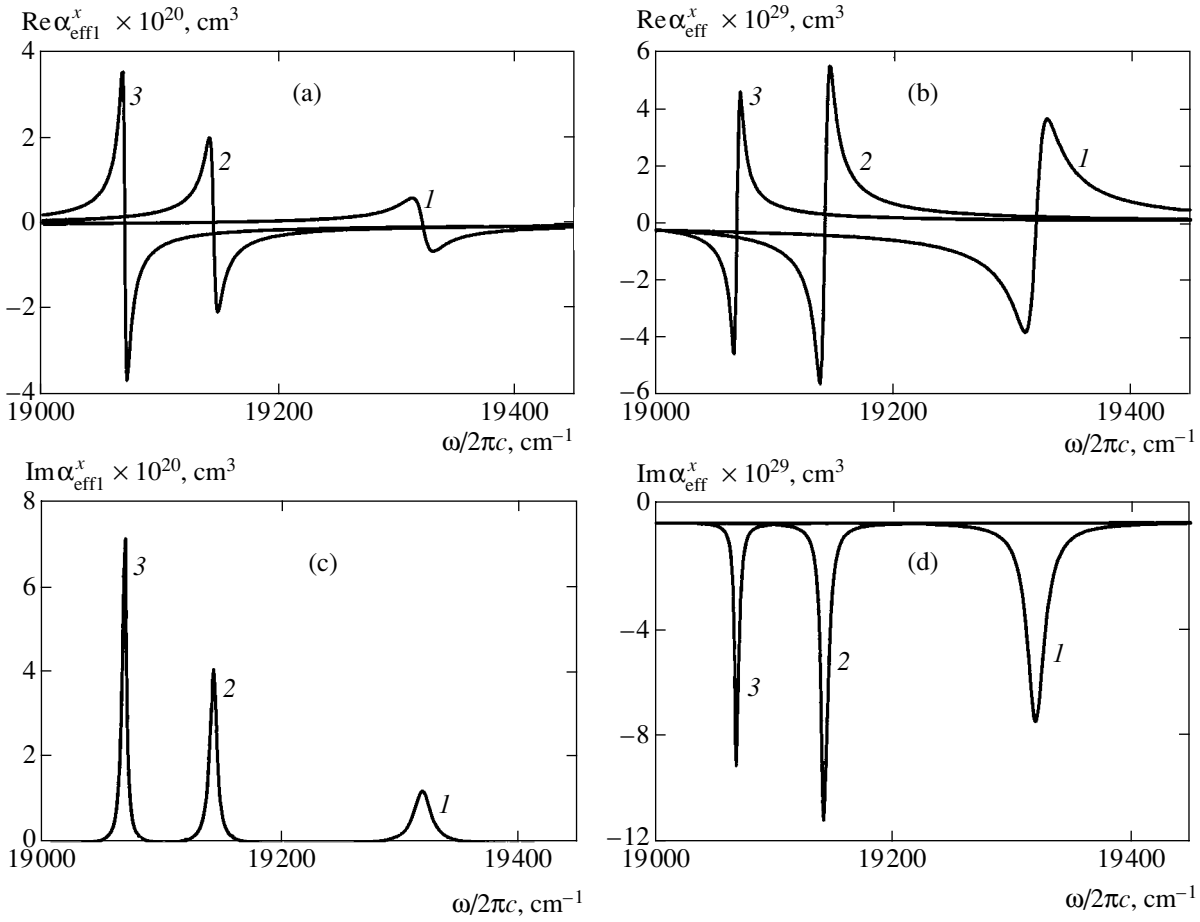

Fig. 3. Typical dispersion dependences of the real and imaginary parts of the effective polarizabilities of (a, c) two-level Na atoms in the probe and (b, d) Cr^{3+} ions in the sample, calculated with allowance for the polarizing action of the glass and corundum, respectively, for probe–sample distances of $z_0 = 30$ (*1*), 50 (*2*), and 60 nm (*3*); and $n_{01} = 1.5$, $n_0 = 1.78$, and $a_1 = 20$ nm.

systems both in the probe and in the sample. Figure 2 shows a plot of the geometrical factor $a_{R1}^{x,y}$ versus the relative distance z_0/a_1 from the top of the probe center to be sample surface. For small distances, such that $k_0 z_0 \ll 1$ (e.g., in the near-field zone relative to the surface of the semiinfinite medium), the main role is played by terms proportional to $1/z_0^4$ (rather than $1/z_0^3$, as in the case of point dipoles [15]). At large distances from the probe, such that $k_0 z_0 \gg 1$, the main role is played by terms proportional to $1/z_0$.

Figure 3 presents the dispersion dependences for the effective polarizabilities of two-level Na atoms in the probe and Cr^{3+} ions in the sample, calculated with allowance for the polarizing action of the glass and corundum, respectively.

For external radiation polarized in the incidence plane *xz*, the probe–sample system exhibits a single optical size resonance at a frequency significantly different from the natural frequencies of both Na atoms and Cr^{3+} ions (with neglect of the interaction). The optical size resonance frequency depends strongly on the distance z_0 from the probe center to the sample. Indeed, the frequency is $\omega/2\pi c = 19320$ cm^{-1} for $z_0 = 30$ nm and decreases to 19140 cm^{-1} for $z_0 = 50$ nm and to 19070 cm^{-1} for $z_0 = 70$ nm (Fig. 3). For external radiation polarized perpendicularly to the incidence plane, the optical size resonance frequencies are the same: in a system of the ruby–glass probe with sodium, the effect is independent of polarization. According to the results of calculations, a change in the spectral line widths of Na atoms and Cr^{3+} ions leads to broadening of the optical size resonances, but the resonance frequencies remain unchanged.

6. NEAR-FIELD OPTICAL MICROSCOPY IN THE REGION OF CONTINUOUS SPECTRA

Thus, we have solved the boundary-value problem corresponding to the scheme of Fig. 1a, in which a dielectric ball probe interacts with the flat surface of a semiinfinite dielectric optical medium. The spectra of both probe and sample contain narrow lines corresponding to two-level impurity atoms. Using expressions (5.4) and (5.6), it is possible to take into account the polarizing action of the probe and the sample. There are various possible situations in which the polarizing effects on the optical fields in the near-field and wave zones of impurity atoms and the environment are either comparable or the polarizing action of a nonresonant surrounding is significantly greater than that of the impurity. In the latter case, the probe–sample interaction is determined by the static polarizabilities of atoms in the probe and the sample.

The interaction of a sample and a probe made of semiconductor materials can also be described based on the solution of the boundary-value problem, with allowance for the corresponding quantum polarizabilities. In the approximation of a homogeneously broadened line, the quantum polarizability is given by Eq. (3.2) [22, 23], where $d_{0(1)}$ are the dipole moments of the recombination transitions in the sample and probe and $\omega_{0(1)}$ are the frequencies of transitions between levels in the corresponding two-level systems. For a bulk GaAs sample, the dipole moment of the transition is about 1.5×10^{-17} esu [24]. Although the probe represents a quantum dot, it is still possible to assume that the dipole moment of the transition coincides with that in the bulk semiconductor. According to [24], the homogeneous broadening for a bulk semiconductor at room temperature amounts to $2\pi/T_2' \approx 10^{13}$ s^{-1}. The main contribution to this value is due to the interaction between charge carriers. For a quantum dot, we may expect that the corresponding $2\pi/T_{21}'$ value is determined by the electron–phonon and hole–phonon interactions. According to the theory of semiconductor lasers, the carrier–phonon interaction is one order of magnitude smaller than that between the charge carriers; therefore, we can assume that $2\pi/T_{21}' \approx 10^{12}$ s^{-1}.

However, we must also take into account the line broadening related to the probe size scatter in various directions. The positions of energy levels for a spherical quantum dot depend on the radius a_1 as [25]

$$E_{nk} = \frac{\hbar^2}{2ma_1^2}\alpha_{nk}^2, \tag{6.1}$$

where α_{nk} is the *k*th root of the Bessel function $J_{n+1/2}(z)$. This formula can be obtained by using a stationary spherically symmetrical solution of the Schrödinger equation [26]. The uncertainty ΔE_{nk} of the energy levels is proportional to the uncertainty Δa_1 of the quantum dot radius, so that

$$\Delta E_{nk}^{e,h} = \frac{\hbar^2}{m_{e,h}a_1^2}\alpha_{nk}^2\frac{\Delta a_1}{a_1}, \tag{6.2}$$

where $m_{e,h}$ are the effective masses of a hole and electron, respectively. The inhomogeneous line broadening,

$$\frac{\Delta E_{01}}{\hbar} = \frac{\Delta E_{01}^e + \Delta E_{01}^h}{\hbar},$$

is the sum of two terms representing the contributions from the electron and hole levels. Assuming that the linear dimensions of a quantum dot can be set to within 99% ($\Delta a_1/a_1 \approx 0.01$), we obtain for GaAs ($m_e = 0.067m$, $m_h = 0.45m$, m being the mass of free electron) an esti-

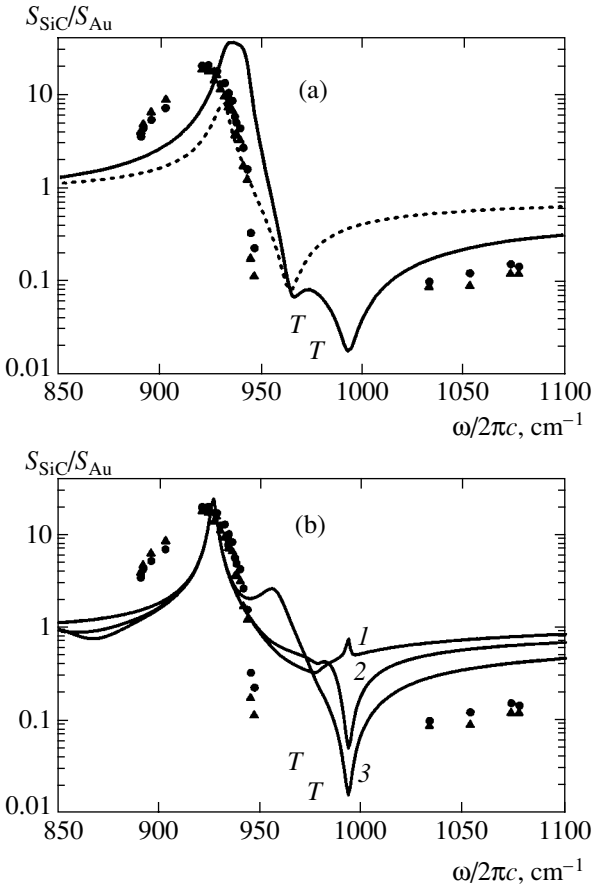

Fig. 4. Dispersion dependences of the backscattered signal in the system of a Pt probe over a semiinfinite medium (S_{SiC}/S_{Au} is the ratio of the signals scattered from a SiC substrate and Au islands): (a) experimental and theoretical data from [27]; (b) experimental data from [27] compared to the results of calculations according to formula (4.6) with z_0 = 30 (*1*), 50 (*2*), and 80 nm (*3*); "*T*" indicates experimental points on the background level.

mate of $\Delta E_{01}/\hbar = 10^{12}$ s^{-1}, which is comparable with the homogeneous broadening $2\pi/T'_{21}$.

Now let us consider the interaction of a metal probe with the surface of a semiinfinite dielectric medium. Hillenbrand *et al.* [27] experimentally studied light backscattering from a sample with gold islands. A ball probe with a radius of a_1 = 10–30 nm was made of gold, silver, or platinum, while the sample was made of polar dielectric SiC featuring residual rays in a range from 790 to 950 nm. In this frequency interval, the reflectances of massive gold and SiC are frequency-independent and amount to approximately 0.99 and 0.97, respectively [27]. For a 10-nm gold ball probe, the transition frequency according to formula (6.1) is $E_{01}/\hbar = 1.44 \times 10^{12}$ s^{-1}.

In the range of residual rays for SiC, the quantum polarizability of the probe is virtually independent of frequency and can be determined using the formula [20]

$$\alpha_1(a_1, \omega) = 12\pi\varepsilon_0 a_1^3 \int_0^1 \frac{\varepsilon(R_1, \omega) - 1}{\varepsilon(R_1, \omega) + 2} R_1^2 dR_1, \qquad (6.3)$$

where

$$\varepsilon(R_1, \omega) = 1 - n(R_1)\frac{\omega_p^2}{\omega(\omega + i\Gamma)}, \qquad (6.4)$$

ω_p is the plasma frequency, Γ is a phenomenological damping constant, and $n(R_1)$ is the normalized electron density in the ball. The latter quantity is determined as

$$n(R_1) = \frac{4}{3N_1V_1}\sum_{k,l}(2l+1)\left(\frac{J_l(\alpha_{kl}R_1)}{J_{l+1}(\alpha_{kl})}\right)^2, \qquad (6.5)$$

where J_l, J_{l+1} are the Bessel functions of the *l*th and $(l + 1)$th order, α_{kl} is the *k*th root of the *l*th-order Bessel function, N_1V_1 is the number of electrons in the ball, R_1 is the distance from the ball center normalized to the radius a_1, and $\varepsilon_0 = 1/4\pi$ is the dielectric constant. Thus, the probe–sample interaction in the near-field microscopy experiments described in [27] takes place in the region of a continuous spectrum with respect to the probe.

Figure 4a presents the experimental data obtained in [27] on the frequency dependence of a normalized signal $|S_{SiC}/S_{Au}|$ of dipole radiation from a SiC surface partly covered with gold islands. A plane wave of the external radiation was incident on the sample at an angle of $\vartheta_I = \pi/4$, and the response signal was detected in the direction opposite to that of the incident wave. The probe radius was a_1 = 20 nm and the external wave was polarized in the incidence plane *xz* (Fig. 1a). The distance between the probe and the sample surface was varied from 0 to 40 nm [27]. In Fig. 4a, the solid curve shows the results of theoretical calculations performed in [27] based on the theory developed in [5–14], according to which $S_{SiC} \sim |\alpha_{eff}|^2$, where

$$\alpha_{eff} = \frac{\alpha(1+\beta)}{1 - \alpha\beta/16\pi(a_1 + z_0)^3}. \qquad (6.6)$$

Here (in notations of [27]),

$$\beta = \frac{\varepsilon_s - 1}{\varepsilon_s + 1},$$

ε_s is the dielectric permittivity of the sample,

$$\alpha = 4\pi a_1^3(\varepsilon_p - \varepsilon_m)(\varepsilon_p + 2\varepsilon_m)$$

is the polarizability of a spherical probe, and ε_m is the permittivity of the probing metal. In Fig. 4a, the symbols (triangles and circles) correspond to various regions on the SiC sample surface surrounded by gold islands. As can be seen from these data, the theoretical interpretation proposed in [27] provides for a rather rough approximation. The dashed curve in Fig. 4a shows the signal of dipole radiation from a platinum probe near the SiC surface.

Figure 4b shows the frequency dependences of the normalized signal $|S_{\mathrm{SiC}}/S_{\mathrm{Au}}|$ for various distances z_0 from the probe center to the SiC sample surface, which were calculated using the theory proposed in this paper. These curves were constructed using formulas (4.6), (4.7), (3.9), (3.2), and (6.3); in the region of the continuous spectrum of the probe, the refractive index is $n_1 = 2.63 + 3.54i$, while the frequency dependence of the dielectric permittivity of SiC is described by the expression

$$\varepsilon = \varepsilon_\infty + \frac{4\pi\rho\omega_0^2}{\omega_0^2 - \omega^2 - i\Gamma\omega}, \tag{6.7}$$

where $\varepsilon_\infty = 6.49$, $\omega_0 = 788$ cm^{-1} is the transition frequency, $\rho = 0.257$ is the oscillator strength for the phonon transition, and $\Gamma = 6.8 \times 10^{-3}$ cm^{-1}. The refractive index n in (3.10) is determined using formula (4.7), where n^2 is replaced by ε.

As can be seen in Fig. 4b, the signal intensity reaches a maximum at 927 cm^{-1}, which coincides to a high precision with the experimental value. According to our theory, this peak corresponds to the optical size resonance formed as a result of the self-consistent interaction of the platinum ball probe with the SiC surface. For narrow resonances, even a small change in the distance between the interacting dipoles leads to a significant shift in the optical size resonance frequency. This takes place for the interacting atoms considered as point dipoles [15]. The same situation is observed for a probe–sample system possessing narrow resonances (see Section 5). In the experiments reported in [27], the probe possessed a continuous spectrum and, hence, the optical size resonance at $\omega/2\pi c = 927$ cm^{-1} was less sensitive to a change in the distance z_0 from the probe center to the sample surface. However, as can be seen from Fig. 4b, an additional optical size resonance appears for $z_0 = 30$ nm at $\omega/2\pi c = 990$ cm^{-1}. Since the frequency interval from 950 to 1030 cm^{-1} was not studied in [27], it is impossible to check for the presence of this resonance in the experimental spectrum.

A comparison of our theory to the experimental data presented in Fig. 4a suggests that a polarizing action of the edges of gold islands is rather insignificant for frequencies close to resonance, but this action can make a significant contribution to the resulting response at small amplitudes of the signal scattered from the sample surface. This probably explains the discrepancy between experiment [27] and the results of our calculations using formula (4.6) in the region of 1000–1100 cm^{-1}. The polarizing action can be readily taken into account based on the proposed theory, but this requires more detailed information about the shape of island edges. In our opinion, the significant difference in resonance width between the theoretical and experimental curves can be explained by different values of the relaxation times used in [27] and in our calculations, but this question requires additional investigation.

Thus, we have developed an electrodynamic theory of near-field optical microscopy based on the description of a self-consistent interaction between a ball probe and a flat sample surface. The proposed theory is applicable to the probes and samples of various materials in the regions of both discrete and continuous spectra. The presence of foreign particles or islands on the sample surface can be taken into account by introducing those corresponding volume integrals into the theory that take into account the polarizing action of such objects.

ACKNOWLEDGMENTS

This study was supported by the Ministry of Education of the Russian Federation, project no. RD02-1.2-68.

APPENDIX

Calculation of Geometrical Factors

Let us calculate the volume integral entering Eqs. (2.5),

$$\begin{aligned}\mathbf{P}_{0M}(\mathbf{r}) &= \int \mathrm{curl}\,\mathrm{curl}\,\mathbf{P}_0(\mathbf{r}')G_M(R)dV' \\ &\equiv N\hat{a}_T\mathbf{X}\exp(-i\omega t),\end{aligned} \tag{A.1}$$

for the observation points inside the medium, where

$$\mathbf{P}_0 = \frac{1}{2}N\mathbf{X}, \quad G_M = \frac{\exp(ik_0R')}{R'}.$$

Taking into account that

$$\mathbf{P}_0 = (n^2 - 1)k_0^2\mathbf{Q}, \quad k_0 = \frac{\omega}{c},$$

where n is the refractive index of the medium and $\mathbf{Q}$ is the function of coordinates which obeys the equations

$$\nabla^2\mathbf{Q} + n^2k_0^2\mathbf{Q} = 0, \quad \mathrm{div}\mathbf{Q} = 0,$$

and using a lemma [28] for exporting the (rotrot) operator out of the integral and the Green theorem, we obtain

$$\mathbf{P}_{0M} = \mathrm{curl}\,\mathrm{curl}\,\mathbf{I}_\Sigma + \frac{4\pi}{3}(n^2+2)k_0^2\mathbf{Q}. \qquad (A.2)$$

Here, the surface integral is expressed as

$$\mathbf{I}_\Sigma = \int_\Sigma \left(\mathbf{Q}\frac{\partial G_M}{\partial \nu'} - G_M\frac{\partial \mathbf{Q}}{\partial \nu'}\right)dS', \qquad (A.3)$$

and $\partial/\partial\nu'$ denotes differentiation along the external normal to the surface Σ. The value of $\mathbf{P}_{0M}$ in Eq. (A.2) is determined at observation point $\mathbf{r}$ inside or on the surface of the medium. If the point is outside the medium, for example, inside the probe, then

$$\mathbf{P}'_{0M} = \mathrm{curl}\,\mathrm{curl}\,\mathbf{I}'_\Sigma, \qquad (A.4)$$

where $\mathbf{I}'_\Sigma$ differs from $\mathbf{I}_\Sigma$ by the position of the observation point.

Let $\mathbf{Q}$ have the following form:

$$\mathbf{Q} = \mathbf{Q}_0\exp(ik_0n\mathbf{r}\cdot\mathbf{s}_T), \qquad (A.5)$$

where $\mathbf{Q}_0$ is a constant vector and $\mathbf{s}_T$ is a unit vector determined by the refraction angle ϑ_T in the incidence plane xz. Introducing a unit vector $\mathbf{s}$, such as

$$s_x = -\sin\varphi, \quad s_y = 0, \quad s_z = -\cos\varphi,$$

the surface integral (A.3) can be written as

$$\mathbf{I}_\Sigma = I_\Sigma^{(0)}\mathbf{Q}_0\exp(ik_0\mathbf{r}\cdot\mathbf{s}), \qquad (A.6)$$

where

$$I_\Sigma^{(0)} = -2\pi\frac{\sin(\varphi+\vartheta_T)}{\cos\varphi\sin\vartheta_T}. \qquad (A.7)$$

As was demonstrated in [15], the surface integral has the same value on the surface Σ and in the vicinity of this surface inside the medium. Assuming the surface Σ to be homogeneous, so that

$$\mathbf{s} = \mathbf{s}_I, \quad \varphi = \vartheta_I,$$

and considering a particular case of the normal incidence ($\vartheta_I = 0$), we obtain

$$I_\Sigma^{(0)} = -2\pi(n+1).$$

Substituting this value into (A.2), we obtain the following expression (valid on the surface Σ):

$$\mathbf{P}_{0M} = \mathbf{P}_0\frac{4\pi}{n^2-1}\left(\frac{n^2+2}{3} - \frac{n+1}{2}\right) \equiv \mathbf{P}_0 a_T. \qquad (A.8)$$

For observation points outside the medium, including point $\mathbf{r}_1$ (Fig. 1a), Eq. (A.4) yields

$$\mathbf{P}_{0M} = -2\pi\mathbf{P}_0\frac{n+1}{n^2-1} \equiv \mathbf{P}_0 a_R. \qquad (A.9)$$

For observation points outside the medium ($z > 0$), the surface integral $\mathbf{I}'_\Sigma$ is calculated similarly to (A.6). The difference is that the coordinate z' changes sign, which is equivalent to changing s_z to $-s_z$, or φ to ϑ_R (ϑ_R is the angle of reflection). Now let us introduce a unit vector $\mathbf{s}_R$ (instead of $\mathbf{s}$) such that

$$s_{Rx} = -\sin\vartheta_R, \quad s_{Ry} = 0, \quad s_{Rz} = -\cos\vartheta_R.$$

Then, the surface integral is as follows:

$$\mathbf{I}'_\Sigma = -2\pi\frac{\sin(\varphi-\vartheta_T)}{\cos\varphi\sin\vartheta_T}\mathbf{Q}\exp(ik_0\mathbf{r}\cdot\mathbf{s}_R). \qquad (A.10)$$

For $\vartheta_I \neq 0$, the values of $\hat{a}_R$ and $\hat{a}_T$ are determined using general expressions for the integrals (A.6) and (A.10).

Now let us calculate the geometrical factor $\hat{a}_{R1}$ entering expression (2.5b) by placing the observation point at a distance z_0 from the probe center:

$$\int_{V_1}\mathrm{curl}\,\mathrm{curl}\,N_1\frac{\mathbf{p}_1(t-R'_c/c)}{R'_1}dV'_1$$

$$= \exp(-i\omega t)\hat{a}_{R1}N_1\mathbf{X}_1 = \exp(-i\omega t)(n_1^2-1)k_0^2 \qquad (A.11)$$

$$\times\,\mathrm{curl}\,\mathrm{curl}\int_{V_1}\mathbf{Q}_1G_p(R'_1)dV'_1,$$

where $V_1 = (4\pi/3)a_1^3$, a_1 is the probe radius, n_1 is the refractive index of the probe material,

$$N_1\mathbf{p}_1 = \mathbf{P}_1 = (n_1^2-1)k_0^2\mathbf{Q}_1\exp(-i\omega t),$$

and

$$\nabla^2\mathbf{Q}_1 + n_1^2k_0^2\mathbf{Q}_1 = 0, \quad \nabla^2G_p + k_0^2G_p = 0. \qquad (A.12)$$

According to the Green theorem, the volume integral

in (A.11) has the following form:

$$\int_{V_1} \mathbf{Q}_1 G_p(R_1')\,dV_1' = \frac{1}{(n_1^2-1)k_0^2} \times \int_{\Sigma_1}\left(\mathbf{Q}_1\frac{\partial G_p}{\partial \nu'} - G_p\frac{\partial \mathbf{Q}_1}{\partial \nu'}\right)dS_1', \tag{A.13}$$

where $\partial/\partial\nu_1'$ denotes a derivative along the external normal ν_1' to the probe surface Σ_1.

The surface integral in (A.13) can be calculated taking into account that

$$\mathbf{Q}_1 = \mathbf{Q}_{01}\frac{\exp(ik_0 n_1 R_1)}{R_1} \tag{A.14}$$

satisfies wave equation (A.12), where $\mathbf{Q}_{01}$ is a constant vector and R_1 is a point inside the probe. Passing to the spherical coordinates and calculating integral (A.13), we obtain

$$a_{R1}^{x,y} = -\frac{2\pi a_1}{(n_1^2-1)k_0^2}\frac{d^2}{dz_0^2} \times\left[\frac{1}{z_0}\frac{a_1}{\sqrt{a_1^2+z_0^2}}\exp(ik_0\sqrt{a_1^2+z_0^2}) - \frac{1}{z_0}\exp(ik_0\sqrt{a_1^2+z_0^2})\left(n_1+\frac{i}{a_1k_0}\right) - \frac{1}{z_0}\exp(ik_0|a_1-z_0|)\left(1-n_1-\frac{i}{a_1k_0}\right)\right], \quad a_{R1}^{z} = 0. \tag{A.15}$$

By the same token, the geometrical factor $\hat{a}_{T1}$ for the observation points inside the ball probe can be calculated taking into account that

$$\int_{V_1}\mathrm{curl}\,\mathrm{curl}\,N_1\frac{\mathbf{p}_1(t-R_1'/c)}{R_1'}dV_1' = \exp(-i\omega t)\hat{a}_{T1}N_1\mathbf{X}_1 = \exp(-i\omega t)(n_1^2-1)k_0^2\int_{V_1}\mathrm{curl}\,\mathrm{curl}\,\mathbf{Q}_1 G_p(R_1')dV_1'. \tag{A.16}$$

Exporting the (curl curl) operator out of the integral and using the Green theorem, we obtain (A.16) in a different form:

$$\exp(-i\omega t)\,\mathrm{curl}\,\mathrm{curl}\left[\int_{\Sigma_1}\left(\mathbf{Q}_1\frac{dG_p}{dR_1}-G_p\frac{d\mathbf{Q}_1}{dR_1}\right)dS_1' - \int_{\sigma}\left(\mathbf{Q}_1\frac{dG_p}{dR_1}-G_p\frac{d\mathbf{Q}_1}{dR_1}\right)dS_1'\right] - \frac{8\pi}{3}(n_1^2-1)k_0^2\mathbf{Q}_1\exp(-i\omega t), \tag{A.17}$$

where σ_0 is a spherical surface surrounding the observation point at the center of the ball probe. Taking into account the shape of the function $\mathbf{Q}$ (A.14), let us calculate the surface integrals in (A.17). The integral over the external surface Σ_1 of the ball probe is as follows:

$$\int_{\Sigma_1}\left(\mathbf{Q}_1\frac{dG_p}{dR_1}-G_p\frac{d\mathbf{Q}_1}{dR_1}\right)dS_1' = -4\pi\frac{n_1-1}{n_1+1}\mathbf{Q}_{01}\exp(ik_0 z)[\exp(ik_0a_1(n_1+1))-1].$$

An analogous expression for the surface integral over the internal surface σ_0 is obtained by replacing a_1 with L_0, where L_0 is the radius of sphere σ_0. For $k_0a_1 \ll 1$, tensor $\hat{a}_{T1}$ becomes a scalar and, hence,

$$\hat{a}_{T1} = a_{T1} = -\frac{4\pi}{3}. \tag{A.18}$$

REFERENCES

1. R. Hillenbrand and F. Keilmann, Appl. Phys. Lett. **80**, 25 (2002).
2. M. Speht, J. D. Pedaring, W. M. Heckl, and T. W. Hänsch, Phys. Rev. Lett. **68**, 476 (1992).
3. B. Knoll and F. Keilmann, Nature **399**, 134 (1999).
4. W. L. Barnes, J. Mod. Opt. **45**, 661 (1998).
5. L. Salomon, F. De Fornel, and J. P. Goudonnet, J. Opt. Soc. Am. A **8**, 2009 (1991).
6. W. Denk and D. W. Pohl, J. Vac. Sci. Technol. B **9**, 510 (1991).
7. D. Van Lebeke and D. Barchiesi, J. Opt. Soc. Am. A **9**, 732 (1992).
8. S. Berntsen, E. Bozhevolnaya, and S. Bozhevolnyi, J. Opt. Soc. Am. A **10**, 878 (1993).
9. B. Labani, C. Girard, D. Courjon, and D. Van Lebeke, J. Opt. Soc. Am. B **7**, 936 (1990).
10. C. Girard and X. Bouju, J. Chem. Phys. **95**, 2056 (1991).
11. C. Girard and X. Bouju, J. Opt. Soc. Am. B **9**, 298 (1992).
12. C. Girard and D. Courjon, Phys. Rev. B **42**, 9340 (1990).
13. C. Girard, Appl. Opt. **31**, 5380 (1992).

14. C. Girard and M. Sparjes, Appl. Opt. **29**, 3726 (1990).
15. O. N. Gadomskiĭ, Usp. Fiz. Nauk **170**, 1145 (2000) [Phys. Usp. **43**, 1071 (2000)].
16. O. N. Gadomskiĭ and T. T. Idiatullov, Zh. Éksp. Teor. Fiz. **119**, 1222 (2001) [JETP **92**, 1060 (2001)].
17. O. N. Gadomskiĭ and Yu. V. Abramov, Opt. Spektrosk. **93**, 953 (2002) [Opt. Spectrosc. **93**, 879 (2002)].
18. O. N. Gadomskiĭ and K. Yu. Moiseev, Opt. Spektrosk. **92**, 613 (2002) [Opt. Spectrosc. **92**, 560 (2002)].
19. O. Keller, M. Xiao, and S. Bozhevolnyi, Surf. Sci. **280**, 217 (1993).
20. M. Xiao, S. Bozhevolnyi, and O. Keller, Appl. Phys. A **62**, 115 (1996).
21. V. M. Faĭn and Ya. I. Khanin, *Quantum Radiophysics* (Sovetskoe Radio, Moscow, 1965), p. 562.
22. A. N. Oraevskiĭ, *Molecular Generators* (Nauka, Moscow, 1964).
23. A. Allen and J. H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975; Mir, Moscow, 1993).
24. G. P. Agrawal and N. K. Dutta, *Semiconductor Lasers* (Van Nostrand, New York, 1993).
25. A. N. Oraevskiĭ, M. Skalli, and V. L. Velichanskiĭ, Kvantovaya Élektron. (Moscow) **25**, 211 (1998).
26. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, (Fizmatgiz, Moscow, 1963; Pergamon, New York, 1977).
27. R. Hillenbrand, T. Taubner, and F. Keilmann, Nature **418**, 159 (2002).
28. M. Born and E. Wolf, *Principles of Optics* (Pergamon Press, Oxford, 1969; Nauka, Moscow, 1973).

Translated by P. Pozdeev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 479–484.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 529–535.

ATOMS, SPECTRA, RADIATION

On the Change in Electromagnetic Wave Polarization in a Smooth One-Dimensionally Inhomogeneous Medium

K. Yu. Bliokh[a,*] and Yu. P. Stepanovskii[b]

[a]*Institute of Radio Astronomy, National Academy of Sciences of Ukraine, Krasnoznamennaya ul. 4, Kharkov, 61002 Ukraine*

[b]*NSC Kharkov Institute of Physics and Technology, Akademicheskaya ul. 1, Kharkov, 61108 Ukraine*

**e-mail: kostya@bliokh.kharkiv.com*

Received January 27, 2003

Abstract—Propagation of an electromagnetic wave in a smooth one-dimensionally inhomogeneous isotropic medium is considered in the second approximation of geometrical optics. The polarization evolution is studied extensively. It is known that in the first (Rytov) approximation of geometrical optics, there is only the rotation of the plane of polarization (with no change in the polarization shape and sign) for rays with torsion. In the case considered, both the shape of polarization ellipse and the sign of polarization change proportionally to the integral of the squared ray curvature even for plane rays. The effect is of nonlocal geometrical nature and can be described in terms of the generalized geometrical phase incursion between two linear polarizations. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The change in electromagnetic wave polarization in a smoothly inhomogeneous isotropic medium in a geometrical optics approximation was revealed for the first time by Rytov and Vladimirsky [1, 2] (see also [3, 4]). Later, Berry rediscovered this phenomenon in terms of the adiabatic quantum evolution [4–6]. The essence of the effect is as follows. The zero-order geometrical optics approximation does not carry any information on wave polarization: two modes of different polarization are found to be degenerate [3]. In the first approximation, the well-known Rytov law is valid: "The plane of polarization rotates with respect to the natural trihedron of the ray through an angle equal to the integral of the ray torsion along the ray." Interesting geometrical properties of this law were treated in [2] by Vladimirsky and, subsequently, in [4–6] by Berry and his followers. The effect has also been measured experimentally [4, 6]. It is significant that in the first geometrical optics approximation, the polarization ellipse does not change its shape (i.e., the eccentricity), and the polarization does not change its sign, i.e., the direction of the electric field vector rotation [3]. In addition, when the ray represents a plane curve, there is no rotation of the polarization since the torsion is equal to zero. Such a situation takes place, for example, in a one-dimensionally inhomogeneous medium.

In this paper, we study the change in the electromagnetic wave polarization in the second geometrical optics approximation. We will show that, in this case, in contrast to the first-order approximation, there are substantial changes in the shape of the polarization ellipse and the sign of polarization, as well as limited rotations of the polarization plane. The changes occur in such a way that the elliptical polarization, having passed through the linear polarization stage, becomes elliptical with another sign. Similarly, the linear polarization, having passed through an elliptical stage, becomes linear in another plane. Thus, the polarization of the electromagnetic wave may vary substantially as a consequence of the phenomenon in question.

The Rytov rotation of the polarization plane occurs as a result of changes in the phase difference between two modes with opposite circular polarizations. In the case of linear polarization, it is the same as variations in relative amplitudes of two modes with orthogonal linear polarizations. The variations of the polarization shape and sign, derived in this paper, arise from the increase in the phase difference between two modes with orthogonal linear polarizations. The latter phenomenon, like the former one, has a nonlocal geometrical basis. The phase difference accumulating between two modes is in direct proportion to the integral of the squared ray curvature and, hence, is nonzero even for a plane curve. While the Rytov–Vladimirsky–Berry effect of the rotation of the polarization plane is described now in terms of the so-called Berry geometrical phase [4–6], the effect under study (i.e., the effect of the change in the shape of the polarization ellipse) can be described in the context of the generalized geometrical phases introduced in [7, 8]. The generalized geometrical phases are defined as the extension of Berry's geometrical phases to higher orders of the adiabatic approximation, which makes it possible to describe nonlocal geometrical effects of second and higher orders.

1063-7761/03/9703-0479$24.00

Note also that the quantum analog of the classical effect under discussion is given in [9] (see also [4]), where the author discovered that the probability of transition of the radiation from a state with a polarization of a certain sign to the state with a polarization of the opposite sign is proportional to the square of the absolute value of the integral of the ray curvature squared.[1] Since the phase incursion π between two polarization modes changes the polarization sign, it becomes clear that the phase incursion between polarization modes and the probability of the change of polarization sign are directly related to each other.

2. BASIC EQUATIONS

Consider the propagation of a monochromatic electromagnetic wave in an isotropic dielectric nonabsorbing medium characterized by permittivity ε. Let the medium be smoothly inhomogeneous in one coordinate x: $\varepsilon = \varepsilon(x)$. Then the propagation of an electromagnetic wave is described by the stationary Maxwell equations (the fields are assumed to be proportional to $\exp(-i\omega t)$):

$$\begin{aligned} \mathrm{curl}\mathbf{H} + ik_0\varepsilon\mathbf{E} &= 0, \\ \mathrm{curl}\mathbf{E} - ik_0\mathbf{H} &= 0, \end{aligned} \tag{1}$$

where $k_0 \equiv \omega/c$ (ω is the wave frequency and c is the velocity of light in vacuum), and $\mathbf{E}$ and $\mathbf{H}$ are the strengths of the electric and magnetic fields, respectively. Taking into consideration the medium homogeneity in the y and z directions, we carry out the following change of variables:

$$\begin{aligned} \mathbf{E} &\longrightarrow \mathbf{E}(x)\exp(ik_y y + ik_z z), \\ \mathbf{H} &\longrightarrow \mathbf{H}(x)\exp(ik_y y + ik_z z). \end{aligned} \tag{2}$$

We can assume without loss of generality that $k_y = 0$.

System (1) is then reduced to two second-order ordinary differential equations:

$$E_y'' + k_x^2 E_y = 0, \tag{3}$$

$$E_z'' - \frac{k_z^2\varepsilon'}{k_x^2\varepsilon}E_z' + k_x^2 E_z = 0. \tag{4}$$

Here and below, the prime stands for an ordinary derivative with respect to x, and

$$k_x^2(x) = k_0^2\varepsilon(x) - k_z^2. \tag{5}$$

[1] It is interesting to note that, in a two-level system, the probability of the transition to another level is exponentially small in the general case [10]. The fact that in the case of the change in the polarization of the wave, the transition probability is in direct proportion to the fourth power of the curvature (i.e., to the fourth power of the adiabaticity parameter) is a consequence of the polarization degeneracy in a zero-order approximation.

Note also that the derivatives of ε are related to the geometrical optics small parameter,

$$\left|\frac{\varepsilon'}{\varepsilon}\right| \propto \frac{2\pi}{kL} \equiv \delta \ll 1, \tag{6}$$

where L is the characteristic scale of the inhomogeneity. Since we are interested in the second geometrical optics approximation, we retain only the first-order derivatives of ε, their squares, and the second-order derivatives.

After the substitution

$$E_z = \frac{k_x}{k}\tilde{E}_z, \quad k \equiv k_0\sqrt{\varepsilon(x)},$$

Eq. (4) takes a form similar to Eq. (3):

$$\begin{aligned} \tilde{E}_z'' + u\tilde{E}_z &= 0, \\ u \equiv k_x^2 + \frac{k_z^2}{2k_x^2}\frac{\varepsilon''}{\varepsilon} - \frac{k_z^2(k^2 - k_z^2/4)}{k_x^4}&\left(\frac{\varepsilon'}{\varepsilon}\right)^2. \end{aligned} \tag{7}$$

We have derived two equations, (3) and (7), describing the oscillations of y and z components of the electric field. Solving these equations with an accuracy of the second order of the adiabatic approximation in parameter (6), one can find the difference in the evolution of the field phases and amplitudes (the derivation of the higher order adiabatic approximation; see [7, 8, 11]). Let us represent the relative change of the field components as

$$\frac{E_z(x)}{E_y(x)} = \alpha(x)\exp[i\psi(x)], \quad \alpha, \psi \in \mathbf{R}. \tag{8}$$

The quantities α and ψ determine the changes in the relative amplitudes and phases, respectively, of two electric field components.

Equations (3) and (7) coincide in the first geometrical optics approximation; consequently, the amplitudes of E_y and $\tilde{E}_z$ vary identically in the system considered (the second-order terms affect the phase only [8, 11]). Thus, we find that

$$\alpha(x) = \frac{k_x(x)}{k(x)}\frac{k(0)E_z(0)}{k_x(0)E_y(0)}. \tag{9}$$

The change in the phase difference of the field components is determined only by the phase incursion of the two last correction terms in expression (7) for u. Taking into consideration that

$$\sqrt{u} \approx k_x\left[1 + \frac{k_z^2}{4k_x^4}\frac{\varepsilon''}{\varepsilon} - \frac{k_z^2(k^2 - k_z^2/4)}{2k_x^6}\left(\frac{\varepsilon'}{\varepsilon}\right)^2\right],$$

we arrive at

$$\psi = \psi_0 + \int_0^x \left[\frac{k_z^2}{4k_x^3}\frac{\varepsilon''}{\varepsilon} - \frac{k_z^2(k^2 - k_z^2/4)}{2k_x^5}\left(\frac{\varepsilon'}{\varepsilon}\right)^2 \right] dx, \quad (10)$$

where $\psi_0 = \psi(0)$.

3. CHANGES IN THE POLARIZATION SHAPE

Let us proceed from the *y* and *z* components of the electric field to the field projections onto the principal normal and the binormal of the ray. In our case, the binormal points toward the *y* axis ($b_y = -\operatorname{sgn} k_z \operatorname{sgn}\varepsilon'$), and the unit normal vector is

$$\mathbf{\nu} \equiv (\nu_x, \nu_z) = \operatorname{sgn} k_z \operatorname{sgn}\varepsilon' \left(\frac{k_z}{k}, -\frac{k_x}{k}\right).$$

Thus, the ratio of the normal to the binormal component is given by

$$\frac{E_\nu}{E_b} = \frac{k_x E_z - k_z E_x}{k E_y}. \quad (11)$$

Taking into consideration the relation between the *x* and *z* components of the electric field, in the zero-order geometrical optics approximation

$$E_x = -\frac{ik_z E_z'}{k_x^2} \approx -\frac{k_z E_z}{k_x},$$

we obtain

$$\frac{E_\nu}{E_b} = \frac{k}{k_x}\frac{E_z}{E_y} = \tilde{\alpha}\exp(i\tilde{\psi}), \quad (12)$$

where the ratio of the electric field components is represented similarly to (8).

Comparing relation (12) with Eqs. (8) and (9), we obtain

$$\tilde{\alpha} = \frac{k}{k_x}\alpha = \text{const}, \quad \tilde{\psi} = \psi. \quad (13)$$

The quantities $\tilde{\alpha}$ and ψ determine the relative amplitude and phase of two electric field components related to the coordinate system of the natural ray trihedron. It is convenient to use these coordinates in geometrical optics.

In the case of linear polarization ($\psi = 0, \pi$), $\arctan\tilde{\alpha}$ is the angle between the electric field vector and the principal normal to the ray, which varies according to the Rytov law [1–4] in the first approximation. The constancy of $\tilde{\alpha}$ here is also a consequence of the Rytov law, since the ray is a plane curve and the polarization plane does not rotate. With a fixed ψ, the change in the amplitude factor $\tilde{\alpha}$ results in a change of both the ellipse orientation relative to the principal normal to the ray and the eccentricity, whereas the polarization sign (the direction of the field vector rotation in the polarization ellipse) remains unchanged.

The phase difference ψ in oscillations of ν and *b* (or *y* and *z*) components of the electric field affects both the ellipse eccentricity and the polarization sign. The change in ψ with $\tilde{\alpha}$ fixed may reverse the polarization sign. When $\psi = 0, \pi$, the polarization is linear and the angle between the electric field vector and the normal to the ray is equal to $\pm\arctan(\tilde{\alpha}^{-1})$ (Figs. 1a, 1c). When $\psi = \pi/2, 3\pi/2$, the axes of the polarization ellipse are in the *y* and *z* directions, with the ratio of the "normal" semiaxis to the "binormal" one equal to $\tilde{\alpha}$ (Figs. 1b, 1d). When $\psi \in (0, \pi)$ and $\psi \in (\pi, 2\pi)$, the polarization has opposite signs.

Thus, a variation of the relative phase ψ from 0 to π, with fixed $\tilde{\alpha}$, turns the polarization ellipse between two limiting directions (corresponding to linear polarizations) at angles $\pm\arctan(\tilde{\alpha}^{-1})$ to the normal. The polarization sign remains constant in this case. While passing through the linear polarization, the direction of rotation changes and the variation of ψ from π to 2π turns the ellipse in the same way, but in the opposite direction and with the opposite polarization sign (see Fig. 1).

It should be noted that Eqs. (3) and (7) are independent in the approximation considered. If only one electric field component (*y* or *z*) is excited, the other component is not excited in this case and, hence, the wave preserves its linear polarization along the corresponding axis.

4. THE GENERALIZED GEOMETRICAL PHASE ψ

To analyze expression (10), we must invoke the formalism of generalized geometrical phases [7, 8] (the alternative formalism of higher order geometrical phases, see [12]). Indeed, in the integral in Eq. (10), some terms may be *local* and, hence, may cause small deviations (on the order of δ) of ψ. Other terms, on the contrary, may be *nonlocal* with a consequent unbounded growth of ψ. *Local (nonlocal) terms form the terms of the integrand in Eq. (10) that have (do not have) antiderivatives depending only on the boundary values.* The nonlocal terms contain information not only on the boundary points but also on the function values along the whole integration path. These terms can be nonvanishing even in the case when the integration path is a closed curve (loop).

To separate local and nonlocal terms, let us turn to the geometrical formalism and introduce the general-

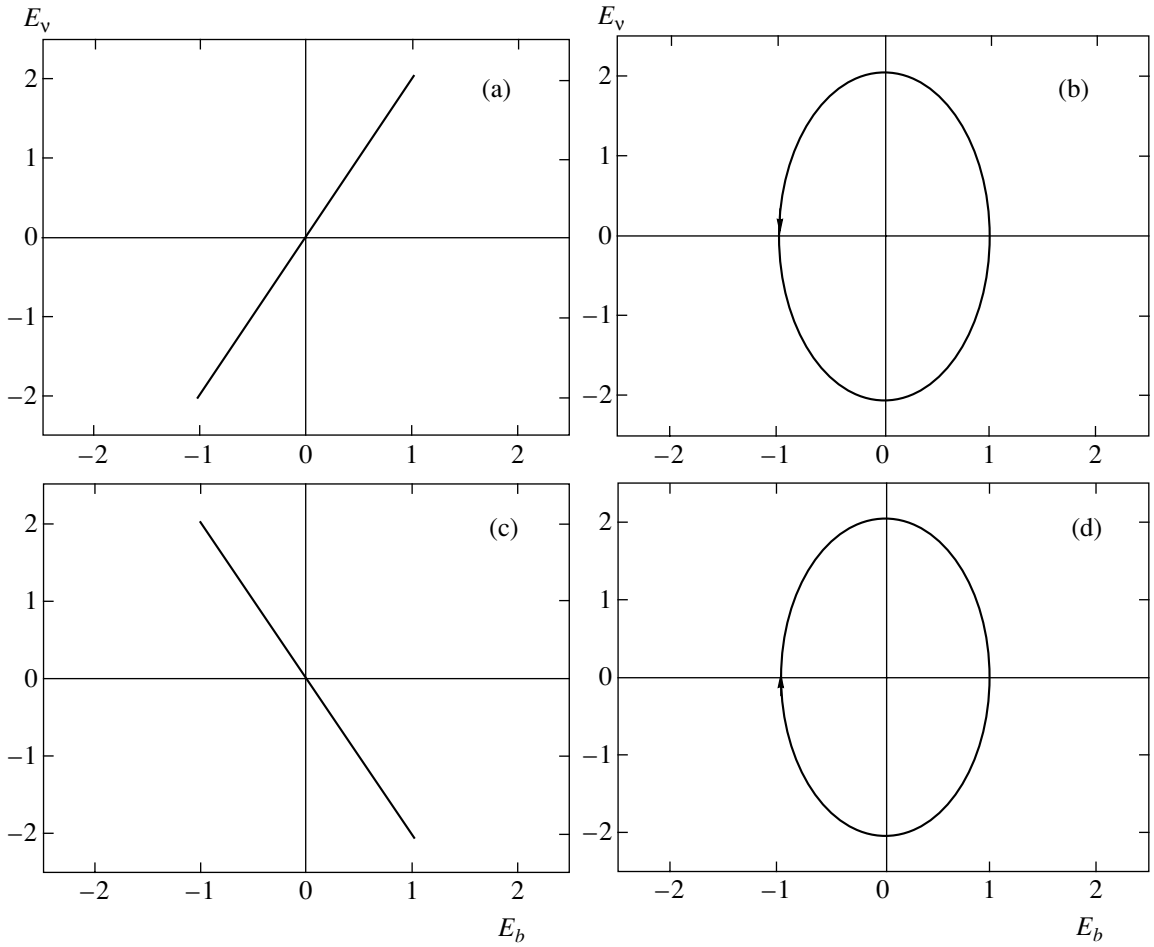

Fig. 1. Evolution of field polarization upon a change in relative phase ψ from 0 to 2π for $\tilde{\alpha} = 2$. Figures 1a–1d correspond to values 0, $\pi/2$, π, and $3\pi/2$.

ized space of the parameter ε: $\vec{M} \equiv (\varepsilon, \varepsilon')$ [7, 8]. Then expression (10) can be written as a contour integral of a certain field over this space:

$$\psi = \psi_0 + \int_l \vec{G} d\vec{M},$$
$$\vec{G} \equiv \left(-\frac{k_z^2(k^2 - k_z^2/4)}{2k_x^5} \frac{\varepsilon'}{\varepsilon^2}, \frac{k_z^2}{4k_x^3} \frac{1}{\varepsilon} \right), \tag{14}$$

where l is the contour along which the representative point of the wave moves in the $\vec{M}$ space. The quantity ψ is nonlocal and can assume significant values only if field $\vec{G}$ is nonpotential, i.e., if

$$\mathrm{curl}\vec{G} \equiv \frac{\partial G_2}{\partial \varepsilon} - \frac{\partial G_1}{\partial \varepsilon'} \neq 0.$$

By substituting Eq. (14) into this expression, we find that field $\vec{G}$ complies with the condition

$$\mathrm{curl}\vec{G} = -\frac{k_z^2}{8\varepsilon^2 k_x^3} \neq 0 \quad \text{for} \quad k_z \neq 0. \tag{15}$$

Hence, we can already state that, for a cyclic evolution of the parameters that correspond to closed contours in the $\vec{M}$ space (e.g., in a periodically inhomogeneous medium), the phase ψ incursion over one cycle is

$$\psi_c = \oint_{l_c} \vec{G} d\vec{M} = \int_S (\mathrm{curl}\vec{G}) \cdot \hat{n} ds$$
$$= -n \int_S \frac{k_z^2}{8\varepsilon^2 k_x^3} d\varepsilon d\varepsilon'. \tag{16}$$

Here, l_c is a closed contour in the $\vec{M}$ space, S is the oriented surface with a unit normal $\hat{n}$ spanning over the contour l_c, and $n = \pm 1$ corresponds to the motion along

the contour in the anticlockwise and clockwise directions, respectively. Thus, for cyclic evolutions, we notice that ψ_c is the generalized geometrical phase determined by the geometry of the closed contour l_c in the $\vec{M}$ space.

In order to isolate the geometrical component from the phase ψ in Eq. (14) in the general case, we separate the nonpotential component from field $\vec{G}$. Then the field in (14) can be written as

$$\vec{G} = \vec{G}_{\text{pot}} + \vec{G}_{\text{curl}}, \quad \text{where } \vec{G}_{\text{pot}} = \text{grad}_{\vec{M}}\varphi,$$
$$\varphi = \frac{k_z^2}{4k_x^3}\frac{\varepsilon'}{\varepsilon}, \quad \vec{G}_{\text{curl}} = \left(\frac{k_z^2}{8k_x^3}\frac{\varepsilon'}{\varepsilon^2}, 0\right), \tag{17}$$

and the phase in (14) takes the form

$$\psi = \psi_0 + \frac{k_z^2}{4k_x^3}\frac{\varepsilon'}{\varepsilon}\Bigg|_0^x + \int_l \vec{G}_{\text{curl}} d\vec{M}. \tag{18}$$

Apparently, the second term arising from $\vec{G}_{\text{pot}}$ is a local term, which is determined only by the boundary values of the parameters. It is always small, being on the order of δ. The last term is the required nonlocal geometrical phase

$$\psi_{\text{geom}} \equiv \int_l \vec{G}_{\text{curl}} d\vec{M} = \int_0^x \frac{k_z^2}{8k_x^3}\left(\frac{\varepsilon'}{\varepsilon}\right)^2 dx. \tag{19}$$

In spite of the fact that the integrand is small (on the order of δ^2), the geometrical phase can be equal to or on the order of unity for $x \geq \delta^{-2}$. This is a direct consequence of its nonlocality or nonpotentiality of field $\vec{G}_{\text{curl}}(\vec{M})$.

Note that initial expression (10) and final expression (19) differ considerably. The phase in Eq. (19) is the essential component of Eq. (10). We could not have isolated it without using the geometrical formalism or the functional approach [7, 8, 13].

Expression (19) can easily be written in ray terms as

$$\psi_{\text{geom}} = \int_0^\sigma \frac{k}{2k_x^2} K^2 d\sigma. \tag{20}$$

Here,

$$K^2 \equiv \left|\frac{d\mathbf{l}}{d\sigma}\right|^2 = \frac{k_z^2}{k^2}\left(\frac{\varepsilon'}{2\varepsilon}\right)^2$$

is the ray curvature squared, $\mathbf{l}$ is the unit tangent vector to the ray ($l_x = k_x/k$, $l_z = k_z/k$), and σ is the ray arc length ($d\sigma = dx/l_x$).

Let us note that the approximation considered is good at distances $x \ll k^{-1}\delta^{-3}$, as long as the remainder term $O(\delta^3 kx)$ in the used asymptotic formulas (see [7, 8]) is small. Taking into consideration the fact that the third-order terms contribute only to the amplitude variations (see [7, 8]), we can conclude that the formulas derived for the difference of the phase incursion of two polarizations are applicable for $x \ll k^{-1}\delta^{-4}$.

The formulas obtained for the polarization evolution follow from the asymptotic solutions to the Maxwell equations and, in the framework of these equations, fail to account for only the backward scattering (reflection). It is known that in the case of one-dimensional inhomogeneity, the reflections are exponentially small in parameter δ (see, for example, [14]). The length over which the reflections begin to manifest themselves can be roughly estimated as $k^{-1}\exp\delta^{-1}$. This length is much larger than the length $k^{-1}\delta^{-m}$ over which the m-order effects manifest themselves, when $(-\delta\ln\delta)^{-1} > m$. For the second-order effect in question, we obtain the restriction $\delta < 0.3$, which is guaranteed to hold by virtue of Eq. (6) (numerical verification of the example in the next section shows that the applicability of the formulas obtained for $\delta = 0.3$ is satisfactory). For $\delta < 0.2$, we can disregard the backward scattering already at lengths $x \sim k^{-1}\delta^{-4}$, where the next phase corrections of the fourth order begin to manifest themselves.

5. EXAMPLE

Consider the inclined propagation of light in a periodic one-dimensionally inhomogeneous medium. Let the $\varepsilon(x)$ dependence be determined by the formula

$$\varepsilon = \varepsilon_0 + \varepsilon_1 \cos(wx), \tag{21}$$

with $\varepsilon_1 < \varepsilon_0$ and $w/k_0 \ll 1$ (in accordance with Eq. (6)).

If the medium is slightly inhomogeneous and $\varepsilon_1 \ll \varepsilon_0$, for nonlocal geometrical phase (19) in the first nonvanishing approximation in $\varepsilon_1/\varepsilon_0$, we have

$$\psi_{\text{geom}} \approx \frac{w^2 k_z^2}{8k_{x0}^3}\left(\frac{\varepsilon_1}{\varepsilon_0}\right)^2 \int_0^x \sin^2(wx) dx, \tag{22}$$

where

$$k_{x0} = \pm\sqrt{k_0^2\varepsilon_0 - k_z^2}.$$

For $x \gg \pi/w$, integral (22) is close to

$$\psi_{\text{geom}} \approx \frac{w^2 k_z^2}{16k_{x0}^3}\left(\frac{\varepsilon_1}{\varepsilon_0}\right)^2 x. \tag{23}$$

As noted above, the substantial change in a shape of polarization occurs when the phase incursion is $\psi \sim 1$. This takes place over distances

$$x \sim \frac{16k_{x0}^3}{w^2 k_z^2}\left(\frac{\varepsilon_0}{\varepsilon_1}\right)^2. \tag{24}$$

In this example, the absolute value of the local term in formula (18) can be estimated as

$$|\psi_{\text{loc}}| \approx \left|\frac{k_z^2}{4k_x^3}\frac{\varepsilon_1 w \sin(wx)}{\varepsilon_0}\right| \leq \frac{wk_z^2}{4k_x^3}\frac{\varepsilon_1}{\varepsilon_0} \ll 1.$$

This confirms the general conclusion about its smallness.

The agreement of the polarization evolution (Fig. 1) and the accumulated incursion of phase ψ with formulas (19) and (23) has been verified via numerical calculation of the initial Maxwell equations in the periodic medium (21). For $k_0 = 1$, $k_z = 0.5$, $\varepsilon_0 = 1$, $\varepsilon_1 = 0.3$, and $w = 0.3$, the discrepancy between Fig. 1 and formulas (19) and (23) is within 10 and 30%, respectively. The above values of the parameters are close to the limits of the applicability of geometrical optics (parameter $\delta = 0.3$, see (6)) and the weak inhomogeneity approximation ($\varepsilon_1/\varepsilon_0 = 0.3$). According to estimate (24), phase ψ by π varies for these values of the parameters over the distance $x_\pi = 800\lambda \approx 250L$, where $\lambda = 2\pi k_0^{-1}$ is the wavelength and $L = 2\pi w^{-1}$ is the spatial period of the inhomogeneity.

6. CONCLUSIONS

The evolution of the polarization of an electromagnetic wave propagating in a smooth one-dimensionally inhomogeneous isotropic medium has been studied. The problem has been considered in the second geometrical optics approximation.

It is known that, in the first geometrical optics approximation, the polarization evolution is described by the Rytov law [1–4]. According to this law, typically the relative *amplitude* of the normal and binormal electric field components changes in the medium where a ray of the wave exhibits torsion. This results in the rotation of the polarization ellipse of the wave; however, the ellipse eccentricity and the polarization sign remain unchanged. In the case of a one-dimensionally inhomogeneous medium, the ray represents a plane curve and the change of polarization in the Rytov approximation does not occur.

We have shown that in the second geometrical optics approximation, in contrast to the first-order approximation, a change in the relative *phase* of the normal and binormal electric field component is observed. This leads to changes in the eccentricity of the polarization ellipse, to its limited rotation, and to a change in the polarization sign (Fig. 1). It is shown above that the change in the relative phase, similar to the Rytov rotation of the polarization plane, is a geometrical nonlocal effect, the change being proportional to the integral of the ray curvature squared. As a consequence, the relative phase grows even under cyclic evolution, when both the medium parameters and the ray direction return to their initial values. Specifically, in a periodically inhomogeneous medium, the relative phase grows continuously with consequent continuous changes in polarization (see Section 5).

Since the revealed phenomenon causes the polarization sign reversal, its quantum analog is nonadiabatic transitions between two states of a photon, which correspond to the "+" and "–" polarizations. In [9], the formula for the probability of this kind of transition in a smoothly curved waveguide is presented. Assuming that $k_x = k$ in formula (20) (since the wave propagates in the waveguide transversely to the inhomogeneity), we find that the square of the modulus on the right-hand side of Eq. (20) coincide, to a factor of 1/4, with the probability of the nonadiabatic transition obtained in [9] (if it is assumed there that the torsion is equal to zero).

It should be pointed out that the discovered effect is rather weak, as it corresponds to the second approximation of geometrical optics. As evident from the above example, even for a relatively strong inhomogeneity and at the boundary of the region of applicability of the geometrical optics approximation, the changes in polarization are observed over distances on the order of a thousand wavelengths and a hundred inhomogeneity scales. Nevertheless, this phenomenon is the only one that causes the change in polarization (in the approximation considered) in a one-dimensionally inhomogeneous medium, which is why it can be observed.

REFERENCES

1. S. M. Rytov, Dokl. Akad. Nauk SSSR **18**, 263 (1938).
2. V. V. Vladimirskiĭ, Dokl. Akad. Nauk SSSR **31**, 222 (1941).
3. Yu. A. Kravtsov and Yu. I. Orlov, *Geometrical Optics of Inhomogeneous Media* (Nauka, Moscow, 1980).
4. S. I. Vinitskiĭ, V. L. Debrov, V. M. Dubovik, *et al.*, Usp. Fiz. Nauk **160** (6), 1 (1990) [Sov. Phys. Usp. **33**, 403 (1990)].
5. M. V. Berry, Proc. R. Soc. London, Ser. A **392**, 45 (1984).
6. A. Shapere and F. Wilczek, *Geometrical Phases in Physics* (World Sci., Singapore, 1989).
7. K. Yu. Bliokh, J. Math. Phys. **43**, 5624 (2002).
8. K. Yu. Bliokh, J. Phys. A: Math. Gen. **36**, 1705 (2003).
9. M. V. Berry, Nature **326**, 277 (1987).
10. J.-T. Hwang and P. Pechukas, J. Chem. Phys. **67**, 4640 (1977).
11. C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).
12. M. V. Berry, Proc. R. Soc. London, Ser. A **414**, 31 (1987).
13. K. Yu. Bliokh, Izv. Vyssh. Uchebn. Zaved., Prikl. Nelin. Din. **9**, 45 (2001).
14. G. M. Zaslavskiĭ, V. P. Meĭtlis, and N. N. Filonenko, *Wave Interaction in Inhomogeneous Media* (Nauka, Novosibirsk, 1982).

Translated by K. Bliokh

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 485–492.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 536–544.

ATOMS, SPECTRA, RADIATION

One-Parametric Dependences of the Spectra, Scattering Lengths, and Recombination Coefficients for a System of Three Bosons

F. M. Pen'kov

Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia
e-mail: penkov@thsun1.jinr.ru

Received February 4, 2003

Abstract—A model of a two-pole pair t-matrix for the description of three-boson systems with $r_0/a_0 \ll 1$ is studied. The position of the second pole on the nonphysical sheet is treated as a parameter of the Faddeev integral equation. Close agreement with the calculations performed for realistic potentials was obtained, and interaction parameters for three bosons that were not calculated earlier were predicted. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The discovery of a weakly bound state of two helium atoms [1] and problems concerning stability of the Bose condensate of alkali metal atoms (e.g., see [2]) stimulated the appearance of many calculations for the properties of three-particle systems determined by pair interaction with a large scattering length a_0 compared to the range of pair forces r_0,

$$\frac{r_0}{a_0} \ll 1. \tag{1}$$

Studying such three-particle systems, which is of importance for practice, entails considerable difficulties in numerical calculations, which require the inclusion of very large distances to determine the asymptotic behavior of the wave function of a three-particle system. Calculations of the helium atom spectrum are numerous (e.g., see [3] and the references therein). Little work, however, has been done to determine the lengths of helium atom scattering by the helium dimer [3, 4]. What is more, there is a cutoff dependence of scattering-length values obtained in numerical calculations [4]. Still greater difficulties arise with recombination coefficients, which determine Bose condensate stability.

The three-particle systems that we are considering are of interest not only because of their importance for practice but also because their two-particle interaction parameters are close to the conditions of Efimov effect manifestations [5]. In particular, the $\ln(|a_0|/r_0)/\pi$ value, which determines the number of Efimov levels in a three-particle system for helium–helium interactions, approximately equals 0.9. This leaves open the possibility that the second bound state of the trimer will exhibit anomalous behavior when pair interactions become shallower. Indeed, numerical simulations of such a situation lead to an increase (in magnitude) in the excited-state energy of the system [6]. For this reason, the system of three helium atoms is the first system of three particles with equal masses whose properties are determined by the effective self-similar interaction proportional to $1/R^2$ with the coupling constant $-\gamma^2$ independent of pair interaction details [5]. Pair interaction parameters only limit the range of such an effective potential, $r_0 \ll R \ll |a_0|$. The obvious properties of solutions for this potential are indicative of the determining role played by one dimensionless parameter, namely, $|a_0|/r_0$. When this parameter tends to infinity, the number of bound states of the three-particle system does the same. In particular, if $|a_0| \longrightarrow \infty$, the spectrum crowds to zero, which determines the Efimov effect proper. If $r_0 \longrightarrow 0$, the spectrum is unbounded from below [7], which is equivalent to falling on the center [8]. According to [9], the integral equations for three-particle systems with a nonzero range of pair forces [10] do not have a unique solution and are determined by one free parameter. This situation determines a possible procedure for describing such systems via fixing the free parameter according to a known spectral point [9]. This idea has been repeatedly used to describe real physical systems. In particular, in recent years, the effective field theory (EFT) was applied to construct phenomenological three-particle equations that included a free parameter from the outset, based on the additional condition of the convergence of the EFT series [11]. The use of three-particle equations with zero-range pair forces to describe real physical systems implies the smallness of finite-range corrections. However currently, no rigorous conclusions concerning the magnitude of corrections for finite $|a_0|/r_0$ values have been made. For this reason, three-particle systems are considered in this

work using the effective pair interaction range as a parameter. The main stages of solving the problem remain unchanged. The effective range is used as a parameter in the Faddeev integral equations and is tuned to bring one of the two levels in the system of three helium atoms (trimer) in coincidence with the existing data. Next, the second level and the length of helium atom scattering by a bound pair of two other helium atoms are calculated. It is shown below that the calculations performed according to this scheme agree with the results of calculations using the Faddeev differential equations to within several percent. Finally, recombination coefficients are calculated.

In order that the Faddeev equations explicitly include the effective range of forces, we use the two-pole pair matrix $t(p, p'; p_0^2/m)$. One of its poles corresponds to the bound state of the pair, and the position of the other determines the effective range. In particular, if the separable pair interaction is used in the form of the Yamaguchi potential

$$v(p, p') = -\frac{8\pi}{m}\frac{\beta(\beta+\kappa)^2}{(\beta^2+p^2)(\beta^2+p'^2)}, \qquad (2)$$

the two-particle matrix $t(p, p'; p_0^2/m)$ has poles at $p_0 = i\kappa$ and $p_0 = -i(2\beta + \kappa)$. Here, p is the relative motion momentum; m is the mass of the atom; $\kappa = \sqrt{-m\varepsilon}$ is the wave number of the state of the dimer with energy ε; and β determines the expansion parameters for the effective range,

$$a_0 = \frac{2(\beta+\kappa)^2}{\beta\kappa(\kappa+2\beta)}, \quad r_0 = \frac{\kappa^2+2\beta\kappa+3\beta^2}{\beta(\beta+\kappa)^2}.$$

Note that, if $\beta \gg \kappa$, the scattering length is $a_0 \approx \kappa^{-1}$ and the effective range is $r_0 \approx 3/\beta$. The zero-range limit is attained by turning β to infinity. Because β is explicitly related to the effective range, precisely this value will be used as the problem parameter.

2. SCATTERING LENGTH AND BOUND STATES OF THE HELIUM TRIMER

The energies of the bound states of three helium atoms and the amplitudes of elastic helium atom scattering by the helium dimer (He_2) below the breakup energy were calculated by solving the Faddeev equations with separable pair potentials (2). The procedure for constructing and symmetrizing the Faddeev equations in the specified problem has repeatedly been described (e.g., see monograph [12]). In the next section, we give a scheme for constructing the equation for the inelastic transition amplitude. For comparison, the main stages of deriving the equation for the elastic scattering amplitude will also be presented. Here, we note only that the scheme considered in [13] is used. According to this scheme, the integral equations for the amplitude of boson scattering by a bound state of two other bosons have the simple quasi-two-particle form

$$f(k_i, k_0) = V_0^{\text{eff}}(k_i, k_0) + \frac{2}{\pi}\int_0^\infty \frac{V_0^{\text{eff}}(k_i, k)k^2 dk}{k^2 - (Z-\varepsilon)4m/3} f(k, k_0), \qquad (3)$$

where

$$Z = \frac{3k_0^2}{4m} + \varepsilon + i0$$

is the total energy of the system, $\mathbf{k}$ is the momentum of the relative motion of the particle and the bound pair, and V_0^{eff} is the S-wave projection of the effective potential. The three-particle effective potential for the Yamaguchi pair interaction has the form

$$V^{\text{eff}}(k_i, k_j) = \frac{16}{3}\frac{(\beta+\kappa)^2\beta S(k_i)}{(\beta^2+(\mathbf{k}_j+\mathbf{k}_i/2)^2)(-mZ+k_i^2+k_j^2+\mathbf{k}_i\cdot\mathbf{k}_j)(\beta^2+(\mathbf{k}_i+\mathbf{k}_j/2)^2)}, \qquad (4)$$

where the notation

$$S(k) = (\beta+a_k)^2\frac{\kappa+a_k}{2\beta+\kappa+a_k},$$

$$a_k = \sqrt{-m\left(Z-\frac{3k^2}{4m}\right)}$$

is used. Note that potential (4) remains finite and (3) transforms into the Skornyakov–Ter-Martirosyan equation [10] as $\beta \longrightarrow \infty$.

The scattering length was calculated by two methods to estimate the accuracy of the effective range expansion, namely, by solving the limiting equation for $k_0 = 0$ and by finding the limit of the solution (amplitude) $a^{(3)} = f(k_0, k_0)$ as $k_0 \longrightarrow 0$. A homogeneous equation with the same kernel is solved in the problem for the bound state. Equation (3) was solved numerically following the scheme described in [14]. The parameters

Calculated energies of helium trimer bound states, lengths of helium atom scattering by the helium dimer, and recombination coefficients for helium atoms; see text for explanations

Method	β, Å^{-1}	E_1, K	E_2, mK	$a^{(3)}$, Å	C
HFD-B, ε = –1.685411 mK					
1	0.3305		–2.7899	116.60	0.977
2	0.3200	–0.12597		121.58	0.704
[3]		–0.1325	–2.74	135 ± 5	
[4]				121.9	
LM2M2, ε = –1.303483 mK					
1	0.3305		–2.3163	109.48	2.513
2	0.3220	–0.12070		114.25	2.145
[3]		–0.1259	–2.28	131 ± 5	
[4]				115.4	
SAPT, ε = –1.898390 mK					
1	0.33111		–3.0501	118.23	0.521
2	0.31730	–0.12763		124.35	0.276
[15, 4]		–0.13637	–2.986	123.7	
SAPT1, ε = –1.732405 mK					
1	0.33130		–2.8507	116.74	0.868
2	0.31785	–0.12539		123.00	0.554
[15, 4]		–0.13382	–2.790	122.4	
SAPT2, ε = –1.81500 mK					
1	0.33130		–2.9507	117.51	0.686
2	0.31735	–0.12638		123.86	0.396
[15, 4]		–0.13516	–2.887	123.1	

given below were calculated accurate to no less than six decimal places. Such an accuracy was attained using the algorithm with a floating integrand discretization mesh. The method for selecting a mesh for numerically solving integral equations of the type $f = V + Vg_0 f$ with a completely continuous kernel is based on a preliminary estimative integration of the Vg_0V value with the required accuracy. This can be done because the f value is a "smoother" function than V for compact equations. The details of selecting the class of functions for which a unique solution is attained can be found in monograph [12].

As has been mentioned, the β value was treated as the problem parameter. Its value was adjusted to bring one of the helium trimer levels in coincidence with the level earlier calculated by other authors [3, 15] using the realistic Aziz potentials, namely, HFD-B [16], LM2M2 [17], SAPT [18], SAPT1 [18], and SAPT2 [18]. The helium trimer has two bound states, and we performed calculations to fit both the first and second levels. The results are summarized in the table. The through table rows contain the pair potential types and the corresponding dimer binding energies. The energies for the HFD-B and LM2M2 potentials were calculated in this work, whereas those for the SAPT potentials were taken from [15]. The references in the leftmost column (under the heading "method") indicate the sources of information for fitting the ground E_1 or excited E_2 trimer state, and the index 1 or 2 is the number of the level (E_1 or E_2) that was fitted by the β parameter given in the second column. Note that [3] contains both helium trimer binding energies and the lengths of helium atom scattering by the dimer, whereas the calculated trimer binding energies and scattering lengths for the SAPT potentials are given in different works, [15] and [4], respectively. To simplify the table, the fitted levels are omitted. The column with C values refers to the results described in the next section, which will be considered later on. The tabulated calculation results are evidence of satisfactory agreement for such a sim-

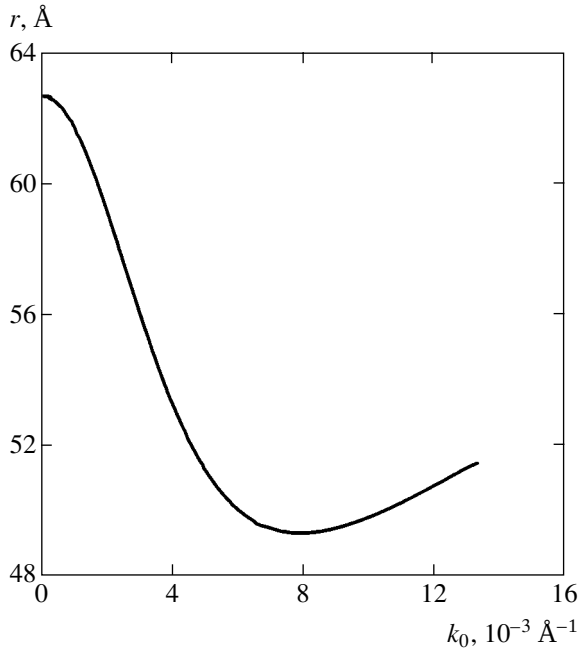

Fig. 1. Dependence $r(k_0)$. See text for explanations.

ple model. The maximum error in the position of the ground state obtained by fitting the excited level is 6.5%, and the maximum error in the excited state energy obtained by fitting the ground-state level is 2.2%. As to agreement between the scattering lengths obtained by fitting the excited state level and calculated in [4], it is close (relative error less than 1%) to the extent that it can serve as a criterion for the accuracy of calculating these values. Recall that the difference in the scattering lengths reported in [3] and [4] arises because of different cutoff conditions used to numerically solve the Faddeev differential equations [4].

High-accuracy calculations of scattering amplitudes allow us to draw conclusions about the possibility and applicability region of the quasi-two-particle effective range expansion

$$k_0 \cot\delta(k_0) = -\frac{1}{a^{(3)}} + \frac{r_{\text{eff}} k_0^2}{2} + \ldots .$$

Here, δ is the phase of helium atom elastic scattering by the dimer in the S-wave and r_{eff} is the effective scattering range. This expansion in even powers of momentum is often used in analyzing low-energy scattering and is valid for fairly rapidly decreasing potentials (e.g., see [19]).

The first and second derivatives were calculated to determine the momentum dependence of $k_0 \cot\delta(k_0)$. The HFD-B potential was selected by way of example. This function and its first derivatives are not plotted here because they contain little information. Note only that the coefficient of the linear term in the expansion of $k_0 \cot\delta(k_0)$ is zero to within the error of computations. Figure 1 shows the second derivative,

$$r(k_0) = \frac{d^2(k_0 \cot\delta(k_0))}{dk_0^2},$$

which coincides with the effective range as $k_0 \longrightarrow 0$.

It is clearly seen from Fig. 1 that the parabolic momentum dependence near zero (admixture of k^4) changes in the region of momenta on the order of 10^{-3} Å^{-1} and then exhibits a well-defined linear dependence up to momenta on the order of 4×10^{-3} Å^{-1}. It follows that the applicability region of the expansion in integer powers of energy does not exceed 10^{-5} K, which prevents us from describing the amplitude of helium atom scattering by the dimer with the simple pole equation

$$f(k) = \frac{1}{-1/a + r_{\text{eff}} k^2/2 - ik}$$

(e.g., see [20]). Moreover, substituting $r_{\text{eff}} = 62.6$ Å and $a^{(3)} = 121.58$ Å into this equation shows that the amplitude has poles at energies containing both real and imaginary components rather than at the trimer binding energy. Note that substituting $a_0 = 88.601$ Å and $r_0 = 7.278$ Å, which are the parameters of the expansion of the effective range for pair collisions of helium atoms (the HFD-B potential), into the equation for the amplitude gives positive poles with a relative error of 4×10^{-5}.

3. RECOMBINATION COEFFICIENTS

The recombination coefficient α is related to the number of $3 \longrightarrow 2$ transitions ($1 + 1 + 1 \longrightarrow 2 + 1$) per unit time from unit volume as αn^3, where n is the concentration of atoms. By definition, the total energy of the system of three particles is $Z = 0$.

The recombination of atoms at low temperatures is of special interest because this phenomenon breaks up Bose condensation. At the same time, calculations of this recombination coefficient involve large uncertainties. For instance, it is claimed in [2] that the dimensionless coefficient C introduced from dimensional considerations,

$$\alpha = C\frac{a_0^4 \hbar}{m},$$

is universal for all systems and equals 3.9 if condition (1) is satisfied. The authors of [21] note that such a universality cannot be ensured, and the give a $C = 0$–65 spread

of C values. The EFT model is used in [22] to obtain the phenomenological equation

$$C = C_{\max}\cos^2[s_0\ln(a_0\Lambda) + \delta],$$
$$C_{\max} = 67.9 \pm 0.7, \quad \delta = 1.74 \pm 0.02, \tag{5}$$

whose parameters are determined from the integral equation of the EFT model. In this equation, $s_0 \approx 1.0063$ and Λ is the parameter of the theory and can be determined by the procedures described above, that is, by comparing it with binding energy or scattering length calculations for three-particle systems.

Our task was to find a solution from the exact three-particle Faddeev equations and determine its asymptotic behavior as $\beta \longrightarrow \infty$ ($r_0 \longrightarrow 0$).

The recombination coefficient can be written via the $3 \longrightarrow 2$ transition amplitude as

$$\alpha = \pi\frac{k_f}{m}|f_{3\to 2}|^2, \tag{6}$$

where $k_f = \sqrt{4/3}\,\kappa$. Note that this equation describes the two-particle flow in the exit channel multiplied by a factor of 1/6. This flow arises when we take into account the number of triple combinations.

Equation (3) for the elastic scattering amplitude was given above without derivation. For this reason, we will describe the scheme for deriving the Faddeev integral equation that determines the $3 \longrightarrow 2$ inelastic transition amplitude for an arbitrary separable pair interaction $v = |v\rangle\langle v|$. Let the *in* state of the continuous spectrum be labeled by index "0," and the bound-pair state, by a Greek letter. The Greek momentum index will determine one of the three sets of Jacobi coordinates. Below, we use Liepmann–Schwinger-type integral equations. The methods for deriving these equations can be found in many monographs (e.g., see [19]). To emphasize the generality of constructing equations for elastic and inelastic processes, we will write common equations as far as it is possible. For instance, common equations will be given for transitions from the continuous spectrum and from the subsystem numbered one. For this reason, we define the channel T matrix by the equation (see [19]) with a double subscript in parentheses,

$$T_{\alpha,(0,1)} = \langle\varphi_\alpha|\bar{v}_\alpha|\Psi^+_{(0,1)}\rangle,$$

where $\bar{v}_\alpha$ is the standard denotation of the sum of the pair potentials except for the interaction that binds subsystem α, $|\Psi^+_{(0,1)}\rangle$ is the three-particle wave function in the continuous spectrum for index 0 or the three-particle wave function with a bound pair subsystem in the *in* state, and $\langle\varphi_\alpha|$ is the wave function of the bound pair state in the end channel. For simplicity, the quantum numbers of continuous spectrum states are omitted. In particular, the state of the continuous spectrum of two interacting particles in channel α is denoted by $|\varphi^+_\alpha\rangle$.

Since the particles are identical, the inelastic transition amplitude is given by the sum of three amplitudes,

$$f_{3\to 2} = -\frac{m^*}{2\pi}T^{\text{inel}}, \quad T^{\text{inel}} = \sum_\alpha T_{\alpha,0}.$$

Here, $m^* = (2/3)m$ is the reduced mass of the three-particle system. Note that the elastic scattering amplitude is written similarly,

$$f_{2\to 2} = -\frac{m^*}{2\pi}T^{\text{el}}, \quad T^{\text{el}} = \sum_\alpha T_{\alpha,1}.$$

The integral equations for the continuous spectrum wave function,

$$|\Psi^+_0\rangle = |\varphi^+_\alpha\rangle + G_\alpha\bar{v}_\alpha|\Psi^+_0\rangle, \quad \alpha = 1, 2, 3,$$

can be used to obtain the Faddeev equations for $U_{\alpha,0} = v_\alpha|\Psi^+_0\rangle$,

$$U_{\alpha,0} = t_\alpha + t_\alpha G_0(U_{\beta,0} + U_{\gamma,0}),$$
$$\alpha \neq \beta \neq \gamma. \tag{7}$$

Here, $G_0 = (Z - H_0)^{-1}$ is the Green function of the three-particle system with free Hamiltonian H_0, $G_\alpha = (Z - H_0 - v_\alpha)^{-1}$ is the Green function with interaction in the pair subsystem α, t_α is the pair t matrix in the three-particle space that satisfies the integral equation

$$t_\alpha = v_\alpha + v_\alpha G_0(Z)t_\alpha,$$

and Z is the total three-particle system energy whose positive imaginary part tends to zero.

Note that the wave function of transitions from the *in* state with bound subsystem 1 is described by other equations, namely,

$$|\Psi^+_1\rangle = |\varphi_\alpha\rangle\delta_{1,\alpha} + G_\alpha\bar{v}_\alpha|\Psi^+_1\rangle, \quad \alpha = 1, 2, 3.$$

Accordingly, the $U_{\alpha,1} = v_\alpha|\Psi^+_1\rangle$ values are given by the equations

$$U_{\alpha,1} = \tau_\alpha\delta_{1,\alpha} + t_\alpha G_0(U_{\beta,1} + U_{\gamma,1}),$$
$$\alpha \neq \beta \neq \gamma, \tag{8}$$

which have the same kernel as (7) but different free terms. Here, $\tau_\alpha = v_\alpha|\varphi_\alpha\rangle$ and $\delta_{i,j}$ is the standard Kronecker symbol. We will use two different two-particle Green functions with differently written energy parameters, namely, $g_0(\varepsilon_\alpha)$ for the free motion of a pair of par-

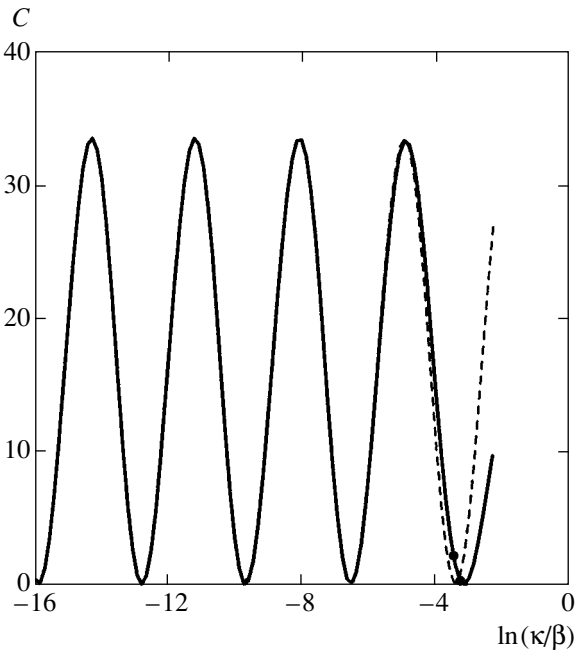

Fig. 2. Dependence of C on pair interaction parameters. The solid line corresponds to calculations by (11), and the dashed line, to calculations by EFT model (5). The two solid circles show the maximum and minimum C values for realistic potentials. See text for explanations.

ticles in subsystem α and $g_0(Z-\varepsilon_\alpha)$ for the free motion of the third particle with respect to subsystem α. Taking into account the representation of the pair t matrix for separable potentials [13]

$$t_\alpha G_0 = v_\alpha|\varphi_\alpha\rangle g_0(Z-\varepsilon_\alpha)\langle\tilde{\varphi}_\alpha|$$

via the off-shell wave function of the subsystem

$$\langle\tilde{\varphi}_\alpha| = \frac{\langle\varphi_\alpha|v_\alpha G_0(Z)}{\langle\varphi_\alpha|v_\alpha g_0(\varepsilon_\alpha)G_0(Z)|v_\alpha\rangle},$$

whose projection $\langle\tilde{\varphi}_\alpha|\mathbf{k}_\alpha, \mathbf{p}_\alpha\rangle$ coincides with $|\mathbf{k}_\alpha\rangle\langle\varphi_\alpha|\mathbf{p}_\alpha\rangle$ on the energy shell, when

$$\frac{3k_\alpha^2}{4m} + \frac{p_\alpha^2}{m} = \frac{3k_f^2}{4m} + \varepsilon = Z,$$

we can introduce the off-shell T matrix

$$\tilde{T}_{\alpha,(0,1)} = \langle\tilde{\varphi}_\alpha|(U_{\beta,(0,1)} + U_{\gamma,(0,1)})$$

with the obvious property $\tilde{T}_{\alpha,(0,1)} \longrightarrow T_{\alpha,(0,1)}$ as $k \longrightarrow k_f$. Faddeev equation (7) can then be used to obtain the equation for the off-shell T matrix,

$$\tilde{T}^{\text{inel}} = 6R + 2Vg_0(Z-\varepsilon_\alpha)\tilde{T}^{\text{inel}}, \qquad (9)$$

where

$$R = \langle\tilde{\varphi}_\alpha|t_\beta|p_\beta\rangle, \quad V = \langle\tilde{\varphi}_\alpha|v_\beta|\varphi_\beta\rangle,$$
$$\alpha \neq \beta, \quad p_\beta = 0.$$

Equations (8) give equations for elastic transitions in the form

$$\tilde{T}^{\text{el}} = 2V + 2Vg_0(Z-\varepsilon_\alpha)\tilde{T}^{\text{el}}. \qquad (10)$$

Precisely this equation written for the Yamaguchi potential (2) gives (3). In the preceding section, (3) was used to determine the amplitude of elastic particle scattering by a bound pair and the spectrum of the three-particle system.

Prior to further transformations, note that the inhomogeneous term in (9) is written via V,

$$R = VD, \quad D = \frac{v_\alpha(p_\alpha)g_0(Z-\varepsilon_\alpha)}{\langle\varphi_\alpha|v_\alpha g_0(\varepsilon_\alpha)G_0(Z)|v_\alpha\rangle}.$$

This allows $3D$ to be factored out from the equation by the mere substitution $f_{3\to 2} = 3DF$. The complex equation for F can conveniently be reduced to two real equations using the simple relation $F = F_0 + ik_fF_1F$.

After transforming the Faddeev equations this way to the form convenient for solving them, we can express C via two-particle interaction parameters and the solutions to the Faddeev equations for F_0 and F_1,

$$C = \frac{4\pi^2 27}{\sqrt{3}}\frac{\beta(\beta+\kappa)}{(2\beta+\kappa)^2}\frac{k_f^2F_0^2(k_f)}{1+k_f^2F_1^2(k_f)},$$

where the F_0 and F_1 amplitudes correspond to the solutions of the integral equations with the same kernel but different free terms,

$$F_n(k_i) = V^{\text{eff}}(k_i, k_n) + \frac{2}{\pi}\int_0^\infty \frac{V^{\text{eff}}(k_i, k)k^2dk}{k^2-k_f^2}F_n(k), \qquad (11)$$
$$n = 0, 1, \quad k_0 = 0, \quad k_1 = k_f.$$

The integration in the vicinity of the Cauchy singularity is understood in the sense of the principal value. The effective potential $V^{\text{eff}} = -(m^*/\pi)V$ is as previously determined by (4).

Note that one of the equations in (11) coincides with the equation for elastic scattering. More exactly, $F_1(k_f) = \tan(\delta)/k_f$, where δ is the phase of atom elastic scattering by the dimer at the three-particle threshold energy.

Equations (11) were solved numerically. The C values corresponding to different pair potentials are listed in the table. It is noteworthy that, the difference in scattering lengths being small, the recombination coeffi-

cients differ almost by an order of magnitude. To study the reason for such a difference, we plotted the dependence of C on the κ/β ratio (Fig. 2). This dependence shows that C logarithmically oscillates with a large amplitude.

Our numerical calculations lead us to conclude that the remark made in [21] concerning the nonuniversality of the C value was justified. It can also be stated that the empirical formula from [22] is only valid in the asymptotic region, far outside the region of real r_0/a_0 values. To illustrate this statement, two curves are plotted in Fig. 2. The results of our calculations are given by the solid line, and the results of calculations by (5), by the dashed line. The solid circles are the C values for two helium–helium potentials (see table), which give the largest and smallest C values. These values lie on the curve that depends only on the r_0/a_0 ratio (the κ/β ratio in this work) but does not coincide with the curve given in [22]. Note that our curve differs in amplitude from the curve constructed according to (5) exactly by a factor of 2. Curve (5) was therefore normalized by a factor of 0.5. The difference of C values by a factor of 2 was already mentioned in [21], where a comparison with the data from [2] was made, and that the results of [22] were in agreement with those obtained in [21]. Unfortunately, the reason for this difference is unclear. The limiting form of (11) ($\beta \longrightarrow \infty$) shows that the inhomogeneous terms of this equation and the EFT model equation obtained in [22] by a factor of $\sqrt{2}$. The final results differ by a factor of 2 precisely for this reason. The brief description of the derivation of the integral equation for the ($3 \longrightarrow 2$) amplitude given in [22] is insufficient for drawing conclusions on the origin of this difference.

4. ONE-PARAMETER DEPENDENCES

Equations (3) and (11) can be made dimensionless by the simple change of variables $k = x\kappa$. We then obtain equations that depend on a single dimensionless parameter (κ/β) and describe not only the C value but also the dimensionless scattering amplitude $f\kappa$ and the dimensionless trimer binding energies E_1/ε and E_2/ε. We can therefore exclude the κ/β adjustable parameter and construct the curves that interrelate the bond energies, scattering length, and C.

The dependences of the C value, dimensionless length of boson scattering by a bound pair, and the dimensionless trimer ground state energy on the dimensionless trimer excited-state energy are shown in Fig. 3. The values marked by crosses were taken from [3, 4, 15]; they refer to bond-energy and scattering-length calculations for the system of three helium atoms. The degree to which the results of our calculations performed using the simple pair interaction model agree with those obtained for realistic atom–atom interaction potentials inspires hopes that, generally, these dependences apply to arbitrary three-boson systems that satisfy condition (1).

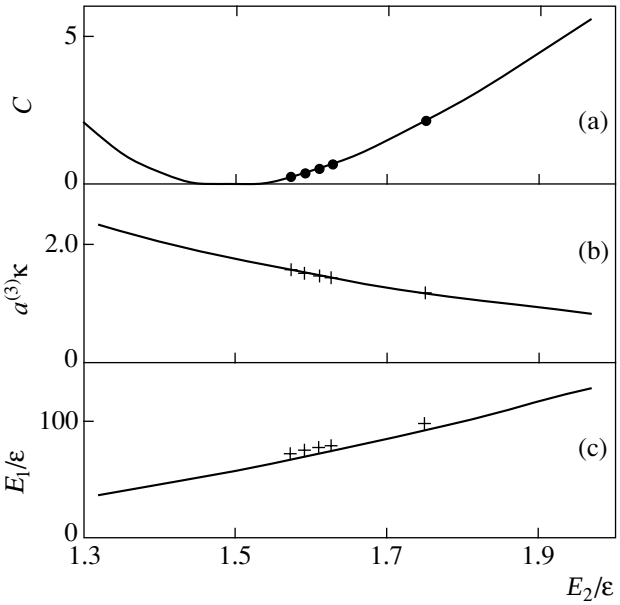

Fig. 3. Dependences of the dimensionless characteristics of the three-boson system on the dimensionless trimer excited state energy. (a) The C value that determines the recombination coefficient α; solid circles are the calculation results obtained in this work for realistic helium–helium potentials. (b) The dimensionless length of boson scattering by a bound pair; crosses are the calculation results obtained in [4] for realistic helium–helium interaction potentials. (c) Dimensionless trimer ground state energy; crosses correspond to the calculation results obtained in [3, 15] for realistic helium–helium interaction potentials.

5. CONCLUSIONS

The calculations performed with the simple two-pole t matrix for the helium–helium pair interaction give spectra and scattering lengths for the system of three helium atoms that closely agree with those obtained using realistic potentials. We can therefore hope that the recombination coefficients, which were calculated for the first time, will also be close to the results of future calculations with realistic pair interactions.

ACKNOWLEDGMENTS

This work was financially supported by the Russian Foundation for Basic Research (project no. 01-02-17575).

REFERENCES

1. F. Luo, G. C. McBane, G. Kim, *et al.*, J. Chem. Phys. **98**, 3564 (1993).
2. P. O. Fedichev, M. W. Reynolds, and G. V. Shlyapnikov, Phys. Rev. Lett. **77**, 2921 (1996).

3. A. K. Motovilov, W. Sandhas, S. A. Sofianos, and E. A. Kolganova, Eur. Phys. J. D **13**, 33 (2001).
4. V. Roudnev, Chem. Phys. Lett. **367**, 95 (2003); V. A. Roudnev, Preprint No. E4-2002-85, JINR (Joint Inst. for Nuclear Research, Dubna, 2002).
5. V. Efimov, Yad. Fiz. **12**, 1080 (1970) [Sov. J. Nucl. Phys. **12**, 589 (1970)].
6. E. A. Kolganova, A. K. Motovilov, and S. A. Sofianos, J. Phys. B **31**, 1279 (1998).
7. R. A. Minlos and L. D. Faddeev, Zh. Éksp. Teor. Fiz. **41**, 1850 (1961) [Sov. Phys. JETP **14**, 1315 (1961)].
8. L. H. Thomas, Phys. Rev. **47**, 903 (1935).
9. G. S. Danilov, Zh. Éksp. Teor. Fiz. **40**, 498 (1961) [Sov. Phys. JETP **13**, 349 (1961)].
10. G. V. Skornyakov and K. A. Ter-Martirosyan, Zh. Éksp. Teor. Fiz. **31**, 775 (1956) [Sov. Phys. JETP **4**, 648 (1956)].
11. P. F. Bedaque, H.-W. Hammer, and U. van Kolck, Phys. Rev. Lett. **82**, 463 (1999).
12. S. P. Merkur'ev and L. D. Faddeev, *Quantum Scattering Theory for Few-Body Systems* (Nauka, Moscow 1985).
13. N. Zh. Takibaev and F. M. Pen'kov, Yad. Fiz. **50**, 373 (1989) [Sov. J. Nucl. Phys. **50**, 234 (1989)].
14. F. M. Pen'kov and N. Zh. Takibaev, Yad. Fiz. **57**, 1300 (1994) [Phys. At. Nucl. **57**, 1232 (1994)].
15. V. Roudnev, S. Sofianos, and S. Yakovlev, physics/0204025; Few-Body Syst. (in press).
16. R. A. Aziz, F. R. W. McCourt, and C. C. K. Wong, Mol. Phys. **61**, 1487 (1987).
17. R. A. Aziz and M. J. Slaman, J. Chem. Phys. **94**, 8047 (1991).
18. A. R. Janzen and R. A. Aziz, J. Chem. Phys. **107**, 914 (1997).
19. J. R. Taylor, *Scattering Theory: the Quantum Theory of Nonrelativistic Collisions* (Wiley, New York, 1972; Mir, Moscow, 1975).
20. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 3rd ed. (Nauka, Moscow, 1974; Pergamon, New York, 1977).
21. E. Nielsen and J. H. Macek, Phys. Rev. Lett. **83**, 1566 (1999).
22. P. F. Bedaque, E. Braaten, and H.-W. Hammer, Phys. Rev. Lett. **85**, 908 (2000).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 493–504.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 545–557.

ATOMS, SPECTRA, RADIATION

Resonant Charge Exchange in Slow Collisions Involving Halogens and Oxygen¶

B. M. Smirnov

Joint Institute for High Temperatures, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 127412 Russia

e-mail: smirnov@oivtran.iitp.ru

Received February 19, 2003

Abstract—The coupling of electron momenta is considered for the resonant charge exchange process in slow collisions. Because the electron transfer in this process occurs at large distances between the colliding atomic particles, where ion–atom interactions are relatively weak, we can separate different types of interaction and find the character of coupling of the electron momenta in the quasi-molecule, consisting of the colliding ion and its atom, for real collision pairs. Since the real number of interaction types for colliding particles exceeds that used in the classical Hund coupling scheme, there are intermediate cases of momentum coupling outside the standard Hund scheme. This occurs for the resonant charge exchange involving halogens and oxygen where the quantum numbers of the quasi-molecule in the course of the electron transfer are the total momenta J and j of the colliding ion and atom and the projection M or M_J of the atom orbital or total momentum on the quasi-molecule axis. The ion–atom exchange interaction potential is independent of the ion fine state, and under these conditions, the resonant charge exchange process is not entangled with the rotation of electron momenta, as in case "a" of the Hund coupling. The partial cross section of the resonant charge exchange process depends on quantum numbers of the colliding particles. The average cross sections depend weakly on the coupling scheme.

1. INTRODUCTION

The process of resonant electron transfer in slow collisions of an ion and the parent atom results in transition of a valence electron from one core to another. In the simplest case of the transition of an *s*-electron, this process is determined by the interference of two electron terms of the quasi-molecule consisting of the colliding ion and the atom. Correspondingly, the probability of this transition is expressed through the energy difference for the even and odd quasi-molecule states and the cross section of this process [1]. In the case of the transition of a *p*-electron involving an ion and an atom with unfilled electron shells, the resonant charge process becomes more complex because the electron transfer can be entangled with the processes of rotation of electron momenta and transitions between fine states of the colliding particles. One can simplify the analysis of this process by constructing a hierarchy of interactions in the quasi-molecule and thus choosing a suitable case of the Hund coupling [2–4] that corresponds to certain quantum numbers of the quasi-molecule. Although this scheme is related to diatomic molecules when the distance between the nuclei is fixed, it can be extended to a quasi-molecule consisting of colliding particles [5–8]. According to the general method by Nikitin [5–8], the trajectory is then divided into several parts such that the Hund coupling of a certain type is realized in each part. If the transition range between different cases of the Hund coupling is narrow, one can construct the wave function of colliding particles and the *S*-matrix of the transition by sewing the wave functions on different sides of the transition range [7]. This allows one to separate different processes and to find the probabilities for the variation of quantum numbers of the colliding particles at a given collision trajectory.

This general scheme can be used in analyzing the resonant charge exchange process involving an ion and an atom with unfilled electron shells when the electron momenta can be coupled via different schemes, and the resonant charge exchange process can therefore be entangled with other processes (rotation of the electron momenta and transitions between fine structure states) in different ways. Indeed, within the framework of the classical Mulliken scheme of the momentum summation [2], three types of interactions are introduced for a quasi-molecule: the electrostatic interaction V_e is responsible for the energy splitting of different angular momentum projections onto the molecule axis; δ_f corresponds to the spin-orbit interaction and other relativistic interactions; and the rotational energy or the Coriolis interaction V_{rot} accounts for the interaction between the orbital and spin electron momenta with the rotation of the molecular axis. Depending on the ratio between these interaction energies, one can construct six cases of the Hund coupling [2–4]; each of these cases corresponds to a certain scheme of momentum summation

¶This article was submitted by the author in English.

Table 1. The cases of Hund coupling

Hund case	Relation	Quantum numbers
a	$V_e \gg \delta_f \gg V_{\text{rot}}$	Λ, S, S_n
b	$V_e \gg V_{\text{rot}} \gg \delta_f$	Λ, S, S_N
c	$\delta_f \gg V_e \gg V_{\text{rot}}$	Ω
d	$V_{\text{rot}} \gg V_e \gg \delta_f$	L, S, L_N, S_N
e	$V_{\text{rot}} \gg \delta_f \gg V_e$	J, J_N

and is characterized by certain quantum numbers of the diatomic molecule. These cases are used as model ones in the analysis of some transitions in atomic collisions [7–9].

For the resonant charge exchange in slow collisions, the electron transfer from one core to another proceeds at large distances between the colliding particles, where interactions of different types are weak. This simplifies the general analysis of the charge exchange process and gives additional experience in understanding the momentum coupling. As a result of this analysis, we find a number of interactions to be actually greater than within the framework of the Hund scheme. Indeed, the electrostatic interaction V_e includes the exchange interaction V_{ex} inside the atom, which leads to certain orbital momenta L of the atom and I of the ion, and to certain spins S and s of these atomic particles, that is, V_{ex} characterizes the energy splitting of states with different quantum numbers $LSls$. The long-range interaction $U(R)$ and the ion–atom exchange interaction $\Delta(R)$ are added to this. In addition, the fine splitting of levels refers to the atom (δ_a) and ion (δ_i) separately, and competition between all these interactions gives rise to many other cases of momentum coupling compared to the Hund coupling scheme. In analyzing this problem for real ion–atom systems, we deal with a restricted number of momentum couplings. Below, we consider this problem for the resonant charge exchange of halogens and oxygen in the case where the ions and atoms are found in the ground state and the collision energies vary from thermal ones up to tens of electronvolts (this energy range is of interest for a low-temperature plasma).

2. ASYMPTOTIC THEORY OF RESONANT ELECTRON TRANSFER

We first formulate a general method to analyze this problem. We use the asymptotic theory of the resonant charge exchange [9–12], where the electron transfer has the tunnel character and large impact parameters of collisions make the main contribution to the cross section of this process. A reciprocal value of a typical impact parameter of collisions is a small parameter in the asymptotic theory of the resonant charge exchange process. Expanding the cross section over this small parameter and restricting ourselves by two expansion terms, we can express the ion–atom exchange interaction potential and the cross section via the asymptotic parameters of the transferring electron in an isolated atom and quantum numbers of the ion and atom electron shells. In contrast to models, the asymptotic theory allows us to find the correct value of the cross section with an estimated accuracy. For p-electron transitions in the collision energy range under consideration, this accuracy is better than 10% [13, 14].

The asymptotic theory allows us to determine the ion–atom exchange interaction potential. The cross section of the electron charge exchange is then expressed via the exchange interaction potential for given quantum numbers of the quasi-molecule consisting of the colliding ion and atom at a given distance R between them. In constructing this interaction potential, we start from the Hund coupling scheme [2–4], represented in Table 1 together with the quantum numbers of the quasi-molecule for each case of Hund coupling. We introduce the following notation: **L** is the total electron angular momentum of the molecule, **S** is the total electron spin, **J** is the total electron momentum of the molecule, **n** is the unit vector along the molecular axis, **N** is the rotation momentum of nuclei, Λ is the projection of the angular momentum of electrons on the molecular axis, Ω is the projection of the total electron momentum **J** on the molecular axis, S_n is the projection of the electron spin on the molecular axis, and L_N, S_N, and J_N are projections of these momenta onto the direction of the rotation momentum **N** of the nuclei. We adopt this scheme as a basis in considering the momentum coupling of the quasi-molecule consisting of a colliding ion and the parent atom.

We note two momentum coupling schemes for the atom and ion [4, 15, 16]. In the LS scheme, which is realized under $V_{\text{ex}} \gg \delta_a, \delta_i$, we have LM_LS as the quantum numbers of the atom and lms as the quantum numbers of the ion, where L and l are the angular momenta, M and m are their projections onto the molecular axis, and S and s are the spins of the atom and the ion, respectively. In the j–j coupling scheme for an individual atomic particle, which occurs at $V_{\text{ex}} \ll \delta_a, \delta_i$, we use JM_J as the atom quantum numbers and jm_j as the ion quantum numbers, where J and j are the total electron momenta and M_J and m_j are their projections on the molecular axis for the atom and the ion respectively. These quantum numbers are the basis for the limiting cases of electron terms of the quasi-molecule.

The possible interaction potentials in the quasi-molecule under consideration are

$$V_{\text{ex}}, \quad U_M = \frac{Q_{MM}}{R^3}, \qquad U_m = \frac{Q_{MM}q_{mm}}{R^5}, \ \Delta(R), \ \delta_i, \ \delta_a, \text{ and } V_{\text{rot}}. \tag{1}$$

We divide the electrostatic interaction V_e in Table 1 into four parts: the exchange interaction V_{ex} inside the atom and ion responsible for electrostatic splitting of levels inside the isolated atom and ion with given electron shells; the long-range interaction U_M of the ion with the quadruple atom moment; the long-range interaction U_m responsible for splitting of the ion level; and the ion–atom exchange interaction potential Δ that determines the resonant charge exchange cross section. The fine splitting δ_f of levels in Table 1 is written separately for the ion (δ_i) and the atom (δ_a). Here, M and m are the projections of the atom and ion angular momenta onto the molecular axis, R is the ion–atom distance, Q_{ik} is the tensor of the atom quadruple moment, and q_{ik} is the quadruple moment tensor of the ion. As can be seen, the number of possible coupling cases increases significantly in this description in comparison with the classical case. Of course, a small part of these cases can be realized, and we verify this below for certain cases of resonant charge exchange.

To find the suitable momentum coupling scheme, we evaluate the above ion–atom interaction potentials at distances that determine the resonant charge exchange cross section. Constructing the hierarchy of interactions, we find the quasi-molecule quantum numbers in this distance range and the partial cross sections corresponding to these quantum numbers. This allows us to ascertain the momentum summation scheme in slow ion–atom collisions with a resonant electron transfer. We note that the character of momentum coupling influences the value of the average cross sections. Below, we realize this operation for certain cases of ion–atom collisions where the colliding ion and atom are found in the ground electron states.

3. HIERARCHY OF ION–ATOM INTERACTIONS FOR HALOGENS

We start with the resonant electron transfer involving a halogen atom and an ion in the ground electron states,

$$X^+ + X \longrightarrow X + X^+, \tag{2}$$

where X is the halogen atom. In this case, the atom and ion ground states are $(p^5)^2P$ and $(p^4)^3P$, respectively, and all the interactions in (1) are therefore realized in this case. In Table 2, we collect some parameters of the colliding atomic particles in this case. We note that the lower fine structure states include states with the total electron momenta $j = 2, 1, 0$. The parameter δ_i in Table 2 is the splitting of the $^2P_2$–$^2P_0$ ion levels. Next, the value V_{ex} is the splitting between the ground ion level $^2P_2$ and the level $^1D_2$ in the notation of the *LS* momentum coupling scheme. Because the ratio δ_i/V_{ex} is small for all ions, the *LS* momentum coupling scheme is valid in the ion, and we adopt it as a basis.

As a characteristic of the resonant charge exchange process, we take the average cross section σ_{ex} of the resonant charge exchange in the case "a" of the Hund coupling [13, 14, 19], and the corresponding impact parameter R_0 of collision is determined from the relation

$$\sigma_{ex} = \frac{\pi}{2}R_0^2. \tag{3}$$

In Table 2, we give the values of R_0 (in units of the Bohr radius a_0) at the collision energies 0.1, 1 (in parentheses), and 10 eV (in square brackets). These energies

Table 2. Parameters of halogen atoms and ions [17, 18]

	F	Cl	Br	I
δ_a, cm^{-1}	404	882	3685	7603
δ_i, cm^{-1}*	490	996	3840	7087
V_{ex}, cm^{-1}	20873	11654	11410	13727
δ_i/V_{ex}	0.023	0.085	0.34	0.52
R_0, a_0	11.7 (10.6) [9.54]	15.1 (13.8) [12.3]	16.5 (15.1) [13.6]	19.1 (17.2) [15.8]
$\overline{r^2}$, a_0^2	1.54	4.06	5.22	7.20
U_M, cm^{-1}	253 (341) [467]	311 (407) [575]	306 (448) [546]	272 (372) [481]
U_M/δ_a	0.63 (0.84) [1.16]	0.35 (0.46) [0.65]	0.08 (0.12) [0.15]	0.036 (0.049) [0.063]
V_{rot}, cm^{-1}	8.6 (30) [106]	4.9 (17) [60]	3.0 (10) [36]	2.0 (7.1) [25]
$\Delta(R_0)$, cm^{-1}	7.0 (23) [78]	4.2 (14) [46]	2.6 (8.4) [29]	1.8 (6.1) [21]

* Energy differences for levels of the states $^3P_2$ and $^3P_0$.

pertain to the laboratory reference frame, where the atom is motionless and the parent ion has its energy indicated. Using these values of R_0, we evaluate various interaction potentials and refer them to a given collision energy. The value U_M in Table 2 is equal to

$$U_M = U_{00} - U_{11} = \frac{Q_{00} - Q_{11}}{R_0^3} = \frac{6e^2\overline{r^2}}{5R_0^3}, \quad (4)$$

where e is the electron charge and r is the distance of a valence atomic p-electron from the nucleus inside the atom. In Table 2, we list the corresponding values of U_M and the ratio U_M/δ_a, which is usually less than unity, and therefore the fine structure of level splitting is important for processes involving halogens.

If the colliding particles move along straight trajectories, the quasi-molecule rotation energy is given by

$$V_{\text{rot}} = \frac{\hbar v}{R_0} \quad (5)$$

at the closest approach and at the impact parameter R_0 of the ion–atom collision, where v is the relative ion–atom velocity. According to the data in Table 2, the rotation energy is smaller than the other interaction potentials (U_M, δ_i, δ_a). This determines the character of momentum coupling in this case.

Based on the above analysis, we can construct a hierarchy of interactions for a quasi-molecule consisting of a halogen ion and the parent atom at the distances between these particles that determine the cross section of resonant charge exchange (2). The following hierarchy of interactions is valid for more or less all halogens in the range of collision energies 0.1–10 eV:

$$V_{\text{ex}} \gg \delta_i, \delta_a \gg U_M \gg U_m, V_{\text{rot}}. \quad (6)$$

In terms of the data in Table 1, this is an intermediate case between cases "a" and "c" of the Hund coupling. In addition, we evaluate the exchange ion–atom interaction potential $\Delta(R)$ using the formula for the resonant charge exchange cross section σ_{ex} for the transition of an s-electron [1, 10, 11],

$$\sigma_{\text{ex}} = \frac{\pi R_0^2}{2}, \quad (7a)$$

where

$$\frac{1}{v}\sqrt{\frac{\pi R_0}{2\gamma}}\Delta(R_0) = 0.28 \quad (7b)$$

and γ is the asymptotic parameter of the wave function of the transferred valence electron ($\Delta(R) \propto \exp(-\gamma R)$). From this, we can compare the exchange interaction potential $\Delta(R_0)$ at a distance of R_0 (Eq. (3)) with rotation energy (5) of the quasi-molecule at the impact parameter R_0 of collision and the minimal distance between the colliding ion and atom,

$$\Delta(R_0) = 0.28v\sqrt{\frac{2\gamma}{\pi R_0}} = 0.22\sqrt{\gamma R}V_{\text{rot}}. \quad (8)$$

A small parameter of the asymptotic theory is

$$\frac{1}{\gamma R_0} \ll 1. \quad (9)$$

At collision energies of several electronvolts, we have $\gamma R_0 \approx 10$–15, and the above values are therefore comparable, $\Delta(R_0) \sim V_{\text{rot}}$. The values of the exchange interaction potential given in Table 2 confirm this statement.

The hierarchy of interactions in (6) leads to the quantum numbers $LSJM_Jlsjm_j$ of the quasi-molecule, where L and l are the atom and ion angular momenta, S and s are the atom and ion spins, J and j are the total electron momenta of the atom and ion, and M_J and m_j are their projections onto the molecular axis. The wave function is then given by

$$\Psi_{LSJM_Jlsjm_j} = \sum_{MM_Smm_s} \begin{bmatrix} L & S & J \\ M & M_S & M_J \end{bmatrix} \times \begin{bmatrix} l & s & j \\ m & m_s & m_j \end{bmatrix} \Phi_{LMSM_S}\psi_{lmsm_s}, \quad (10)$$

where Φ and ψ are the respective wave functions of the weakly interacting atom and ion; the atom ($LMSM_S$) and ion ($lmsm_s$) quantum numbers are given with the spin-orbit and other relativistic interactions neglected.

Guided by the hierarchy (6) of interactions, we now find positions of the energy levels for the quasi-molecule under consideration and apply this to the case of halogen atoms. In the first approach, $LSls$ are the quasi-molecule quantum numbers, and in the case of halogen atoms, we are restricted by the lowest electron terms ^2P for the atom and ^3P for the ion. The second approach gives the quantum numbers J and j of the quasi-molecule with the splitting between the fine-structure levels determined by the corresponding values δ_a and δ_i for the isolated atom and ion. The third approach leads to the quantum number M_J of the quasi-molecule, i.e., the projection of the total atom electron moment on the molecular axis; the splitting between the levels with

different M_J is then determined by the interaction of the ion charge and the atom quadruple moment,

$$\Delta U(M_J) = \left\langle \Psi_{LSJM_{J_j}lsjm_j} \left| \frac{Q_{MM}}{R^3} \right| \Psi_{LSJM_{J_j}lsjm_j} \right\rangle = \sum_{MM_S} \frac{Q_{MM}}{R^3} \begin{bmatrix} L & S & J \\ M & M_S & M_J \end{bmatrix}^2. \quad (11)$$

For the interaction of the halogen atom and the ion $X(^2P) + X^+(^3P)$, where X is the halogen atom, this formula becomes

$$\Delta U(JM_Jjm_j) = V_0 \left(2 \begin{bmatrix} L & S & J \\ 0 & M_J & M_J \end{bmatrix}^2 - \begin{bmatrix} L & S & J \\ 1 & M_{J-1} & M_J \end{bmatrix}^2 \right), \quad U_0 = \frac{2\overline{r_a^2}}{5R^3}, \quad (12)$$

where Q_{MM} is the component of the quadruple moment tensor of the atom, r_a is the distance of the valence electron from the nucleus, and the bar denotes the average over electron positions in the atom.

The fourth approach corresponds to the quantum number m_j of the quasi-molecule, with the interaction potential between the ion and atom quadruple momenta given by

$$\Delta U(JM_Jjm_j) = \frac{1}{R^5} \sum_{MM_S} Q_{MM} \begin{bmatrix} L & S & J \\ M & M_S & M_J \end{bmatrix}^2 \times \sum_{mm_s} q_{mm} \begin{bmatrix} l & s & j \\ m & m_s & m_j \end{bmatrix}^2, \quad (13)$$

where q_{mm} is the component of the ion quadruple moment tensor. We note that the electron terms of the quasi-molecule under consideration are degenerate with respect to the sign of the total momentum projections. For the interaction of atomic particles $X(^2P) + X^+(^3P)$, where X is a halogen atom, this formula can be rewritten as

$$\frac{\Delta U(JM_Jjm_j)}{V_0} = -\left(2 \begin{bmatrix} L & S & J \\ 0 & M_J & M_J \end{bmatrix}^2 - \begin{bmatrix} L & S & J \\ 1 & M_J - 1 & M_J \end{bmatrix}^2 \right) \left(2 \begin{bmatrix} l & s & j \\ 0 & m_j & m_j \end{bmatrix}^2 - \begin{bmatrix} l & s & j \\ 1 & m_j - 1 & m_j \end{bmatrix}^2 \right), \quad (14)$$

Table 3. The lowest energy E of the quasi-molecule Cl^+–Cl at a distance of $R = 14a_0$ between the nuclei. The statistical weight g refers only to even (or odd) quasi-molecular states

JM_Jj	g	$\Delta U(M_J)$, cm^{-1}	E, cm^{-1}
$\frac{3}{2}\frac{3}{2}2$	20	–130	0
$\frac{3}{2}\frac{1}{2}2$	20	130	260
$\frac{3}{2}\frac{3}{2}1$	12	–130	696
$\frac{3}{2}\frac{1}{2}1$	12	130	956
$\frac{3}{2}\frac{3}{2}0$	4	–130	996
$\frac{1}{2}\frac{1}{2}2$	20	0	1102
$\frac{3}{2}\frac{1}{2}0$	4	130	1256
$\frac{1}{2}\frac{1}{2}1$	12	0	1708
$\frac{1}{2}\frac{1}{2}0$	4	0	2009

where

$$V_0 = \frac{4\overline{r_a^2}\overline{r_i^2}}{25R^5} \quad (15)$$

and the distances r_a and r_i pertain to the atom and ion, respectively.

As an example, we construct the lowest-energy levels for the chlorine ion–atom system at $R = 14a_0$ that characterizes the resonant charge exchange cross section at a collision energy 1 eV (see Table 2). At this distance, the energy of charge-quadruple interaction (Eq. (12)) and quadruple-quadruple interaction (Eq. (15)) are $U_0 = 130$ cm^{-1} and $V_0 = 0.6$ cm^{-1}. In this case, we therefore have the following hierarchy of interactions (6):

$$V_{\text{ex}} \gg \delta_i, \delta_a \gg U_M \gg \Delta, V_{\text{rot}} \gg U_m. \quad (16)$$

In Table 3, we list the level energies E for chlorine given by

$$E = \delta_i + \delta_a + U_M + \varepsilon_0, \quad (17)$$

where ε_0 is taken such that the lowest electron term has zero energy, $E = 0$. The quasi-molecule energies in Table 3 pertain to the ground electron state of the atom and ion, i.e., $L = 1$, $l = 1$, $S = 1$, $s = 1/2$ for this term. In this approximation, we include the quantum numbers JM_Jj for interacting particles; the states with other quantum numbers, i.e., $LSls$, are characterized by higher energies. Indeed, the excitation energy of the ion state $^1D_2$ is 11654 cm^{-1} and the excitation energy of the ion state $^1S_0$ is 27878 cm^{-1}; these ion states pertain to the same electron shell $3p^4$. A nonexcited electron shell of halogen atoms is characterized by one electron term $L = 1$, $s = 1/2$, which simplifies the analysis.

The data in Table 3 are obtained with the interaction potentials Δ, V_{rot}, and U_m neglected. These potentials give additional quantum numbers for the quasi-molecule, and therefore the accuracy of the data in Table 3 is determined by these values: $\Delta(R) \sim V_{\text{rot}} \sim 10$ cm^{-1} and $U_m \sim 1$ cm^{-1}. In this approximation, the statistical weight of the quasi-molecule states is

$$g = 2 \times 2(2j + 1), \tag{18}$$

where the first factor accounts for the degeneration with respect to the sign of M_J, the second factor corresponds to the separation of quasi-molecule states into odd and even ones, and this degeneration is therefore removed by the exchange interaction $\Delta(R)$; and the third factor in (18) accounts for the degeneration with respect to m_j, which is removed by all the neglected interactions Δ, V_{rot}, and U_m.

Thus, the above analysis of the interaction of a halogen ion and atom at large separations shows that the character of electron momentum coupling differs from that of the Hund coupling scheme. Along with the quantum numbers of electron shells of an isolated atom and an ion in the framework of the LS coupling scheme, the quantum numbers of the quasi-molecule are JjM_J (the total electron momenta of the atom and the ion and the projection of the total atom momentum on the molecular axis). Other quantum numbers are mixed due to the rotation energy V_{rot}, exchange Δ, and quadruple-quadruple interaction potentials U_m between the atom and the ion.

4. ION–ATOM EXCHANGE INTERACTION FOR HALOGENS

We now determine the exchange ion–atom interaction potential that allows us to evaluate the resonant charge exchange cross section. For this, we represent the wave function of the atom having n valence electrons with the momentum l_e in the framework of the LS coupling scheme as [15, 16, 20]

$$\begin{aligned}\Phi_{LSM_LM_S}(1, 2, \ldots, n) &= \frac{1}{\sqrt{n}}\hat{P} \\ &\times \sum_{lmsm_s\mu\sigma} G_{ls}^{LS}(l_e, n)\begin{bmatrix} l_e & l & L \\ \mu & m & M_L \end{bmatrix} \\ &\times \begin{bmatrix} \frac{1}{2} & s & S \\ \sigma & m_s & M_S \end{bmatrix} \varphi_{l_e\frac{1}{2}\mu\sigma}(1)\psi_{lsmm_s}(2, \ldots, n),\end{aligned} \tag{19}$$

where Φ, ψ, and φ are the respective wave functions of the atom, the ion, and the valence electron with the quantum numbers indicated; μ and σ are the projections of the angular momentum and spin of the valence electron; the argument of the wave function indicates the electrons contained by this atomic particle; the operator $\hat{P}$ permutes the electrons; and the parentage coefficient $G_{ls}^{LS}(l_e, n)$ is responsible for addition of the valence electron to the ion for construction of an atom for given quantum numbers of these atomic particles.

The exchange interaction potential is given by [8, 12]

$$\begin{aligned}\Delta(R) &= 2\langle \Psi_1|\hat{H}|\Psi_2\rangle \\ &- 2\langle \Psi_1|\hat{H}|\Psi_1\rangle\langle \Psi_1|1|\Psi_2\rangle,\end{aligned} \tag{20}$$

where Ψ_1 is the wave function of the quasi-molecule with the valence electron located near the first core (the atom is located near the first nucleus), Ψ_2 corresponds to the atom location near the second nucleus, and $\hat{H}$ is the Hamiltonian of electrons. We note that an accurate evaluation of this interaction requires the accurate wave functions of the quasi-molecule that take into account the interaction of the valence electron with both cores simultaneously; this is achieved in the framework of the asymptotic theory. Using the general method to evaluate the exchange interaction potential $\Delta(R)$ similarly to case "a" of the Hund coupling in [12, 14, 19, 21], we obtain

$$\begin{aligned}\Delta(R) &= n(G_{ls}^{LS})^2 \sum_{\substack{\mu mm'MM' \\ \sigma\sigma' m_s m_s' m_s'' \\ M_S M_S' m''' m_s'''}} \begin{bmatrix} l_e & l & L \\ \mu & m' & M \end{bmatrix} \\ &\times \begin{bmatrix} \frac{1}{2} & s & S \\ \sigma & m_s' & M_S \end{bmatrix}\begin{bmatrix} L & S & J \\ M & M_S & M_J \end{bmatrix}\begin{bmatrix} l & s & j \\ m & m_s & m_j \end{bmatrix}\begin{bmatrix} l_e & l & L \\ \mu & m'' & M' \end{bmatrix} \\ &\times \begin{bmatrix} \frac{1}{2} & s & S \\ \sigma' & m_s'' & M_S' \end{bmatrix}\begin{bmatrix} L & S & J \\ M' & M_S' & M_J \end{bmatrix}\begin{bmatrix} l & s & j \\ m''' & m_s''' & m_j \end{bmatrix}\Delta_{l_e\mu}.\end{aligned} \tag{21}$$

We here take into account the character of coupling of the electron momenta in the quasi-molecule, such that the atomic core quantum numbers $lsm'm_s'$ and the valence electron atomic numbers $l_e\mu\frac{1}{2}\sigma$ are first summed in the atomic quantum numbers LSM_LM_S and the atom quantum numbers are then summed over the quantum numbers $LSJM_J$; and the ion quantum numbers $lsmm_s$ are summed over the ion quantum numbers lsj. We sum or average over the other quasi-molecule quantum numbers and use the relations

$$\sum_{mm_sm'''m_s'''}\begin{bmatrix} l & s & j \\ m & m_s & m_j\end{bmatrix}\begin{bmatrix} l & s & j \\ m''' & m_s''' & m_j\end{bmatrix} = 1,$$

$$\sum_{\sigma\sigma' m_s'm_s''}\begin{bmatrix} \frac{1}{2} & s & S \\ \sigma & m_s' & M_S\end{bmatrix}\begin{bmatrix} \frac{1}{2} & s & S \\ \sigma' & m_s' & M_S'\end{bmatrix} = \delta_{M_SM_S'}$$

for the Clebsch–Gordan coefficients. In Eq. (21), $\Delta_{l_e\mu}$ is the one-electron exchange interaction potential that corresponds to the case where a valence electron with these quantum numbers is located in the field of two structureless cores and has the same asymptotic wave function as in real atoms. As a result, we find by analogy with [6, 12, 14, 19, 21] that

$$\Delta(l_e\mu, ls, LS, JM_Jj, R) = n(G_{ls}^{LS})^2 \times \sum_{\mu mMM_S}\begin{bmatrix} l_e & l & L \\ \mu & m & M\end{bmatrix}^2\begin{bmatrix} L & S & J \\ M & M_S & M_J\end{bmatrix}^2\Delta_{l_e\mu}(R), \quad (22)$$

where the argument contains the quantum numbers of the quasi-molecule and the distance R between the interacting ion and atom. This formula reduces the problem of the exchange interaction between an atom and an ion with unfilled electron shells to the transition of one electron between structureless cores. It is important that the exchange interaction potential is independent of the ion momentum j.

The one-electron exchange interaction potential $\Delta_{l_e\mu}$ is given by [6, 8, 12, 22]

$$\Delta_{l_e\mu}(R) = A^2R^{2/\gamma-1-|\mu|}\exp\left(-R\gamma-\frac{1}{\gamma}\right) \times \frac{(2l+1)(l_e+|\mu|)!}{(l_e-|\mu|)!|\mu|!(\gamma)^{|\mu|}}. \quad (23)$$

It decreases with the increase of μ as $R^{-|\mu|}$. Here, l_e and μ are quantum numbers of the valence electron and γ and A are the parameters of the asymptotic wave function of this electron. This formula contains the first term of the asymptotic expansion over the small parameter $1/\gamma R$ for the ion–atom exchange interaction potential at large distances between the nuclei.

From (22), we thus obtain the exchange interaction potential involving the halogen atom X(^2P) and its ion $X^+(^3P)$,

$$\Delta(l_e\mu, ls, LS, JM_Jj, R) = 3\Delta_{10}(R) \times \sum_M\begin{bmatrix} 1 & 1 & 1 \\ 0 & M & M\end{bmatrix}^2\begin{bmatrix} 1 & 1/2 & J \\ M & M_J-M & M_J\end{bmatrix}^2 = \frac{3\Delta_{10}(R)}{2}\begin{bmatrix} 1 & 1/2 & J \\ 1 & M_J-1 & M_J\end{bmatrix}^2, \quad (24)$$

where we extract the dominant term in the sum in (22) that is proportional to $\Delta_{10}(R)$ (see Eq. (23)). In Table 4, we give the values of the exchange interaction potential for the ground electron states of the halogen atom X(^2P_J) and ion $X^+(^3P_J)$ with different fine-structure quantum numbers for these particles.

To demonstrate these results, we return to the above example of the interaction $Cl(^2P) + Cl^+(^3P)$ at a distance of $R = 14a_0$ between the nuclei. The energy splittings between even and odd quasi-molecule states are $\Delta_{10} = 14$ cm^{-1} and $\Delta_{11} = 2.0$ cm^{-1} if we consider the cores structureless. Table 4 contains the values of the exchange interaction potential under these conditions for given quantum numbers of the interacting particles. We ignore the quadruple-quadruple ion–atom interaction and the rotation energy; the energy of the even or odd state with given quantum numbers is $E \pm \Delta(R)/2$. The data in Table 4 confirm the above hierarchy of interactions between halogen atoms and their ions.

5. RESONANT CHARGE EXCHANGE FOR HALOGENS

The above results allow us to determine the resonant charge exchange cross section in slow collisions of halogen atoms and their ions in the ground electron states. To determine the partial cross section of the resonant charge exchange, we use the asymptotic formula [10, 11] (see Eq. (8))

$$\sigma_{\text{res}} = \frac{\pi R_0^2}{2}, \qquad \frac{1}{v}\sqrt{\frac{\pi R_0}{2\gamma}}\Delta(R_0) = 0.28, \quad (25)$$

where v is the collision speed and the asymptotic coefficient γ is expressed via the atom ionization potential I as $\gamma = \sqrt{2I}$ in atomic units (also see Eq. (23)). Equation (25) is valid for s-electron transitions or in the case where electron transfer transitions for states with given quantum numbers can be separated from other transi-

Table 4. The exchange interaction potential $\Delta(R)$ for the halogen atom and ion in the ground electron states, $Cl(^2P_{JM_J}) + Cl^+(^3P_j)$, and different states of the fine structure JM_Jj for these particles. The total energy E of these states (Table 3) refers to the distance $R = 14a_0$ between the nuclei and is obtained ignoring the quadruple-quadruple ion–atom interaction and the rotation energy of the quasi-molecule

JM_Jj	$\Delta(R)$	Δ, cm^{-1}	E, cm^{-1}	σ_{ex}, Å^2 ε = 0.1 eV	σ_{ex}, Å^2 ε = 1 eV	σ_{ex}, Å^2 ε = 10 eV
$\frac{3}{2}\frac{3}{2}2$	$\frac{3}{2}\Delta_0$	87	0	110	92	76
$\frac{3}{2}\frac{1}{2}2$	$\frac{1}{2}\Delta_0$	29	260	93	77	62
$\frac{3}{2}\frac{3}{2}1$	$\frac{3}{2}\Delta_0$	87	696	110	92	76
$\frac{3}{2}\frac{1}{2}1$	$\frac{1}{2}\Delta_0$	29	956	93	77	62
$\frac{3}{2}\frac{3}{2}0$	$\frac{3}{2}\Delta_0$	87	996	110	92	76
$\frac{1}{2}\frac{1}{2}2$	Δ_0	58	1012	104	86	71
$\frac{3}{2}\frac{1}{2}0$	$\frac{1}{2}\Delta_0$	29	1256	93	77	62
$\frac{1}{2}\frac{1}{2}1$	Δ_0	58	1708	104	86	71
$\frac{1}{2}\frac{1}{2}0$	Δ_0	58	2009	104	86	71

tions. In particular, the partial cross sections of resonant charge exchange are given in Table 4 for chlorine.

We introduce the resonant charge exchange cross section averaged over fine states assuming the initial population of atom and ion fine states to be proportional to their statistical weights,

$$\overline{\sigma_{ex}} = \frac{1}{3}\sigma_{ex}\left(\frac{3}{2},\frac{3}{2}\right) + \frac{1}{3}\sigma_{ex}\left(\frac{3}{2},\frac{1}{2}\right) + \frac{1}{3}\sigma_{ex}\left(\frac{1}{2},\frac{1}{2}\right), \quad (26)$$

where the atom quantum numbers J and M_J are given in the partial cross section argument. If we expand the resonant charge exchange cross section over the small parameter $1/R\gamma$, keep only two terms in the expansion, and take Eq. (24) into account, we can write Eq. (25) for average cross section (26) as

$$\sigma_{res} = \frac{\pi R_0^2}{2}, \quad (27a)$$

where

$$\frac{2.7}{v}\sqrt{\frac{\pi R_0}{2\gamma}}\Delta_{00}(R_0) = 0.28 \quad (27b)$$

and $\Delta_{00}(R)$ is the ion–atom exchange interaction potential for the transferred s-electron with the given asymptotic parameters γ and A of its wave function. This value is related to the one-electron exchange interaction potential $\Delta_{10}(R)$ for a transferred p-electron $\Delta_{10}(R) =$

Table 5. The average cross sections (in 10^{-15} cm^2) for the halogen atom and ion in the ground electron states $X(^2P) + X^+(^3P)$ at the indicated collision energies ε in the laboratory reference frame for hierarchy (6) of interactions and in the case "a" of the Hund coupling [13, 19] (in parentheses)

	ε = 0.1 eV	ε = 1 eV	ε = 10 eV
F (γ = 1.132, A = 1.6)	6.2 (6.0)	5.1 (4.9)	4.1 (4.0)
Cl (γ = 0.976, A = 1.8)	10 (10)	8.7 (8.4)	7.1 (6.9)
Br (γ = 0.932, A = 1.8)	13 (12)	11 (10)	8.9 (8.2)
I (γ = 0.876, A = 1.9)	16 (16)	14 (13)	12 (11)

$3\Delta_{00}(R)$ in accordance with (23). Table 5 contains the average cross sections of resonant charge exchange for halogen atoms and their ions in the ground electron state, $X(^2P) + X^+(^3P)$, for hierarchy (6) of interactions for the quasi-molecule constituted by the colliding atom and ion. These cross sections practically coincide with the average cross sections for the ground fine states of the colliding particles, i.e., for the process $X(^2P_{3/2}) + X^+(^3P_2)$. Thus, averaging over fine states of the ground electron states and over momentum projections of the ground fine states of colliding particles leads to results that are close to each other. In addition, these data are compared with the cross sections in case "a" of the Hund coupling taken from [13, 14]. As follows from the comparison, the real hierarchy of interactions in a quasi-molecule increases the resonant charge exchange cross section by several percent compared with case "a" of the Hund coupling.

One more feature of the resonant charge exchange for momentum coupling follows from hierarchy (6) of the interactions. The exchange interaction potential that determines the cross section of this process is given by (24), where we restrict ourselves to only a transition of a *p*-electron with zero momentum projection on the molecular axis. As follows from this formula, such states are present in any fine-structure state, and we can therefore ignore the transition of the electron whose momentum projection on the molecular axis is unity. We note that in contrast to the case "a" of the Hund coupling, where rotation of the molecular axis leads to transitions between states with different momentum projections on the molecular axis, such transitions are absent in the case of halogen atoms and ions because of separation of fine-structure states by energy. Next, we evaluate the resonant charge exchange cross section with accounting for the coupling between the transferred electron momentum and the momenta of atomic cores. We can estimate the error in the cross sections if we compare the cross sections with and without the momentum coupling taken into account. We take a *p*-electron that has the same asymptotic parameters as valence electrons of halogen atoms, but which is located in the field of structureless cores. For the resonant charge exchange cross section in chlorine at the respective energies 0.1, 1, and 10 eV, we then obtain the values 87, 71, and 57 Å instead of those in Table 5. We see that ignoring the coupling between the momenta of the transferred electron and atomic cores leads to a significant error.

Thus, it follows from the above analysis that in the course of collision and electron transfer, a quasi-molecule consisting of the colliding halogen ion and atom is characterized by the quantum numbers JM_Jj, and transitions between these states are absent during the electron transfer. The partial cross sections of resonant charge exchange depend on quantum numbers, whereas the average cross sections for the correct scheme of momentum coupling and in case "a" of the Hund coupling are close to each other.

6. RESONANT CHARGE EXCHANGE FOR OXYGEN

We now consider one more example of resonant charge exchange with a *p*-electron transition,

$$O^+(^4S) + O(^3P) \longrightarrow O(^3P) + O^+(^4S), \qquad (28)$$

involving the oxygen atom and ion in the ground electron states. Constructing the hierarchy of interactions (1) in this case, we take the previous case in Eq. (6) as a basis, with the quantum numbers JM_J of the quasi-molecule consisting of the colliding particles. In accordance with (12), the interaction potential of the ion charge with the atom quadruple moment is then given by

$$\Delta U(JM_J) = \frac{2\overline{r_a^2}}{5R^3} \times \left(2\begin{bmatrix} 1 & 1 & J \\ 0 & M_J & M_J \end{bmatrix}^2 - \begin{bmatrix} 1 & 1 & J \\ 1 & M_J - 1 & M_J \end{bmatrix}^2 \right) = \frac{2\overline{r_a^2}}{5R^3} a(JM_J), \qquad (29)$$

where we use the same notation as in (12); the values $a(JM_J)$ are contained in Table 6. It follows from (22) that instead of Eq. (24) for the ion–atom exchange interaction potential for halogens, the exchange interaction potential for oxygen is given by

$$\Delta(J(M_J, R)) = \frac{4}{3}\sum_{\mu} \begin{bmatrix} 1 & 1 & J \\ \mu & M_J - \mu & M_J \end{bmatrix}^2 \Delta_{1\mu}(R), \qquad (30)$$

with the coefficients in this formula given in Table 6.

We note that the excitation energies of oxygen atom fine states from the ground fine state $^3P_2$ are 158 cm^{-1} for the state $^3P_1$ and 220 cm^{-1} for the state $^3P_0$. These values are comparable to the long-range ion-quadruple interaction potential (29) at distances that make the main contribution to the resonant charge exchange cross section. Hence, we have an intermediate case of momentum coupling for oxygen, and in the above-mentioned halogen example with $\delta_a \gg \Delta U$, the quantum numbers of the quasi-molecule are JM_Jj, while in the other limiting case where $\delta_a \ll \Delta U$, the quantum numbers of the quasi-molecule are JMj (where M and M_J are the projections of the orbital and total atom momentum on the quasi-molecule axis). We consider the first limiting case below; Table 6 contains the values of the quasi-molecule energies E calculated in accordance with Eq. (17), where we take into account the fine-structure splitting of levels and charge-quadruple ion–atom interaction in the limit $\delta_a \gg \Delta U$. The energies are taken at an ion–atom distance of $R = 12a_0$ corre-

Table 6. The interaction potentials for the system $O^+(^4S_{3/2})$–$O(^3P_{JM_J})$ at $R = 12a_0$ and the partial cross sections σ_{ex} of resonant charge exchange at the indicated collision energy ε

JM_J	$a(JM_J)$	$\Delta(JM_J)$	E, cm^{-1}	Δ, cm^{-1}	σ_{ex}, Å^2 ε = 0.1 eV	σ_{ex}, Å^2 ε = 1 eV	σ_{ex}, Å^2 ε = 10 eV
22	–1	$\frac{4}{3}\Delta_0$	0	6.8	63	51	40
21	1/2	$\frac{2}{3}\Delta_0$	153	20	77	62	49
20	7/6	$\frac{8}{9}\Delta_0$	220	27	81	66	52
11	1/2	$\frac{2}{3}\Delta_0$	311	20	77	62	49
10	–1/2	$\frac{2}{3}\Delta_1$	209	4.4	56	44	34
00	1/3	$\frac{4}{9}\Delta_0$	363	13	72	57	45

sponding to the impact parameters typical of the resonant charge exchange cross section at the collision energy about 1 eV. We note that rotation energy (5) is 29 cm^{-1} under these conditions and exceeds or is comparable to the exchange interaction potential.

The average resonant charge exchange cross section is

$$\overline{\sigma_{ex}} = \frac{2}{9}\sigma_{ex}(22) + \frac{2}{9}\sigma_{ex}(21) + \frac{1}{9}\sigma_{ex}(20) + \frac{2}{9}\sigma_{ex}(11) + \frac{1}{9}\sigma_{ex}(10) + \frac{1}{9}\sigma_{ex}(00), \quad (31)$$

where the quantum numbers of the fine-structure atom state are given in parentheses, and we assume the population of these states to be proportional to their statistical weights. If we ignore the electron transitions due to rotation of the molecular axis, we find by analogy with Eqs. (25) and (27) that the average cross section is given by

$$\sigma_{res} = \frac{\pi R_0^2}{2}, \quad (32a)$$

where

$$\frac{2.9}{v(R_0\gamma)^{1/3}}\sqrt{\frac{\pi R_0}{2\gamma}}\Delta_{00}(R_0) = 0.28, \quad (32b)$$

and we use the same notation as in (25) and (27). Averaging the cross sections in Table 6 in accordance with (31) and (32) gives the values 71, 57, and 45 Å^2 for the average cross section of resonant charge exchange at the respective collision energies 0.1, 1, and 10 eV. On the other hand, the respective cross sections in case "a" of the Hund coupling are given by 73, 60, and 48 Å^2 [13, 14]. Similarly, for the partial cross sections for the ground fine-structure state, i.e., for the process $O^+(^4S_{3/2}) + O(^3P_2)$, we obtain the values 72, 55, and 46 Å^2 at the respective collision energies of 0.1, 1, and 10 eV; the difference of these average cross sections from the above average cross sections does not exceed their accuracy. Next, if we ignore the coupling of the transferred electron with cores, i.e., if we consider the transition of a *p*-electron between structureless cores at the same asymptotic parameters of the electron wave function as in the above cases, we obtain the values 64, 51, and 40 Å^2 for the average cross sections of resonant charge exchange at the indicated collision energies.

The other limiting case of the interaction hierarchy, $\delta_a \ll U_0$, between the fine splitting of atom levels and the ion–atom quadruple interaction potential leads to the molecular quantum numbers *JMj*, where *M* is the projection of the atom angular momentum on the quasimolecule axis. In this limiting case, Eq. (22) for the ion–atom exchange interaction potential for process (28) becomes

$$\Delta(R) = \frac{4}{3}\sum_{\mu M_S}\begin{bmatrix}1 & 1 & 1\\ \mu & M-\mu & M\end{bmatrix}^2 \times \begin{bmatrix}1 & 1 & J\\ M & M_S & M_S+M\end{bmatrix}^2 \Delta_{1\mu}(R). \quad (33)$$

Averaging the cross sections in Table 7 over fine states of the ground electron state gives the values 70, 56, and

Table 7. The ion–atom interaction potential for the quasi-molecule $O^+(^4S_{3/2})$–$O(^3P_J)$ with its quantum numbers given by J and M (the total atom momentum and the projection of the atom orbital momentum on the molecular axis) and the partial resonant charge exchange cross sections σ_{ex} for the indicated quantum numbers and collision energies ε in the laboratory reference frame

JM	$\Delta(JM_J)$	σ_{ex}, Å^2 ε = 0.1 eV	σ_{ex}, Å^2 ε = 1 eV	σ_{ex}, Å^2 ε = 10 eV
20	$\frac{20}{9}\Delta_{11}$	69	56	40
21	$\frac{10}{9}\Delta_{10}$	84	68	55
10	$\frac{4}{3}\Delta_{11}$	63	51	45
11	$\frac{2}{3}\Delta_{10}$	77	62	49
00	$\frac{4}{9}\Delta_{11}$	51	41	31
01	$\frac{2}{9}\Delta_{10}$	63	50	38

44 Å^2 for the average cross section of resonant charge exchange at the respective collision energies of 0.1, 1, and 10 eV. For the ground fine state $J = 2$, these values are 79, 64, and 50 Å^2, respectively. In this case, the average cross section differs from that for the lowest fine state. Next, with the logarithmic dependence of the cross section on the collision velocity taken into account, we find that instead of (32), the cross section averaged over fine states is given by

$$\sigma_{res} = \frac{\pi R_0^2}{2}, \tag{34a}$$

where

$$\frac{2.6}{v(R_0\gamma)^{1/3}}\sqrt{\frac{\pi R_0}{2\gamma}}\Delta_{00}(R_0) = 0.28, \tag{34b}$$

which only slightly differs from Eq. (32). We find that the average cross sections are close to those for case "a' of the Hund coupling. However, the partial cross sections can be different in these cases.

7. CONCLUSIONS

Studying the character of momentum coupling for the resonant charge exchange process in slow collisions, we have found that the number of real cases of momentum coupling is considerably larger than that following from the classical Hund scheme of momentum coupling. Constructing the hierarchy of interactions for the quasi-molecule consisting of the colliding ion and atom has allowed us to find a suitable scheme for momentum coupling. The strongest interaction for nonheavy atoms is the exchange interaction of electrons inside these atomic particles, and therefore, the quasi-molecule quantum numbers for the ion and atom electron shells are *LSls* (the orbital momentum and spin of the atom and the same quantum numbers for the ion). For halogen and oxygen, the rotation energy V_{rot} of colliding particles is small compared to the fine splitting of the atom (δ_a) and ion (δ_i) levels; it is also small compared to the long-range charge-quadruple interaction U between the ion and the atom. Hence, the resonant charge exchange proceeds at certain quantum numbers JM_Jj or JMj depending on the ratio δ_a/U (where J and j are the total atom and total ion momenta and M and M_J are the projections of the atom orbital and total atom momenta on the quasi-molecule axis). This character of momentum coupling does not correspond to cases in the Hund coupling scheme.

In contrast to the case "a" of the Hund coupling, where the electron transfer process is entangled with the rotation of the atom and ion orbital momenta, the resonant charge exchange process for halogen and oxygen proceeds separately for each fine-structure state. This increases the accuracy of evaluation of the electron transfer cross section. The resonant charge exchange cross section depends on the initial quantum numbers of the quasi-molecule; according to the analysis for halogens and oxygen, the cross sections in case "a" of the Hund coupling scheme and for the real momentum coupling are close to each other.

REFERENCES

1. O. B. Firsov, Zh. Éksp. Teor. Fiz. **21**, 1001 (1951).
2. R. S. Mulliken, Rev. Mod. Phys. **2**, 60 (1930).
3. F. Hund, Z. Phys. **36**, 637 (1936).
4. L. D. Landau and E. M. Lifshitz, *Course of Theoretical Physics*, Vol. 3: *Quantum Mechanics: Non-Relativistic Theory*, 4th ed. (Nauka, Moscow, 1989; Pergamon Press, London, 1980).
5. E. E. Nikitin, Opt. Spectrosk. **22**, 379 (1966).
6. E. E. Nikitin and B. M. Smirnov, Usp. Fiz. Nauk **124**, 201 (1978) [Sov. Phys. Usp. **21**, 95 (1978)].
7. E. E. Nikitin and S. Ya. Umanskiĭ, *Theory of Slow Atomic Collisions* (Atomizdat, Moscow, 1979; Springer, Berlin, 1984).
8. E. E. Nikitin and B. M. Smirnov, *Atomic and Molecular Processes* (Nauka, Moscow, 1988).
9. B. M. Smirnov, *Physics of Atoms and Ions* (Springer, New York, 2003).
10. B. M. Smirnov, Zh. Éksp. Teor. Fiz. **46**, 1017 (1964) [Sov. Phys. JETP **19**, 692 (1964)].

11. B. M. Smirnov, Zh. Éksp. Teor. Fiz. **47**, 518 (1964) [Sov. Phys. JETP **20**, 345 (1965)].
12. B. M. Smirnov, *Asymptotic Methods in Theory of Atomic Collisions* (Atomizdat, Moscow, 1972).
13. B. M. Smirnov, Phys. Scr. **61**, 595 (2000).
14. B. M. Smirnov, Usp. Fiz. Nauk **171**, 233 (2001) [Phys. Usp. **44**, 221 (2001)].
15. E. U. Condon and G. H. Shortley, *Theory of Atomic Spectra*, 2nd ed. (Cambridge Univ. Press, Cambridge, 1949; Inostrannaya Literatura, Moscow, 1949).
16. I. I. Sobel'man, *Atomic Spectra and Radiative Transitions* (Nauka, Moscow, 1977; Springer, Berlin, 1979).
17. S. Bashkin and J. Stoner, *Atomic Energy Levels and Grotrian Diagrams* (North-Holland, Amsterdam, 1978–1982), Vols. 1–4.
18. A. A. Radzig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer, Berlin, 1985).
19. B. M. Smirnov, Zh. Éksp. Teor. Fiz. **119**, 1099 (2001) [JETP **92**, 951 (2001)].
20. G. Racah, Phys. Rev. **61**, 186 (1942); Phys. Rev. **62**, 438 (1942).
21. E. L. Duman and B. M. Smirnov, Zh. Tekh. Fiz. **40**, 91 (1970) [Sov. Phys. Tech. Phys. **15**, 61 (1970)].
22. B. M. Smirnov, Teplofiz. Vys. Temp. **4**, 429 (1966).

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 505–521.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 558–577.

ATOMS, SPECTRA, RADIATION

Limiting Efficiencies of Nonlinear-Optical Processes in Microstructure Fibers

A. M. Zheltikov

Physics Department, International Laser Center, Moscow State University, Vorob'evy gory, Moscow, 119899 Russia

e-mail: zheltikov@top.phys.msu.su

Received February 19, 2003

Abstract—The ways of achieving limiting waveguide enhancement of nonlinear-optical processes in microstructure and photonic-crystal fibers are studied. The waveguide enhancement of nonlinear-optical processes is shown to be physically limited because of the competition of diffraction and refractive-index-step radiation confinement. In the case of the limiting refractive-index step values for fused silica fibers, the maximum waveguide enhancement of nonlinear-optical processes is achieved with submicron fiber core diameters. The maximum waveguide enhancement of coherent anti-Stokes Raman scattering in a hollow microstructure fiber relative to the regime of tight focusing is shown to scale as $\lambda^2/\alpha^2 a^4$ with radiation wavelength λ, the inner fiber radius a, and the magnitude of radiation losses α. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The advent of microstructure and photonic-crystal fibers [1–10] has opened a new phase in the nonlinear optics of guided waves. Such fibers provide a strong confinement of electromagnetic radiation in the fiber core [11, 12] and offer many degrees of freedom in dispersion tailoring via variations in the core–cladding geometry of the fiber [13, 14]. Due to their unique properties, microstructure and photonic-crystal fibers enhance the whole catalog of nonlinear-optical processes, making nonlinear optics accessible to unamplified femtosecond laser pulses and suggesting new solutions for the frequency conversion and spectral transformation of ultrashort laser pulses, as well as for the phase and temporal control of such pulses [4, 15]. The enhancement of a broad class of nonlinear-optical phenomena, including self-phase modulation [12], four-wave mixing and parametric processes [16–21], stimulated Raman scattering [17, 18], and soliton formation [22], leads to efficient generation of radiation with a very broad continuous spectrum—a supercontinuum [23, 24]. Supercontinuum generation in microstructure fibers is the backbone of femtosecond systems for high-precision measurements [25–29], changing the paradigm of optical frequency metrology. Supercontinuum generation also holds much promise for the measurement and control of the phase of ultrashort pulses [30, 31] and offers attractive solutions for ultrafast photonics [32], spectroscopy [33], and biomedical optics [34]. Supercontinuum generation is a result of a complicated interplay between many nonlinear-optical processes, in fact, serving as an anthology of basic nonlinear-optical phenomena. Physical scenarios of supercontinuum generation in microstructure fibers is an exciting subject for basic research in the interdisciplinary area of ultrafast nonlinear optics and the physics of micro- and nanostructures.

Fibers with a cladding in the form of a two-dimensionally periodic microstructure (two-dimensional photonic crystal) and a hollow core, first demonstrated by Cregan *et al.* [35], is one of the most interesting and promising types of microstructure fibers. Photonic band gaps in the transmission spectrum of a two-dimensional periodic cladding in these fibers provide high reflection coefficients for electromagnetic radiation propagating along the hollow core of the fiber, allowing a specific regime of waveguiding to be implemented [35–37]. This mechanism of waveguiding is of special interest for telecommunication applications, opening, at the same time, ways of enhancing nonlinear-optical processes. Benabid *et al.* [38] recently demonstrated a radical enhancement of stimulated Raman scattering in hollow-core photonic-crystal fibers. Such fibers can be also employed for laser manipulation of small-size particles [39]. The structure of hollow-core photonic-crystal fibers is ideally suited for the transportation of high-power laser radiation [40], high-order harmonic generation [41], transmission of ultrashort laser pulses [35, 37], and fiber-optic delivery of laser radiation in technological laser systems [42].

In view of the exciting possibilities offered by microstructure and hollow-core photonic-crystal fibers in nonlinear optics, the physical factors determining the limiting efficiencies of nonlinear-optical processes in such fibers are of special interest. Understanding of these factors would allow the strategy for optimizing

the structure and parameters of microstructure and hollow-core photonic-crystal fibers to be developed for the maximum enhancement of nonlinear-optical processes. The analysis of these factors is the main task of this paper. We will consider the processes of stimulated Raman scattering (SRS) and coherent anti-Stokes Raman scattering (CARS) in hollow-core photonic-crystal fibers, examine the factors determining the limiting SRS and CARS efficiencies in such fibers, and find the optimal fiber lengths for these processes taking into consideration waveguide losses, as well as group-delay and phase-mismatch effects. We will show that microstructure fibers open ways toward limiting waveguide enhancement of nonlinear-optical processes. Physically, the existence of this limit for waveguide enhancement is associated with competition between diffraction and radiation confinement in a waveguide due to the refractive-index step. We will derive asymptotic expressions providing in several important cases an adequate qualitative description of the waveguide enhancement of nonlinear-optical processes and making it possible to estimate with satisfactory accuracy the optimal fiber core diameters for the maximum enhancement of nonlinear-optical interactions.

2. THE PHYSICS BEHIND THE WAVEGUIDE ENHANCEMENT OF NONLINEAR-OPTICAL PROCESSES

Optical fibers represent a natural way to increase the interaction length and to reduce the beam diameter of light fields involved in nonlinear processes. These two factors underlie a successful and rapidly growing area of nonlinear fiber optics [43], substantially improving the efficiency of many nonlinear-optical processes (including stimulated Raman scattering, as well as self- and cross-phase modulation) and lowering thresholds for a certain class of nonlinear-optical effects (e.g., processes involving stimulated Raman scattering). The product Il_{eff}, where I is the light field intensity in an optical fiber and l_{eff} is the effective interaction length, gives a figure of merit [43] for the efficiency of a broad class of nonlinear-optical interactions, including self- and cross-phase modulation, stimulated Raman scattering, and certain types of parametric wave-mixing processes. The interaction length for tightly focused light beams is confined to the beam waist length. In the case of a Gaussian beam, the beam waist length is approximately equal to

$$l_{\text{eff}}^{f} \approx \pi w_0^2/\lambda,$$

where w_0 is the waist radius of the focused beam and λ is the radiation wavelength. The waveguide enhancement of a nonlinear-optical process in a fiber with the magnitude of losses α ($\alpha l \gg 1$, l is the fiber length) with respect to the regime of tightly focused pump beams is then given by the well-known expression [43]

$$\frac{I_f l_{\text{eff}}^{f}}{I_t l_{\text{eff}}^{t}} \approx \frac{\lambda}{\pi w_0^2 \alpha}, \tag{1}$$

where I_f and I_t are the radiation intensities calculated for the fiber and the regime of tight focusing, respectively.

It is straightforward from Eq. (1) that the waveguide enhancement of nonlinear-optical processes is associated with the possibility of achieving high radiation intensities for a given radiation power due to the light-field confinement in a fiber core with a small radius and large interaction lengths attainable for nonlinear-optical processes in fibers with low radiation losses ($l_{\text{eff}}^{f} \approx 1/\alpha$). The waveguide enhancement factor and the figure of merit quantifying the lowering of the threshold for nonlinear-optical processes involving stimulated Raman scattering, as can readily be seen from Eq. (1), grows with a decrease in the fiber core diameter.

This recipe for enhancing nonlinear-optical processes remains applicable, however, within a finite range of fiber core diameters. The physical limit is, of course, determined by diffraction effects. As the fiber core radius becomes smaller and smaller, less and less radiation power remains confined to the fiber core. To quantify the influence of this factor on the waveguide enhancement of nonlinear-optical processes, we represent the product of the intensity of electromagnetic radiation and the effective interaction length in an optical fiber in the following form:

$$I_f l_{\text{eff}}^{f} = \int_0^l \frac{P\eta}{\pi a^2} \exp(-\alpha z)\, dz, \tag{2}$$

where η is the ratio of the laser power confined to the fiber core to the total laser power guided by a fiber mode and a is the radius of the fiber core.

Calculating the integral in Eq. (2) and assuming that $\alpha l \gg 1$, as before, we arrive at

$$\xi = \frac{I_f l_{\text{eff}}^{f}}{I_t l_{\text{eff}}^{t}} \approx \frac{\lambda\eta}{\pi a^2 \alpha}. \tag{3}$$

Formula (3) allows the distribution of radiation power between the fiber core and the fiber cladding to be included in the waveguide enhancement factor for nonlinear-optical processes in fiber-guided modes. In what follows, we will employ Eq. (3) to determine the optimal core radius for nonlinear fibers providing the maximum enhancement of nonlinear optical processes and explore the applicability of different asymptotic

expressions for estimating the optimal parameters of nonlinear fibers.

3. THE ROLE OF DIFFRACTION AND THE PHYSICAL LIMIT FOR THE WAVEGUIDE ENHANCEMENT OF NONLINEAR-OPTICAL PROCESSES

To calculate the radiation power confined to a fiber core in the HE_{1m} waveguide mode, we use the following expression known from the theory of optical fibers [44]:

$$\eta = \frac{\Xi_{co}}{\Xi_{co} + \Xi_{cl}}, \tag{4}$$

where

$$\Xi_{co} = \frac{kn_{co}^2}{\beta J_1^2(u)}\{a_1a_3[J_0^2(u) + J_1^2(u)] + a_2a_4[J_2^2(u) + J_1(u)J_3(u)]\}, \tag{5}$$

$$\Xi_{cl} = \frac{kn_{co}^2}{\beta K_1^2(W)}\frac{u^2}{W^2}\{a_1a_5[K_0^2(W) - K_1^2(W)] + a_2a_6[K_2^2(W) + K_1(W)K_3(W)]\}, \tag{6}$$

$$a_1 = \frac{F_1 - 1}{2}, \quad a_2 = \frac{F_2 + 1}{2},$$

$$a_3 = \frac{F_1 - 1}{2}, \quad a_4 = \frac{F_1 + 1}{2},$$

$$a_5 = \frac{F_1 - 1 + 2\Delta}{2}, \quad a_6 = \frac{F_1 + 1 - 2\Delta}{2},$$

$$F_1 = (uW)^2\frac{b_1 + (1 - 2\Delta)b_2}{V^2}, \quad F_2 = \frac{V^2}{(uW)^2(b_1 + b_2)},$$

$$b_1 = \frac{1}{2u}\left[\frac{J_0(u)}{J_1(u)} - \frac{J_2(u)}{J_1(u)}\right],$$

$$b_2 = -\frac{1}{2W}\left[\frac{K_0(W)}{K_1(W)} + \frac{K_2(W)}{K_1(W)}\right], \quad \Delta = \frac{n_{co}^2 - n_{cl}^2}{2n_{co}^2},$$

n_{co} is the refractive index of the fiber core, n_{cl} is the refractive index of the fiber cladding, u is the eigenvalue of the characteristic equation for the waveguide mode (also known as the mode parameter in the fiber core), $V = ka(n_{co}^2 - n_{cl}^2)^{1/2}$ is the waveguide parameter, $k = 2\pi/\lambda$, $\beta = [k^2n_{co}^2 - (u/a)^2]^{1/2}$ is the propagation constant of the waveguide mode, $W = a(\beta^2 - k^2n_{cl}^2)^{1/2}$ is the mode parameter in the fiber cladding, $J_n(x)$ are the first-kind Bessel functions, and $K_n(x)$ are the modified second-kind Bessel functions.

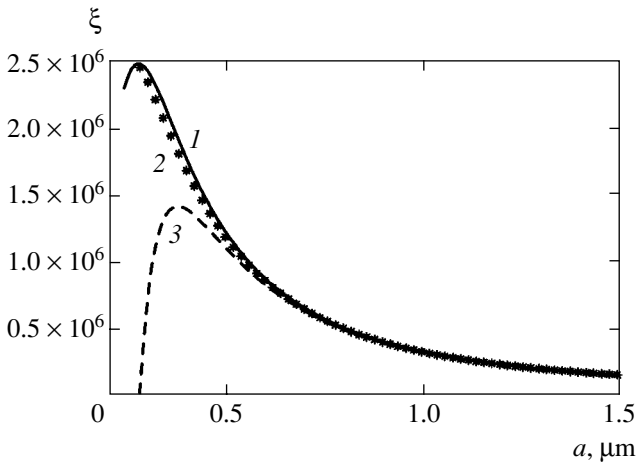

Fig. 1. The factor ξ of the waveguide enhancement of nonlinear-optical processes as a function of the fiber core radius a for a fiber with a refractive index of the core of $n_{co} \approx 1.45$ and a refractive index of the cladding of $n_{cl} \approx 1$ ($\Delta \approx 0.26$) and radiation with a wavelength of 1 μm calculated (*1*) with the use of Eqs. (3)–(6), (*2*) with the use of approximate formula (11) with the mode parameter u determined by solving the relevant characteristic equation, and (*3*) with the use of approximate formula (11) under the assumption that $u \approx 2.4$.

The solid line *1* in Fig. 1 shows the dependence of the factor ξ on a calculated with the use of Eqs. (3)–(6). The maximum values of ξ are achieved with a certain optimal core radius a_{opt}. For the factor ξ represented by curve *1* in Fig. 1, $a_{opt} = 0.27$ μm.

To understand the physical factors controlling the behavior of the waveguide enhancement of nonlinear-optical processes as a function of the fiber core radius, we will examine in greater detail the limiting cases of $V \gg 1$ and $V \ll 1$, corresponding to fibers with small and large ratios of the core radius to the radiation wavelength. Physically, these two situations differ by the relation between diffraction and index-step waveguide confinement effects. To illustrate this argument, we introduce the characteristic angular beam width θ_d, which will serve as a measure of diffraction and which is given, in the case of a Gaussian profile of intensity distribution with a width of w_0, by the well-known formula [44]

$$\theta_d = \arctan\left(\frac{\lambda}{\pi n_{co} w_0}\right) \approx \frac{\lambda}{\pi n_{co} a}. \tag{7}$$

Using Eq. (7), we represent the waveguide parameter V as [44]

$$V = \frac{2\theta_c}{\theta_d}, \tag{8}$$

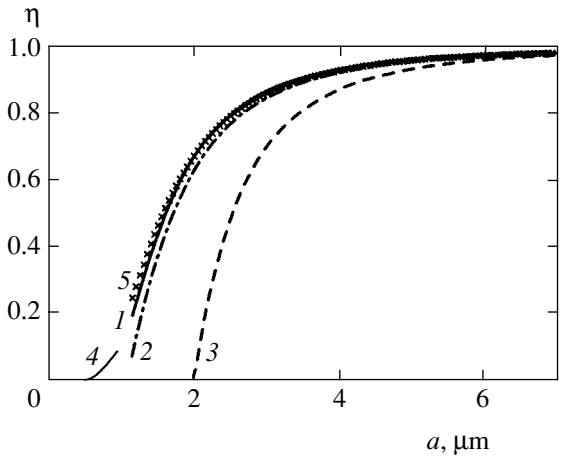

Fig. 2. The fraction η of electromagnetic radiation power confined in the core of a weakly guiding ($\Delta \approx 5 \times 10^{-3}$) fiber as a function of the fiber core radius a for radiation with a wavelength of 1 μm calculated (*1*) with the use of Eqs. (4)–(6), (*2*) with the use of approximate formula (10) with the mode parameter u determined by solving the relevant characteristic equation, (*3*) with the use of approximate formula (10) under the assumption that $u \approx 2.4$, (*4*) with the use of approximate formula (9) in the regime when $V \ll 1$, and (*5*, asterisks) with the use of approximate formulas (10) and (15).

where θ_c is the critical angle of grazing incidence,

$$\theta_c = \arcsin[(1 - n_{cl}^2/n_{co}^2)]^{1/2} \approx (1 - n_{cl}^2/n_{co}^2)^{1/2}.$$

Thus, the waveguide parameter V can be considered as a measure of balance between diffraction and waveguide confinement of a light beam. For fibers with large core radii having $V \gg 1$, diffraction is negligible and most of the laser power is confined to the fiber core (see Fig. 2). In the opposite case of fibers with small core radii having $V \ll 1$, diffraction becomes significant and a considerable fraction of radiation power is guided in the fiber cladding (Fig. 2).

For the fundamental mode of a weakly guiding optical fiber, $\Delta \ll 1$, the fraction η of electromagnetic radiation power confined to the fiber core in the limiting case of $V \ll 1$ can be calculated with the use of the following asymptotic expression [44]:

$$\eta \approx 1.261 \frac{V^2 + 2}{V^4} \exp\left(-\frac{4}{V^2}\right). \quad (9)$$

Because of diffraction effects, the factor η, as can be seen from Eq. (9), rapidly tends to zero as $V \longrightarrow 0$ (curve *5* in Fig. 2).

For large V, the ratio η is given by the following asymptotic formula [44]:

$$\eta \approx 1 - \frac{u^2}{V^3}. \quad (10)$$

For $V \geq 3$, Eq. (10) provides satisfactory agreement with the results of calculations performed for the ratio η with the use of Eqs. (4)–(6) (cf. curves *1* and *3* in Fig. 2). As the fiber core radius increases, the ratio η, in accordance with Eq. (10), tends to unity, since the role of diffraction becomes negligible as compared with the waveguide confinement of the laser beam due to the refractive-index step.

4. ASYMPTOTIC EXPRESSIONS FOR THE FACTOR OF THE WAVEGUIDE ENHANCEMENT OF NONLINEAR-OPTICAL PROCESSES AND NUMERICAL SIMULATIONS

We now use the approximation of Eq. (10) to estimate the waveguide enhancement of nonlinear-optical processes. Substituting Eq. (10) into Eq. (3), we find

$$\xi \approx \frac{\lambda}{\pi a^2 \alpha}\left(1 - \frac{u^2}{V^3}\right). \quad (11)$$

The applicability of the estimate given by Eq. (11) requires additional analysis, since this formula was derived with the use of asymptotic expressions valid for the regime of $V \gg 1$. Such an approach is intrinsically contradictory as the existence of the physical limit for the factor ξ and the optimal value of the fiber core radius is related to diffraction effects, which become negligibly small when $V \gg 1$ [see Eq. (8)]. The waveguide enhancement factor for nonlinear-optical processes calculated with the use of Eq. (11) is shown by curve *2* (asterisks) in Fig. 1. Comparison of the results of these calculations using approximate relations with the predictions of Eqs. (4)–(6) (curve *1* in Fig. 1) demonstrates that Eq. (11) provides a rather accurate description of the factor ξ for large fiber core radii, i.e., in the regime of weak diffraction. The estimate for the optimal fiber core radius obtained with the use of Eq. (11) ($a_{opt} \approx 0.24$ μm, curve *2* in Fig. 1) also agrees well with the optimal value of the fiber core radius predicted by Eqs. (4)–(6) ($a_{opt} \approx 0.27$ μm, curve *1* in Fig. 1).

The use of Eqs. (10) and (11) implies the solution of the characteristic equation for the mode parameter u for the relevant waveguide mode. Let us examine now the accuracy of a simplified estimation procedure that uses Eq. (11) with limiting values of the waveguide mode parameter corresponding to large V. Setting $u \approx 2.4$ for the fundamental guided mode, differentiating Eq. (11) in a, equating the resulting expression to zero, and solving the equation thus obtained, we derive the following simple formula for the optimal value of the fiber core radius, providing the maximum waveguide enhancement of nonlinear-optical processes:

$$a_{opt} = \left(\frac{5}{2}\kappa\right)^{1/3}\lambda, \quad (12)$$

where

$$\kappa = \frac{u^2}{(2\pi)^3 (n_{co}^2 - n_{cl}^2)^{3/2}}.$$

Setting $n_{co} \approx 1.45$ and $n_{cl} \approx 1$, we arrive at the following estimate for the optimal fiber core radius: $a_{\text{opt}} \approx 0.37\lambda$. In view of the approximation $u \approx 2.4$, the accuracy of Eq. (12) is lower than the accuracy of Eqs. (10) and (11). Curve *3* in Fig. 1 shows the dependence of the factor ξ on the fiber core radius calculated with the use of Eq. (11) in the approximation of $u \approx 2.4$ for radiation with a wavelength of 1 μm and the above-specified refractive indices of the core and the cladding. Expression (12) can provide only order-of-magnitude estimates on the optimal fiber core radius and the maximum enhancement factor for nonlinear-optical processes (curve *3* in Fig. 1). This approximation systematically overestimates factor ξ, since it employs Eq. (11) where the *V*-dependent mode parameter *u* is replaced by its upper-bound value. The maximum value of the factor ξ in the dependence represented by curve *3* in Fig. 1 is achieved with the core radius meeting Eq. (12), $a_{\text{opt}} \approx 0.37$ μm. Being correct in its order of magnitude, this estimate still noticeably differs in its value from the result obtained with the use of Eqs. (3)–(6).

The main advantages of the estimate given by Eq. (12) are associated with its simplicity and the insights it gives into the influence of fiber parameters on the optimal core radius and the maximum value of the factor ξ. The waveguide enhancement factor in this approximation is given by

$$\xi = \frac{4\pi}{\alpha\lambda}(n_{co}^2 - n_{cl}^2)\psi(V), \qquad (13)$$

where

$$\psi(V) = \frac{1}{V^2}\left(1 - \frac{u^2}{V^3}\right). \qquad (14)$$

When the mode parameter *u* is estimated by its upper-bound value ($u \approx 2.4$ for the fundamental mode), the function $\psi(V)$ depends only on the type of the waveguide mode and reaches its maximum value with $V_{\text{opt}} = (5/2)^{1/3}u^{2/3}$. Expression (13) shows, in particular, that a decrease in the optimal core radius and, consequently, the increase [in view of Eq. (3)] in the maximum waveguide enhancement of nonlinear-optical processes can be achieved by increasing the refractive-index step (parameter Δ) between the core and the cladding of the fiber.

Figure 3 displays the dependences of the factor ξ on the fiber core radius for weakly guiding fibers with $\Delta \approx 5 \times 10^{-3}$. The maximum value of the factor ξ for 1-μm radiation is achieved in this case with a core radius equal to 1.5 μm (curve *1* in Fig. 3). The factor ξ under these conditions is nearly 40 times less than in the case of a fiber with a limiting, for fused silica structures, refractive-index step ($\Delta \approx 0.26$, Fig. 1). In the case of weakly guiding fibers, approximate formula (11) also provides a sufficiently accurate estimate for the optimal fiber core radius and the maximum value of the factor ξ (curve *2* in Fig. 3). Finally, Eq. (11), where the mode parameter *u* is replaced by its upper-bound value ($u \approx 2.4$), can provide only order-of-magnitude estimates for these parameters (curve *3* in Fig. 3). This latter approximation, however, gives a qualitatively correct prediction for the decrease in the maximum value of the factor ξ with a decreasing refractive-index step. The difference of the refractive indices squared, $n_{co}^2 - n_{cl}^2$, appearing in Eq. (13) is approximately 50 times higher in the case of the dependences presented in Fig. 1 than the parameter $n_{co}^2 - n_{cl}^2$ for the dependences of Fig. 3. The corresponding ratio of the values of the factor ξ, calculated with the use of the exact formulas (3)–(6), is approximately 40 (cf. curves *1* in Figs. 1 and 3). In practice, the highest values of the parameter Δ can be achieved with tapered fibers [45, 46] and microstructure fibers with high air-filling fractions [12, 20, 47–49]. Fibers of these types, therefore, open ways for attacking the physical limit of the waveguide enhancement of nonlinear-optical interactions.

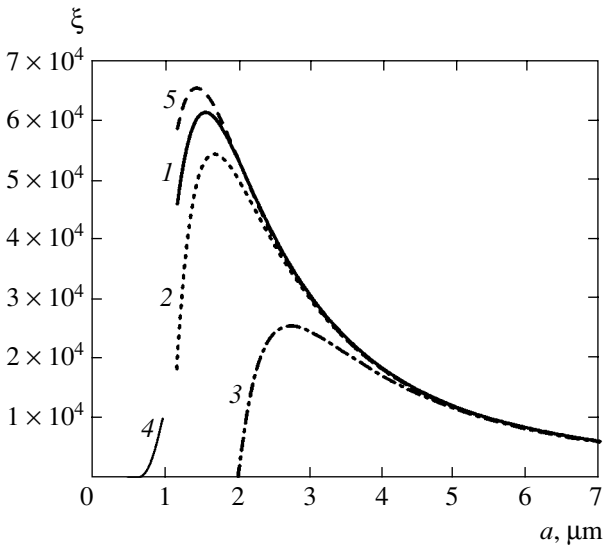

Fig. 3. The factor ξ of the waveguide enhancement of nonlinear-optical processes as a function of the fiber core radius *a* for radiation with a wavelength of 1 μm in a weakly guiding ($\Delta \approx 5 \times 10^{-3}$) fiber calculated (*1*) with the use of Eqs. (3)–(6), (*2*) with the use of approximate formula (11) with the mode parameter *u* determined by solving the relevant characteristic equation, (*3*) with the use of approximate formula (11) under the assumption that $u \approx 2.4$, (*4*) with the use of approximate formula (9) in the regime when $V \ll 1$, and (*5*) with the use of approximate formulas (11) and (15).

The analysis performed above shows that the approximation where the mode parameter in Eqs. (11), (13), and (14) is replaced by its upper-bound value does not always provide a high accuracy of estimates for the factors η and ξ. The accuracy of Eqs. (13) and (14) can sometimes be improved by using the following asymptotic representation for the parameter u of the fundamental waveguide mode [44]:

$$u \approx 2.405 \exp\left(-\frac{1+\Delta}{V}\right). \quad (15)$$

Formulas (13) and (14), in combination with the asymptotic representation of Eq. (15), provide satisfactory accuracy in calculations of the power guided in the fiber core (curve *5* in Fig. 2) and the waveguide enhancement factor ξ (curve *5* in Fig. 3), allowing the optimal values of the fiber core radius and the maximum values of the factor ξ to be calculated with reasonable accuracy. The error in the estimation of the factor Δ with the use of Eqs. (10) and (15) in the intermediate range of V corresponding to the maximum values of the waveguide enhancement increases with the growth in the parameter Δ. The accuracy of estimates on the optimal value of the fiber core radius and the maximum value of the factor ξ decreases under these conditions.

5. STIMULATED RAMAN SCATTERING AND COHERENT ANTI-STOKES RAMAN SCATTERING IN GUIDED MODES OF HOLLOW PHOTONIC-CRYSTAL FIBERS

In this section, we will discuss the generic idea of enhancing SRS- and CARS-type nonlinear-optical processes with the use of hollow-core microstructure and photonic-crystal fibers. Hollow fibers are currently widely employed [50] for the generation of ultrashort pulses using Kerr-nonlinearity-related self- and cross-phase modulation [51–53] and stimulated Raman scattering [54], as well as for high-order harmonic generation [55–58] and improving the sensitivity of gas-phase analysis based on four-wave mixing spectroscopy [59–61].

In hollow fibers, the refractive index of the core, n_{co}, is lower than the refractive index of the cladding, $n_{cl} = \sqrt{\varepsilon_{cl}}$. Therefore, the propagation constants of guided modes in hollow fibers,

$$\beta = \sqrt{k_{co}^2 - (u_n/a)^2} = \sqrt{k_{cl}^2 - (W_n/a)^2}$$

(where $k_{co} = n_{co}\omega/c$, $k_{cl} = n_{cl}\omega/c$, u_n is the eigenvalue of the characteristic equation for a waveguide mode with a system of mode indices n, and W_n is the transverse wave number of the fiber cladding mode) have nonzero imaginary parts. The propagation of light in such fibers is accompanied by radiation losses. For EH_{mn} modes of a hollow fiber with an inner radius a and a refractive index of the gas filling the fiber core of $n_{co} \approx 1$, the radiation intensity attenuation coefficient is given by [62, 63]

$$\alpha = \left(\frac{u_{mn}}{2\pi}\right)^2 \frac{\lambda^2}{a^3} \frac{n^2+1}{\sqrt{n^2-1}}, \quad (16)$$

where $n = n_{cl}$.

Such a behavior of the magnitude of optical losses prevents one from using hollow fibers with very small inner diameters in nonlinear-optical experiments. Using Eq. (16), we find that the magnitude of radiation losses for the fundamental mode of a hollow fiber with a fused silica cladding and an inner radius of 7 μm may exceed 6.5 cm^{-1} for 1-μm radiation, which, of course, imposes serious limitations on the applications of such fibers.

To qualitatively illustrate the idea of lowering the magnitude of optical losses in a hollow fiber with a periodic microstructure cladding relative to the magnitude of optical losses in a hollow waveguide with a solid cladding, we will employ the result well known from the analysis of radiation propagation in a planar waveguide with a periodic cladding [64]. The decrease in the magnitude of optical losses in a hollow planar waveguide with a periodic cladding relative to the magnitude of optical losses in a hollow planar waveguide with a solid cladding can be quantified by determining the ratio of the logarithm of the coefficient of reflection from a periodic structure to the logarithm of the coefficient of reflection from the wall of a hollow waveguide. Around the center of the photonic band gap in the reflection spectrum of the periodic structure in the waveguide cladding with a sufficiently large number of layers N, the coefficient of optical losses in a hollow planar waveguide with a periodic cladding α_{PBG} decreases exponentially, as shown in [64], relative to the coefficient of losses in a hollow waveguide with a solid cladding α_h with the increase in the number of modulation periods of the refractive index in the waveguide cladding:

$$\frac{\alpha_{PBG}}{\alpha_h} \propto a \exp(-2|\kappa|Nd),$$

where κ is the coupling coefficient of the forward and backward waves in the periodic structure of the waveguide cladding and d is the modulation period of the refractive index in the waveguide cladding. Thus, hollow waveguides with a periodic cladding allow a considerable reduction in optical losses characteristic of hollow-waveguide modes.

Hollow-core photonic-crystal fibers seem ideally suited for highly efficient waveguide SRS and CARS interactions. Transmission spectra of hollow photonic-crystal fibers display isolated peaks [35, 37], which can be employed to radically enhance nonlinear-optical interactions of spectrally isolated optical signals. SRS

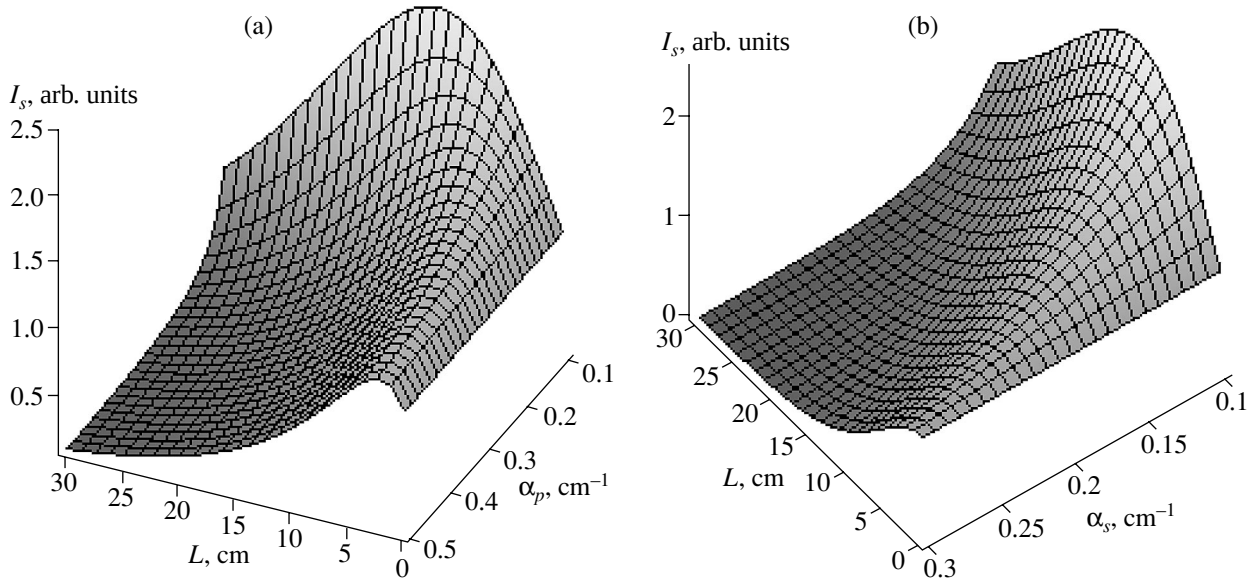

Fig. 4. The intensity of the SRS signal as a function of the fiber length L and the attenuation coefficients of (a) the pump and (b) Stokes signals for $gI_0 = 0.3$ cm^{-1}, (a) $\alpha_s = 0.1$ cm^{-1}, and (b) $\alpha_p = 0.1$ cm^{-1}.

and CARS belong to this class of nonlinear-optical processes. In the following sections, we will consider ways to optimize parameters of hollow photonic-crystal fibers for maximum enhancement of SRS and CARS processes.

6. STIMULATED RAMAN SCATTERING

6.1. The Influence of Waveguide Losses and the Optimal Fiber Length

To assess the influence of waveguide losses on stimulated Raman scattering in a hollow-core fiber with a solid or photonic-crystal cladding having a length l, we employ the solution to the equation for the intensity of the Stokes signal of the stationary SRS neglecting pump-depletion effects [43]:

$$I_s(l) = I_s(0)\exp(gI_0 l_{\text{eff}} - \alpha_s l), \qquad (17)$$

where g is the SRS gain; I_0 is the initial intensity of the pump signal; α_p and α_s are the coefficients of losses at the pump and Stokes frequencies, respectively; and

$$l_{\text{eff}} = \frac{1}{\alpha_p}[1 - \exp(-\alpha_p l)] \qquad (18)$$

is the effective interaction length.

The intensity of the SRS signal, as can be seen from Eqs. (17) and (18), is determined by the SRS gain, the intensity of pump radiation, and the coefficients of losses at the frequencies of pump and Stokes radiation. Hollow-core microstructure fibers with a core diameter on the order of 13–15 μm [35, 37] allow pump radiation intensities on the order of 5×10^{10}–5×10^{12} W/cm^2 to be achieved with 100-fs pulses having an energy of 0.01–1 μJ. Figure 4 displays the Stokes signal intensity as a function of the fiber length and the coefficients of losses at the frequencies of pump (Fig. 4a) and Stokes (Fig. 4b) radiation calculated for the magnitudes of losses characteristic of photonic-crystal fibers with a core diameter of 13–15 μm under the assumption that $gI_0 = 0.3$. Such values of the gI_0 factor can be achieved in the case of hollow-core fibers filled with molecular hydrogen ($b^2 = [\text{Sp}(\partial\alpha_{ij}/\partial Q)/3]^2 \approx 4 \times 10^{-34}$ cm^4, where $\partial\alpha_{ij}/\partial Q$ is the derivative of the electronic polarizability of molecules in the generalized coordinate Q defining the nuclear configuration) or nitrogen ($b^2 \approx 1.6 \times 10^{-33}$ cm^4) at atmospheric pressure with the above-specified levels of laser radiation intensity.

As can be seen from Eqs. (17) and (18), the Stokes signal grows only with $gI_0 > \alpha_p$. Otherwise, waveguide losses result in an exponential decay of the Stokes signal. For small l, such that $\alpha_p l$, $\alpha_s l$, $gI_0 l \ll 1$, the Stokes signal grows linearly as a function of l (Figs. 4, 5). For lengths l substantially exceeding the attenuation length of pump radiation, $\alpha_p l \gg 1$, waveguide losses lead to a noticeable attenuation of pump radiation. The intensity of the Stokes signal exponentially decreases under these conditions upon an increase in l with a characteristic spatial scale defined by the attenuation length of the Stokes signal (Figs. 4, 5). SRS in a fiber with losses can be thus optimized with some fiber length l_{opt}^{SRS}, providing the maximum efficiency for the generation of the Stokes component (see Figs. 4, 5). Using Eqs. (17)

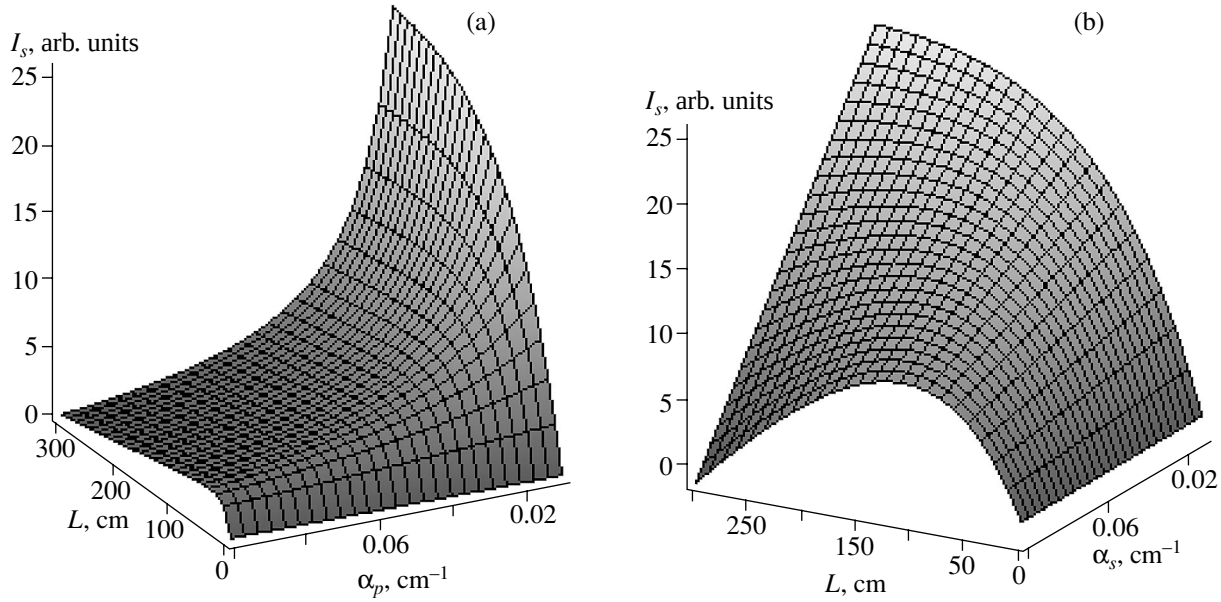

Fig. 5. The intensity of the SRS signal as a function of the fiber length L and the attenuation coefficients of (a) the pump and (b) Stokes signals for $gI_0 = 0.3$ cm^{-1}, (a) $\alpha_s = 0.01$ cm^{-1}, and (b) $\alpha_p = 0.01$ cm^{-1}.

and (18), we derive the following expression for the optimal fiber length l_{opt}^{SRS} (with $gI_0 > \alpha_p$):

$$l_{\text{opt}}^{SRS} = \frac{1}{\alpha_p}\ln\left(\frac{gI_0}{\alpha_s}\right). \tag{19}$$

Substituting this expression for the optimal fiber length into Eq. (17), we find that the maximum integral SRS gain in a hollow fiber is given by

$$G = \frac{gI_0}{\alpha_p} - \frac{\alpha_s}{\alpha_p}\left[1 + \ln\left(\frac{gI_0}{\alpha_s}\right)\right]. \tag{20}$$

In the case of $gI_0 \gg \alpha_s, \alpha_p$, the second term in Eq. (20) is small as compared to the first term. The maximum increase in the integral SRS gain in a hollow fiber relative to the regime of tight focusing is then equal to

$$\zeta = \frac{\lambda}{\pi a^2 \alpha_p}. \tag{21}$$

Comparing Eqs. (3) and (21), we find that the upper bound of the increase in the integral SRS gain predicted by Eqs. (17)–(20) for stationary SRS coincides with the estimate for the waveguide enhancement factor of nonlinear-optical processes given by Eq. (3), which was derived from general physical considerations. As can be seen from Eq. (21), even short, hollow, photonic-crystal and microstructure fibers can provide a substantial SRS enhancement (see Figs. 4, 5). A radical lowering of the SRS threshold in hollow microstructure fibers was observed earlier by Benabid *et al.* [38]. We will show in Section 7 that such fibers may allow an even more substantial enhancement of nonlinear signal generation in the case of coherent anti-Stokes Raman scattering.

6.2. Group-Delay and Group-Velocity-Dispersion Effects

Group-delay effects limit the length of nonlinear-optical interaction, giving rise to a walk-off of the pump and Stokes pulses within the characteristic length

$$l_w = \frac{\tau}{\left|v_p^{-1} - v_s^{-1}\right|},$$

where v_p and v_s are the group velocities of the pump and Stokes pulses, respectively, and τ is the pump pulse duration. Group-velocity dispersion leads to the spreading of light pulses within the characteristic length $l_d = \tau^2/|\beta_2|$. In view of these factors, the effective length of pump–Stokes pulse interaction should be redefined as

$$L_{\text{eff}} = \min(l_{\text{eff}}, l_w, l_d).$$

The influence of the group-velocity mismatch and group-velocity dispersion can be reduced, with an appropriate choice of fiber and gas parameters, by using the dispersion of waveguide modes [65]. Physically, this opportunity can be understood by comparing the

group velocity of a light pulse propagating in a gas-filled hollow fiber (the solid line in Fig. 6a),

$$v^{pq} = \left(\frac{\partial K^{pq}}{\partial \omega}\right)^{-1}, \tag{22}$$

where K^{pq} is the propagation constant corresponding to the relevant waveguide mode of a hollow fiber with mode indices p and q, with the group velocity of a light pulse in the same gas, but in the absence of a waveguide (the dashed line in Fig. 6a),

$$v = \left(\frac{\partial k}{\partial \omega}\right)^{-1} = \frac{c}{n\left(1 + \frac{\omega}{n}\frac{\partial n}{\partial \omega}\right)}, \tag{23}$$

where $k = n\omega/c$ and n is the refractive index of the gas. The propagation constant of a light pulse in a gas-filled hollow fiber, involved in Eq. (22), is related to the wave number k of this pulse in the same gas by the expression

$$K^{pq} = (k^2 - h_{pq}^2)^{1/2},$$

where the quantity h_{pq} can be found from the characteristic equation for the waveguide mode of a hollow fiber.

In particular, using the well-known formulas for the propagation constants of guided modes in hollow fibers [62], we arrive at the following expression for the group velocity of a light pulse with a transverse field distribution corresponding to the EH_{1m} mode of a hollow fiber:

$$v_m = v_0\left[1 - \frac{1}{2}\left(\frac{u_m c}{a\omega n}\right)^2\right], \tag{24}$$

where

$$v_0 = \frac{c}{n\left[1 + \frac{\omega}{n}\frac{\partial n}{\partial \omega}\Big|_{\omega_l}\right]}$$

is the group velocity of the light pulse in the gas in the absence of a hollow fiber.

The group-velocity mismatch in a gas-filled hollow fiber can be then represented as a sum of two terms:

$$\Delta K_m = \Delta K_0 + \Delta K_m^w, \tag{25}$$

where ΔK_0 and ΔK_m^w are the components of the group-velocity mismatch due to the gas and waveguide dispersion, respectively.

The waveguide component of the group-velocity mismatch, as follows from Eq. (24), is inversely proportional to the square of the inner radius of a hollow fiber, scaling as

$$\Delta K_m^w \propto a^{-2}.$$

Physically, this circumstance implies that larger group-velocity mismatches can be compensated in hollow fibers with smaller inner diameters. The dashed lines in Figs. 6a and 6b present the group index and the group-velocity dispersion of molecular hydrogen at a pressure of 0.5 atm as functions of the wavelength. The dotted lines in the same figures show the wavelength dependences of the group index and the group-velocity dispersion for the EH_{11} mode of a hollow fiber with an inner radius of 68 μm. The resulting wavelength dependences of the group index and the group-velocity dispersion, including the waveguide dispersion component, are shown by the solid lines. With an appropriate choice of hollow-fiber parameters, as can be seen from the dependences presented in Figs. 6a and 6b, the

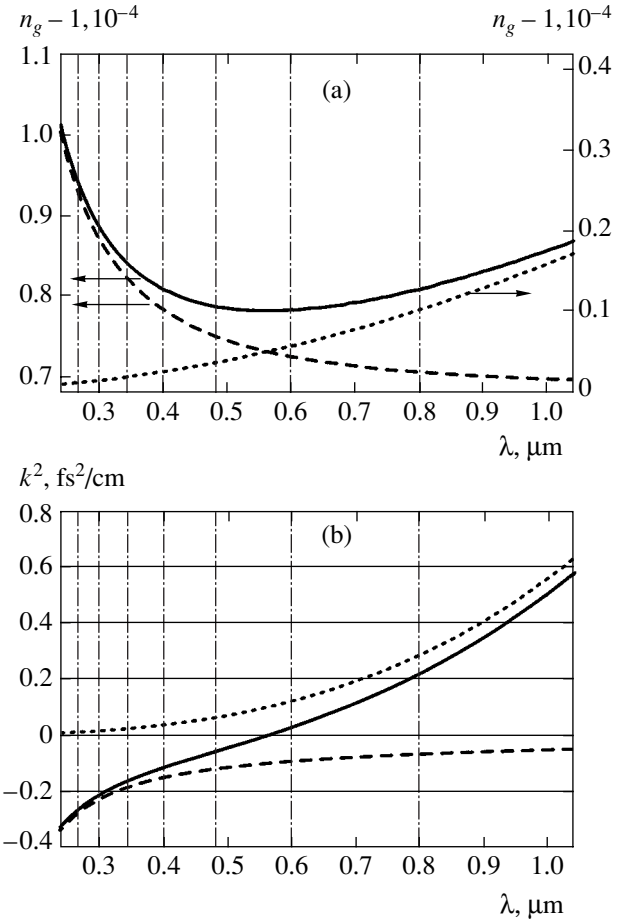

Fig. 6. (a) The group index $n_g = c/v_g$ and (b) group-velocity dispersion calculated as functions of the wavelength (dashed lines) for molecular hydrogen, (dotted lines) the EH_{11} waveguide mode, and (solid lines) the EH_{11} mode of a hollow fiber filled with molecular hydrogen. The gas pressure is 0.5 atm. The inner radius of the fiber is 68 μm.

waveguide dispersion can reduce the group delay and the group-velocity dispersion for the pump and Stokes pulses. At a certain wavelength (560 nm in Fig. 6b), the waveguide component of group-velocity dispersion exactly compensates for the material group-velocity dispersion, giving rise to a point of zero group-velocity dispersion, which is missing from the dispersion profile of the gas (the dashed line in Fig. 6b). The wavelength of zero group-velocity dispersion can be found by doubly differentiating the expression for the propagation constant of guided modes. This procedure yields

$$k_2 = \nu_0^{-2}\left(\frac{\lambda}{2\pi n}\right)^3\left(\frac{u_m}{a}\right)^2, \quad (26)$$

where k_2 is the material component of group-velocity dispersion.

As can be seen from Eq. (26), the wavelength of zero group-velocity dispersion can be tuned by changing the inner radius of the fiber and the type of the waveguide mode, as well as by varying the pressure and the sort of the gas filling the fiber core. The waveguide component of group-velocity dispersion scales as λ^3/a^2 when the radiation wavelength and the inner fiber radius are varied.

7. COHERENT ANTI-STOKES RAMAN SCATTERING

CARS [66–71] is one of the most convenient, efficient, and practical methods of nonlinear coherent spectroscopy. Waveguide regimes [59, 72–75] substantially improve the sensitivity and extend the applicability range of CARS spectroscopy. The enhancement of CARS in gas-filled hollow fibers was experimentally demonstrated back in the 1970s [59]. These experiments have opened a nonlinear-optical chapter in the history of hollow waveguides.

Similar to SRS, CARS is enhanced in hollow fibers due to the confinement of electromagnetic radiation in a small-size fiber core and large interaction lengths. We will show below that the waveguide enhancement in the case of CARS in hollow microstructure fibers with a small inner radius may exceed an analogous enhancement attainable for the SRS process. We will demonstrate that the waveguide enhancement of CARS in a hollow microstructure fiber relative to the regime of tight focusing scales as $\lambda^2/\alpha^2 a^4$ with radiation wavelength λ, radiation losses α, and the fiber inner radius a, allowing substantial enhancement factors to be achieved with hollow-core microstructure fibers having small inner diameters.

We start with the expression for the power of the CARS signal generated via a three-color process $\omega_s = \omega_0 + \omega_1 - \omega_2$ [76]:

$$P_{CARS} = 1.755\times 10^{-5}\frac{\omega_s^4 k_0 k_1 k_2}{c^4 k_s^2 k'} \times D^2\left|\chi_{\text{eff}}^{(3)}\right|^2 P_0 P_1 P_2 F_2. \quad (27)$$

Here, k_0, k_1, k_2, and k_s are the wave numbers of light fields (propagation constants of waveguide modes) with frequencies of ω_0, ω_1, ω_2, and ω_s, respectively; P_0, P_1, and P_2 are the powers of the fields with frequencies of ω_0, ω_1, and ω_2, respectively; $\chi_{\text{eff}}^{(3)}$ is the effective combination of cubic nonlinear-optical susceptibility tensor components corresponding to the chosen set of polarization vectors of pump and signal fields; D is the frequency degeneracy factor of the four-wave mixing process defined after Maker and Terhune [77]; and

$$F_2 = \frac{2k'}{\pi b}\int_0^\infty 2\pi R\,dR \times\left|\int_{-\zeta}^{\xi} d\xi' \frac{\exp\left(\frac{ib\Delta k\xi'}{2}\right)}{(1+i\xi')(k''-ik'\xi')H}\exp\left(-\frac{R^2}{bH}\right)\right|^2 \quad (28)$$

is the phase-matching integral,

$$\Delta k = k_s - (k_0 + k_1 - k_2), \quad k' = k_0 + k_1 - k_2,$$

$$k'' = k_0 + k_1 + k_2,$$

$$\xi = \frac{2(z-f)}{b}, \quad \zeta = \frac{2f}{b},$$

$b = n_j\omega_j w_0^2/c$ is the confocal parameter, w_0 is the beam waist diameter, and

$$H = \frac{(1+\xi')^2}{k'-ik'\xi'} - i\frac{\xi'-\xi}{k'}. \quad (29)$$

In the limiting case of tight focusing, when the confocal parameter b is much less than the length of the nonlinear medium l, $b \ll l$, no increase in the CARS power can be achieved by reducing the pump beam waist radius because of the simultaneous decrease in the interaction length. Mathematically, this well-known result is a consequence of the tight-focusing limit exist-

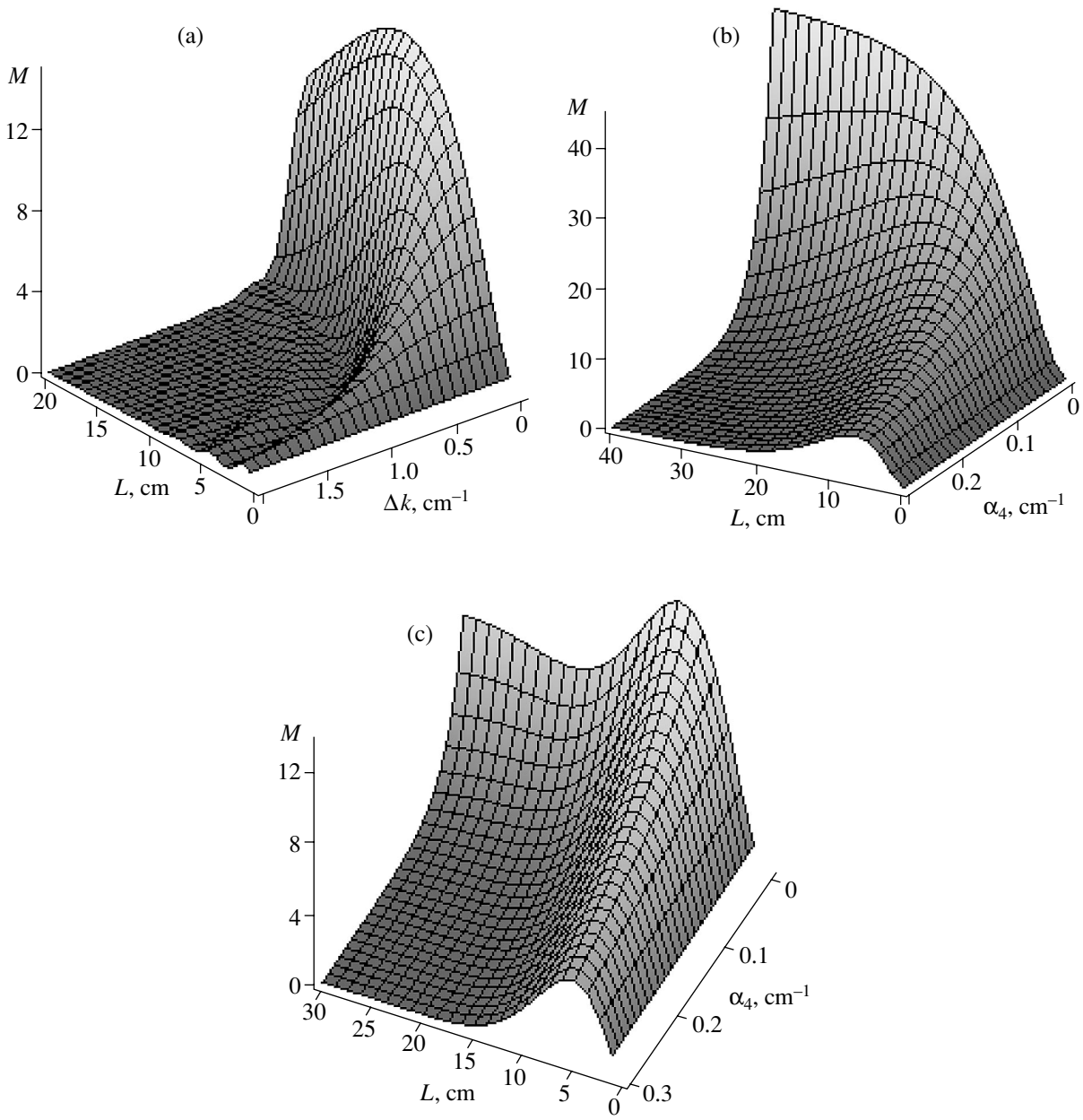

Fig. 7. The factor M, describing the influence of propagation effects on waveguide coherent anti-Stokes Raman scattering, as a function of the fiber length L and (a) the phase mismatch and (b, c) attenuation coefficient α_4 for (a) $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.1$ cm^{-1}, (b) $\alpha_1 = \alpha_2 = \alpha_3 = 0.1$ cm^{-1} and $\Delta\beta = 0$, and (c) $\alpha_1 = \alpha_2 = \alpha_4 = 0.1$ cm^{-1} and $\Delta\beta = 0.3$ cm^{-1}.

ing for the phase-matching integral [76]:

$$F_2 = 4\pi^2 \frac{\exp[-(k''/k')|\Delta k|b]}{(1 + k''/k')^2}. \quad (30)$$

In the case of small phase mismatches, $\Delta kl \ll \pi$, this limit is written as

$$F_2 = \frac{4\pi^2}{(1 + k''/k')^2}. \quad (31)$$

In the opposite limiting case of loosely focused pump beams, $b \gg l$, the phase-matching integral in the regime of weak absorption is reduced to [76]

$$F_2 = \frac{k'}{k''}\frac{4l^2}{b^2}\frac{\sin^2(\Delta kl/2)}{(\Delta kl/2)^2}. \quad (32)$$

In the case of small phase mismatches, Eq. (32) gives

$$F_2 = \frac{k'}{k''}\frac{4l^2}{b^2}. \quad (33)$$

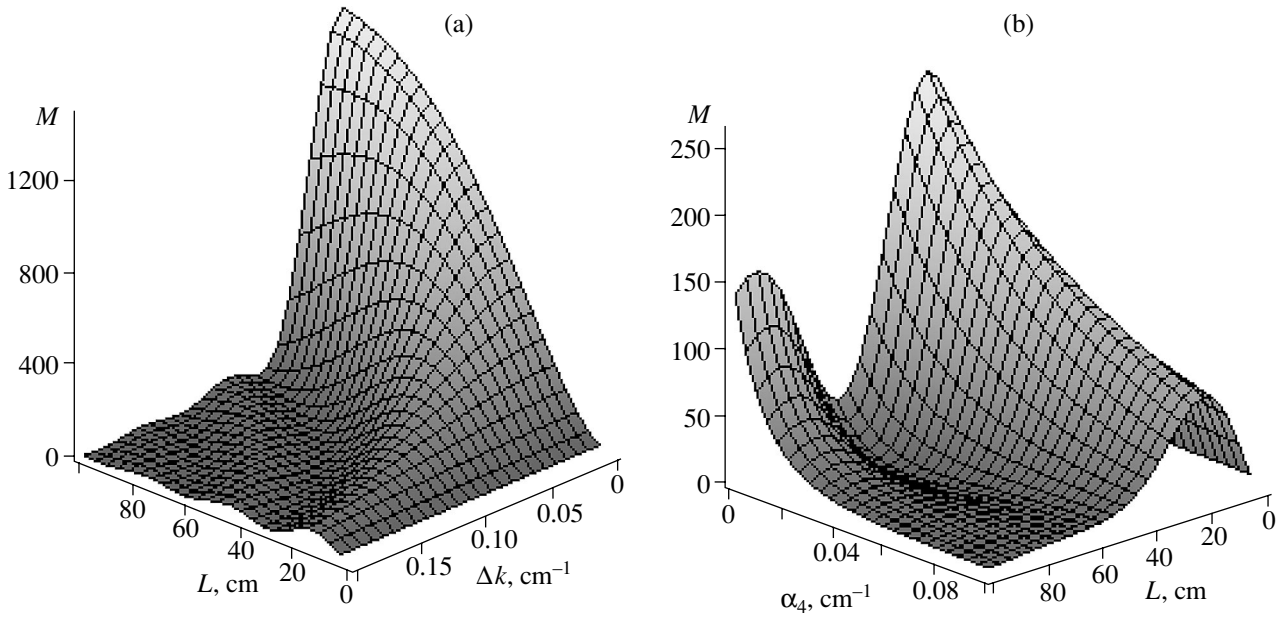

Fig. 8. The factor M, describing the influence of propagation effects on waveguide coherent anti-Stokes Raman scattering, as a function of the fiber length L and (a) the phase mismatch and (b) attenuation coefficient α_4 for (a) $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.01$ cm^{-1} and (b) $\alpha_1 = \alpha_2 = \alpha_3 = 0.01$ cm^{-1} and $\Delta\beta = 0.1$ cm^{-1}.

Assuming that the beam waist radius of focused pump beams is matched with the inner radius of a hollow fiber a having the length l, we find from Eqs. (32) and (33) that the waveguide CARS enhancement factor scales as $\lambda^2 l^2/a^4$ (see also [59, 60, 76]). The mode-overlapping integral should be generally included to describe the influence of transverse field profiles in the waveguide modes involved in the CARS process [72]. The length l can be made very large in the case of fibers, but the fundamental limitation of waveguide CARS in hollow fibers, experimentally demonstrated in [59–61], comes from optical losses, whose magnitude scales as λ^2/a^3 in the case of conventional, solid-cladding hollow fibers [62]. The influence of optical losses and phase-mismatch effects on the CARS process in the loose-focusing regime can be included via the factor [67]

$$M \propto \exp[-(\Delta\alpha + \alpha_4)l] \times \left[\frac{\sinh^2(\Delta\alpha l/2) + \sin^2(\Delta k l/2)}{(\Delta\alpha l/2)^2 + (\Delta k l/2)^2}\right] l^2, \quad (34)$$

where $\Delta\alpha = (\alpha_1 + \alpha_2 + \alpha_3 - \alpha_4)/2$ and α_1, α_2, α_3, and α_4 are the magnitudes of optical losses at frequencies ω_0, ω_1, ω_2, and ω_s, respectively.

It is straightforward to see from Eq. (34) that the amplitude of the CARS signal in a lossy waveguide reaches its maximum at some optimal length l_{opt}^{CARS} (Figs. 7, 8), which is given by

$$l_{\text{opt}}^{CARS} = \frac{1}{\Delta\alpha}\ln\left(\frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_4}\right). \quad (35)$$

With $\alpha_1 \approx \alpha_2 \approx \alpha_3 \approx \alpha_4 = \alpha$, Eq. (35) yields

$$l_{\text{opt}}^{CARS} = \frac{\ln 3}{\alpha}. \quad (36)$$

Then, setting $w_0 = 0.73a$ for the best matching of input beams with the fiber mode radius [62], assuming that the refractive index of the gas filling the fiber core is approximately equal to unity, and taking into consideration that $M = (3^{1/2} - 3^{-1/2})^2/(3\ln 3)^2 \approx 0.123$ for $\Delta k = 0$ and $l = l_{\text{opt}}^{CARS} = \ln 3/\alpha$, we substitute Eq. (36) for the optimal interaction length into Eqs. (33) and (34) to arrive at the following expression for the waveguide CARS enhancement factor in the regime of phase matching:

$$\mu = 1.3 \times 10^{-3}\frac{(k' + k'')^2}{k'k''}\frac{\lambda^2}{\alpha^2 a^4}. \quad (37)$$

Plugging optical losses of a solid-cladding hollow fiber into the CARS enhancement factor by substituting Eq. (16) into Eq. (37) with $u_n = 2.4$ for the limiting

eigenvalue of the EH_{11} mode, we derive the following expression for the factor of CARS enhancement in a solid-cladding hollow fiber relative to the tight-focusing regime in the case of exact phase matching, $\Delta k = 0$:

$$\rho = 6.1 \times 10^{-2} \frac{(k' + k'')^2}{k'k''} \left(\frac{a}{\lambda}\right)^2 \frac{n^2 - 1}{(n^2 + 1)^2}. \qquad (38)$$

The dependence of the CARS enhancement factor ρ in a solid-cladding hollow fiber on the inner fiber radius is shown by curve *1* in Fig. 9. Optical losses, growing with a decrease in the inner radius of the hollow fiber, limit CARS enhancement in hollow fibers of this type.

The situation radically changes in the case of a microstructure fiber. The waveguide enhancement of the CARS process in such fibers relative to the tight-focusing regime is given by Eq. (37). Waveguide losses in this case still represent the main physical factor limiting the waveguide CARS enhancement (Figs. 7, 8). However, the magnitude of optical losses in hollow microstructure fibers may be sufficiently low even for fibers with small inner diameters. The magnitude of optical losses for such fibers, as shown in [38], may be on the order of 1–3 dB/m in the case of fibers with a hollow core diameter of about 15 μm. Curves *2* and *3* in Fig. 9 show the CARS enhancement factor μ calculated as a function of the inner radius of a hollow microstructure fiber for two magnitudes of optical losses, $\alpha = 0.1$ and 0.01 cm^{-1}. In the case of small inner radii, microstructure fibers, as can be readily seen from Eqs. (37) and (38) for the parameters ρ and μ, provide much higher CARS enhancement factors than solid-core hollow fibers. The CARS enhancement factor in hollow microstructure fibers with the magnitude of optical losses equal to 0.1 and 0.01 cm^{-1} (curves *2* and *3* in Fig. 9) starts to exceed the CARS enhancement factor in a solid-cladding hollow fiber (curve *1* in Fig. 9) for core radii less than 20 and 45 μm, respectively. For hollow fibers with small core radii, the waveguide enhancement factor for the CARS process provided by microstructure fibers, as can be seen from Fig. 9, may be several orders of magnitude higher than the waveguide enhancement factor attainable with solid-cladding fibers. We assume in this consideration that the core–cladding geometry in a hollow microstructure fiber supports at least one air-guided mode even for small values of the inner radius of the fiber. A qualitative analysis of the number of air-guided modes supported by hollow photonic-crystal fibers has been provided by Cregan *et al.* [35]. More detailed, numerical simulations of air-guided modes in such fibers have been performed by Broeng *et al.* [78].

Comparison of Eqs. (21) and (37) shows that the scaling law of the waveguide CARS enhancement factor as a function of the magnitude of optical losses, fiber inner radius, and radiation wavelength differs from a similar scaling law of the waveguide SRS enhancement

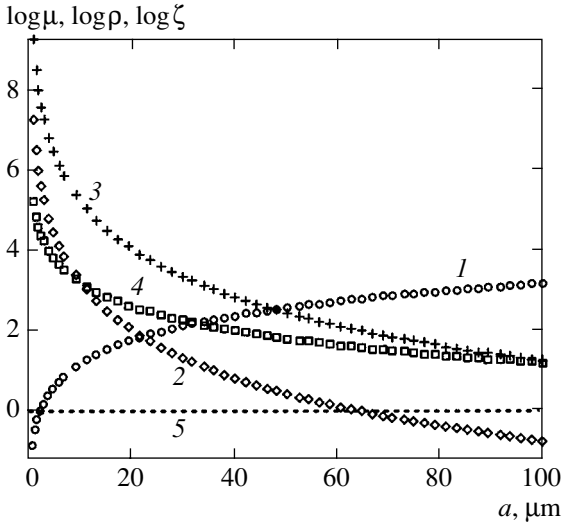

Fig. 9. Figures of merit for the efficiency of waveguide CARS in (*1*, ○) a standard hollow fiber with a solid cladding and (*2*, *3*) a hollow microstructure fiber with an attenuation coefficient $\alpha = 0.1$ (*2*, ◇), 0.01 cm^{-1} (*3*, +) as functions of the inner radius of the fiber *a*. (*4*, □) The figure of merit ζ for SRS efficiency in a hollow microstructure fiber with an attenuation coefficient $\alpha = 0.01$ cm^{-1} as a function of the inner radius of the fiber *a*. Dotted line *5* corresponds to the efficiency of CARS in the regime of tight focusing. The radiation wavelength is 0.5 μm.

factor (curve *4* in Fig. 9). Physically, this difference stems from differences in scattering mechanisms involved in SRS and CARS, with SRS and CARS signals building up in different fashions as functions of the interaction length and pump field amplitudes [cf. Eqs. (17) and (27)]. The difference in waveguide enhancement factors for SRS and CARS suggests different strategies for optimizing fibers designed to enhance these processes. Due to the stronger dependence of the CARS efficiency on the inner radius of a hollow fiber, the limiting waveguide enhancement for CARS in hollow microstructure fibers may noticeably exceed similar factors for SRS (cf. curves *2–4* in Fig. 9).

Phase mismatch, resulting from the difference in propagation constants of guided modes involved in the CARS process, is another important factor limiting the efficiency of CARS in a hollow fiber. In the case of nonzero phase mismatch, the optimal length for the CARS process can be found from a transcendental equation that immediately follows from Eq. (34):

$$\Delta\alpha \sinh(\Delta\alpha l_{\text{opt}}^{CARS}) + \Delta k \sin(\Delta k l_{\text{opt}}^{CARS}) + (\Delta\alpha + \alpha_4)[\cos(\Delta k l_{\text{opt}}^{CARS}) - \cosh(\Delta\alpha l_{\text{opt}}^{CARS})] = 0. \qquad (39)$$

Curve *1* in Fig. 10 shows the optimal length calculated by numerically solving Eq. (39) as a function of the phase mismatch Δk. The growth in the phase mis-

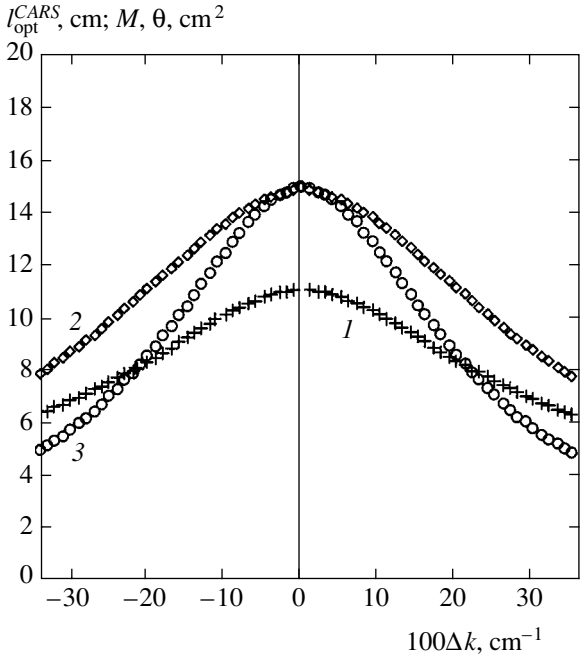

Fig. 10. The optimal length for the CARS process (line *1*, +), parameter $\theta = 0.123(l_{opt}^{CARS})^2$ (line *2*, ◇), and the factor M (line *3*, ○) calculated as functions of the phase mismatch Δk by numerically solving Eq. (39) with $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.1$ cm^{-1}.

match Δk, as can be seen from this dependence, reduces the optimal fiber length for the CARS process. Curves *2* and *3* of the same figure show the parameter $0.123(l_{opt}^{CARS})^2$, corresponding to the approximation of Eq. (37), and the factor M calculated by numerically solving Eq. (39). Comparison of these curves visualizes the deviation of the factor M, which includes the influence of waveguide losses and the phase mismatch, from the approximate dependence given by Eq. (37), which was employed in our analysis to find the limiting waveguide enhancement factors for the CARS process in hollow fibers (Fig. 7). The results presented in Fig. 10 show that the approximation of Eq. (37) works well in the case of perfect phase matching, when the factor M coincides with $0.123(l_{opt}^{CARS})^2$. The deviation of the factor M from $0.123(l_{opt}^{CARS})^2$ increases, as can be seen from Fig. 10, with the growth in the phase mismatch Δk.

Figure 11 displays the waveguide CARS enhancement factor in hollow microstructure fibers as a function of the inner fiber radius for different values of the phase mismatch, $\Delta k = 0$ (curve *1*, +), 0.3 cm^{-1} (curve *2*, ○), and 0.5 cm^{-1} (curve *3*, ◇), and magnitudes of optical losses, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.02$ cm^{-1} (Fig. 11a) and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.1$ cm^{-1} (Fig. 11b). Phase mismatch, as indicated by the results presented in Figs. 10 and 11, reduces the maximum waveguide CARS enhancement attainable with a hollow microstructure fiber, with the power of the CARS signal becoming an oscillating function of the fiber length (Figs. 7a, 8a, 8b, and 10). The characteristic period of these oscillations is determined by the coherence length. Oscillations become less pronounced and then completely flatten out as opti-

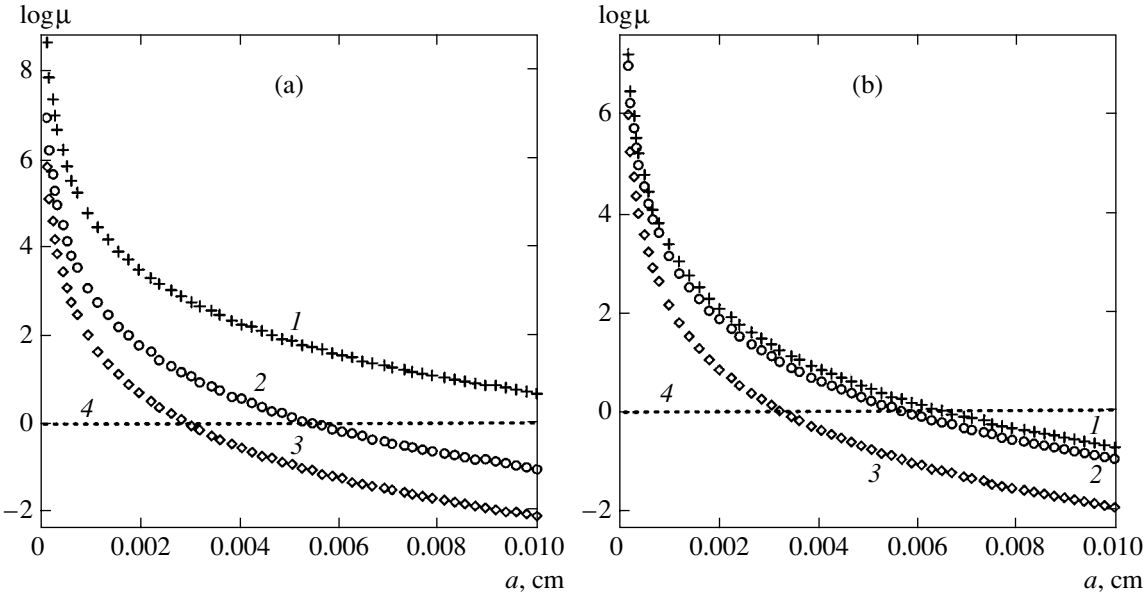

Fig. 11. The figure of merit for CARS efficiency in hollow microstructure fibers as a function of the inner radius of the fiber for different values of the phase mismatch, $\Delta k = 0$ (line *1*, +), 0.3 cm^{-1} (line *2*, ○), and 0.5 cm^{-1} (line *3*, ◇), and attenuation coefficients, $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.02$ cm^{-1} (a) and $\alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0.1$ cm^{-1} (b). Dotted line *4* corresponds to the efficiency of CARS in the regime of tight focusing.

cal losses build up (see also [79]). No oscillations are observed when the attenuation length becomes less than the coherence length (Figs. 7c, 8b). An important option offered by hollow microstructure fibers is the possibility of compensating for the phase mismatch related to the gas dispersion with an appropriate choice of waveguide parameters due to the waveguide dispersion component [50, 65].

8. CONCLUSIONS

The analysis performed in this paper shows that hollow microstructure and photonic-crystal fibers allow coherent anti-Stokes Raman scattering to be substantially enhanced relative to the regime of tight focusing, as well as relative to waveguiding regimes in hollow solid-cladding fibers. Based on the analysis of waveguide losses, as well as phase-mismatch and group-delay effects, we have determined the optimal fiber lengths providing maximum waveguide enhancement for SRS and CARS. Our analysis also reveals the existence of a physical limit for the waveguide enhancement of nonlinear-optical processes. Physically, this limit stems from the competition of diffraction and waveguide confinement of radiation due to the refractive-index step. Optimal confinement of electromagnetic radiation, allowing maximum enhancement of nonlinear-optical processes, is achieved in submicron waveguiding threads with an air cladding. Such waveguide structures with the maximum refractive-index step have been recently implemented in microstructure fibers. We have derived approximate asymptotic expressions that provide in several important cases an adequate qualitative understanding of the influence of fiber parameters on the optimal values of the fiber core radius and the maximum values of the waveguide enhancement factor and that allow these quantities to be estimated with satisfactory accuracy.

The maximum CARS enhancement in a hollow microstructure fiber relative to the tight-focusing regime was shown to scale as $\lambda^2/\alpha^2 a^4$ with radiation wavelength λ, inner fiber radius a, and magnitude of radiation losses α. Due to a rapid growth in CARS efficiency with a decrease in the inner radius of a hollow fiber, the limiting waveguide CARS enhancement factor in hollow microstructure fibers may substantially exceed analogous factors attainable with solid-cladding hollow fibers, as well as the waveguide enhancement for stimulated Raman scattering in hollow microstructure and photonic-crystal fibers. We have shown that the influence of group-delay effects on CARS and SRS processes in hollow fibers can be radically reduced due to the dispersion of waveguide modes with an appropriate choice of the gas pressure, the inner diameter of a hollow fiber, and the waveguide modes involved in the nonlinear-optical process.

Hollow-core microstructure and photonic-crystal fibers thus suggest ways of creating highly efficient gas sensors based on CARS spectroscopy, as well as SRS frequency converters. Waveguiding regimes supported by hollow-core photonic-crystal fibers allow the amount of sample gas required for spectroscopic analysis to be substantially reduced and provide the opportunity for performing nonlinear-optical experiments using low-power laser pulses.

ACKNOWLEDGMENTS

This study was supported in part by the President of Russian Federation Grant MD-42.2003.02, the Russian Foundation for Basic Research (project nos. 00-02-16929 and 02-02-17098), and the Volkswagen Foundation (project I/76 869). This material is also based upon the work supported by the European Research Office of the US Army (contract no. 9449-AM01).

REFERENCES

1. J. C. Knight, T. A. Birks, P. St. J. Russell, and D. M. Atkin, Opt. Lett. **21**, 1547 (1996).
2. J. C. Knight, J. Broeng, T. A. Birks, and P. St. J. Russell, Science **282**, 1476 (1998).
3. Focus Issue of Opt. Express **9** (13) (2001), Ed. by K. W. Koch.
4. Special Issue of J. Opt. Soc. Am. B **19** (2002), Ed. by C. M. Bowden and A. M. Zheltikov.
5. T. M. Monro, P. J. Bennett, N. G. R. Broderick, and D. J. Richardson, Opt. Lett. **25**, 206 (2000).
6. A. B. Fedotov, A. M. Zheltikov, L. A. Mel'nikov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 407 (2000) [JETP Lett. **71**, 281 (2000)].
7. M. V. Alfimov, A. M. Zheltikov, A. A. Ivanov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **71**, 714 (2000) [JETP Lett. **71**, 489 (2000)].
8. A. M. Zheltikov, Usp. Fiz. Nauk **170**, 1203 (2000) [Phys. Usp. **43**, 1125 (2000)].
9. A. M. Zheltikov, M. V. Alfimov, A. B. Fedotov, *et al.*, Zh. Éksp. Teor. Fiz. **120**, 570 (2001) [JETP **93**, 499 (2001)].
10. B. J. Eggleton, C. Kerbage, P. S. Westbrook, *et al.*, Opt. Express **9**, 698 (2001).
11. N. G. R. Broderick, T. M. Monro, P. J. Bennett, and D. J. Richardson, Opt. Lett. **24**, 1395 (1999).
12. A. B. Fedotov, A. M. Zheltikov, A. P. Tarasevitch, and D. von der Linde, Appl. Phys. B **73**, 181 (2001).
13. J. C. Knight, J. Arriaga, T. A. Birks, *et al.*, IEEE Photonics Technol. Lett. **12**, 807 (2000).
14. W. H. Reeves, J. C. Knight, P. St. J. Russell, and P. J. Roberts, Opt. Express **10**, 609 (2002).
15. A. M. Zheltikov, Ultrafast Photonics (in press).
16. J. K. Ranka, R. S. Windeler, and A. J. Stentz, Opt. Lett. **25**, 796 (2000).

17. St. Coen, A. H. L. Chau, R. Leonhardt, *et al.*, Opt. Lett. **26**, 1356 (2001).
18. S. Coen, A. Hing Lun Chau, R. Leonhardt, *et al.*, J. Opt. Soc. Am. B **19**, 753 (2002).
19. A. N. Naumov, A. B. Fedotov, A. M. Zheltikov, *et al.*, J. Opt. Soc. Am. B **19**, 2183 (2002).
20. A. B. Fedotov, A. N. Naumov, A. M. Zheltikov, *et al.*, J. Opt. Soc. Am. B **19**, 2156 (2002).
21. J. M. Dudley, Xun Gu, Lin Xu, *et al.*, Opt. Express **10**, 1215 (2002).
22. J. Herrmann, U. Griebner, N. Zhavoronkov, *et al.*, Phys. Rev. Lett. **88**, 173901 (2002).
23. J. K. Ranka, R. S. Windeler, and A. J. Stentz, Opt. Lett. **25**, 25 (2000).
24. W. J. Wadsworth, A. Ortigosa-Blanch, J. C. Knight, *et al.*, J. Opt. Soc. Am. B **19**, 2148 (2002).
25. S. A. Diddams, D. J. Jones, Jun Ye, *et al.*, Phys. Rev. Lett. **84**, 5102 (2000).
26. D. J. Jones, S. A. Diddams, J. K. Ranka, *et al.*, Science **288**, 635 (2000).
27. R. Holzwarth, T. Udem, T. W. Hansch, *et al.*, Phys. Rev. Lett. **85**, 2264 (2000).
28. S. N. Bagayev, A. K. Dmitriyev, S. V. Chepurov, *et al.*, Laser Phys. **11**, 1270 (2001).
29. Th. Udem, S. A. Diddams, K. R. Vogel, *et al.*, Phys. Rev. Lett. **86**, 4996 (2001).
30. A. Baltuska, T. Fuji, and T. Kobayashi, Opt. Lett. **27**, 1241 (2002).
31. S. N. Bagayev, V. S. Pivtsov, and A. M. Zheltikov, Quantum Electron. **32**, 311 (2002); A. M. Zheltikov, Laser Phys. **12**, 878 (2002).
32. S. Lakó, J. Seres, P. Apai, *et al.*, Appl. Phys. B **76**, 267 (2003).
33. A. B. Fedotov, Ping Zhou, A. P. Tarasevitch, *et al.*, J. Raman Spectrosc. **33**, 888 (2002).
34. I. Hartl, X. D. Li, C. Chudoba, *et al.*, Opt. Lett. **26**, 608 (2001).
35. R. F. Cregan, B. J. Mangan, J. C. Knight, *et al.*, Science **285**, 1537 (1999).
36. J. C. Knight and P. St. J. Russell, Science **296**, 276 (2002).
37. S. O. Konorov, A. B. Fedotov, O. A. Kolevatova, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **76**, 401 (2002) [JETP Lett. **76**, 341 (2002)].
38. F. Benabid, J. C. Knight, G. Antonopoulos, and P. St. J. Russell, Science **298**, 399 (2002).
39. F. Benabid, J. C. Knight, and P. St. J. Russell, Opt. Express **10**, 1195 (2002).
40. O. A. Kolevatova and A. M. Zheltikov, Laser Phys. (2003) (in press).
41. A. N. Naumov and A. M. Zheltikov, Kvantovaya Élektron. (Moscow) **32**, 129 (2002).
42. S. O. Konorov, A. B. Fedotov, O. A. Kolevatova, *et al.*, Laser Phys. **13**, 652 (2003).
43. G. P. Agrawal, *Nonlinear Fiber Optics* (Academic, Boston, 1989; Mir, Moscow, 1996).
44. A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman and Hall, New York, 1983; Radio i Svyaz', Moscow, 1987).
45. T. A. Birks, W. J. Wadsworth, and P. St. J. Russell, Opt. Lett. **25**, 1415 (2000).
46. D. A. Akimov, A. A. Ivanov, M. V. Alfimov, *et al.*, Appl. Phys. B **74**, 307 (2002).
47. A. B. Fedotov, A. N. Naumov, I. Bugar, *et al.*, IEEE J. Sel. Top. Quantum Electron. **8**, 665 (2002).
48. A. B. Fedotov, P. Zhou, A. N. Naumov, *et al.*, Appl. Phys. B **75**, 621 (2002).
49. A. M. Zheltikov, Ping Zhou, V. V. Temnov, *et al.*, Quantum Electron. **32**, 542 (2002).
50. A. M. Zheltikov, Usp. Fiz. Nauk **172**, 743 (2002) [Phys. Usp. **45**, 687 (2002)].
51. M. Nisoli, S. De Silvestri, and O. Svelto, Appl. Phys. Lett. **68**, 2793 (1996).
52. M. Nisoli, S. De Silvestri, O. Svelto, *et al.*, Opt. Lett. **22**, 522 (1997).
53. O. Dühr, E. T. J. Nibbering, G. Korn, *et al.*, Opt. Lett. **24**, 34 (1999).
54. N. Zhavoronkov and G. Korn, Phys. Rev. Lett. **88**, 203901 (2002).
55. C. G. Durfee III, S. Backus, M. M. Murnane, and H. C. Kapteyn, Opt. Lett. **22**, 1565 (1997).
56. A. Rundquist, C. G. Durfee III, Z. Chang, *et al.*, Science **280**, 1412 (1998).
57. Y. Tamaki, K. Midorikawa, and M. Obara, Appl. Phys. B **67**, 59 (1998).
58. E. Constant, D. Garzella, P. Breger, *et al.*, Phys. Rev. Lett. **82**, 1668 (1999).
59. R. B. Miles, G. Laufer, and G. C. Bjorklund, Appl. Phys. Lett. **30**, 417 (1977).
60. A. B. Fedotov, F. Giammanco, A. N. Naumov, *et al.*, Appl. Phys. B **72**, 575 (2001).
61. A. N. Naumov, F. Giammanco, D. A. Sidorov-Biryukov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **73**, 263 (2001) [JETP Lett. **73**, 228 (2001)].
62. E. A. J. Marcatili and R. A. Schmeltzer, Bell Syst. Tech. J. **43**, 1783 (1964).
63. M. J. Adams, *An Introduction to Optical Waveguides* (Wiley, New York, 1981; Mir, Moscow, 1984).
64. A. Yariv and P. Yeh, *Optical Waves in Crystals: Propagation and Control of Laser Radiation* (Wiley, New York, 1984; Mir, Moscow, 1987).
65. A. M. Zheltikov and A. N. Naumov, Quantum Electron. **31**, 471 (2001).
66. S. A. Akhmanov and N. I. Koroteev, *Methods of Nonlinear Optics in Light Scattering Spectroscopy* (Nauka, Moscow, 1981).
67. G. L. Eesley, *Coherent Raman Spectroscopy* (Pergamon Press, Oxford, 1981).

68. S. A. J. Druet and J.-P. E. Taran, Prog. Quantum Electron. **7**, 1 (1981).

69. A. C. Eckbreth, *Laser Diagnostics for Combustion Temperature and Species* (Abacus, Cambridge, MA, 1988).

70. W. Kiefer, J. Mol. Struct. **59**, 305 (1980).

71. A. M. Zheltikov and N. I. Koroteev, Usp. Fiz. Nauk **170**, 1203 (1999) [Phys. Usp. **43**, 1125 (1999)].

72. G. I. Stegeman, R. Fortenberry, C. Karaguleff, *et al.*, Opt. Lett. **8**, 295 (1983).

73. W. P. de Boeij, J. S. Kanger, G. W. Lucassen, *et al.*, Appl. Spectrosc. **47**, 723 (1993).

74. J. S. Kanger, C. Otto, and J. Greve, Appl. Spectrosc. **49**, 1326 (1995).

75. S. O. Konorov, D. A. Akimov, A. N. Naumov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **75**, 74 (2002) [JETP Lett. **75**, 66 (2002)]; S. O. Konorov, D. A. Akimov, A. N. Naumov, *et al.*, J. Raman Spectrosc. **33**, 955 (2002).

76. M. D. Levenson and J. J. Song, in *Coherent Nonlinear Optics*, Ed. by M. S. Feld and V. S. Letokhov (Springer, Berlin, 1980).

77. P. D. Maker and R. W. Terhune, Phys. Rev. A **137**, 801 (1965).

78. J. Broeng, S. E. Barkou, T. Söndergaard, and A. Bjarklev, Opt. Lett. **25**, 96 (2000).

79. N. I. Koroteev, A. N. Naumov, and A. M. Zheltikov, Laser Phys. **4**, 1160 (1994).

Translated by A. Zheltikov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 522–538.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 578–595.

ATOMS, SPECTRA, RADIATION

Dynamics of Spontaneous Radiation of Atoms Scattered by a Resonance Standing Light Wave

M. V. Fedorov[a], M. A. Efremov[a,*], V. P. Yakovlev[b], and W. P. Schleich[c]

[a]*General Physics Institute, Russian Academy of Sciences, Moscow, 119991 Russia*
[b]*Moscow Engineering Physics Institute (State University), Moscow, 115409 Russia*
[c]*Abteilung für Quantenphysik, Universität Ulm, 89069, Ulm, Germany*
**e-mail: efremov@ran.gpi.ru*

Received March 28, 2003

Abstract—The scattering of atoms by a resonance standing light wave is considered under conditions when the lower of two resonance levels is metastable, while the upper level rapidly decays due to mainly spontaneous radiative transitions to the nonresonance levels of an atom. The diffraction scattering regime is studied, when the Rabi frequency is sufficiently high and many diffraction maxima are formed due to scattering. The dynamics of spontaneous radiation of an atom is investigated. It is shown that scattering slows down substantially the radiative decay of the atom. The regions and characteristics of the power and exponential decay are determined. The adiabatic and nonadiabatic scattering regimes are studied. It is shown that the wave packets of atoms in the metastable and resonance excited states narrow down during scattering. A limiting (minimal) size of the wave packets is found, which is achieved upon nonadiabatic scattering in the case of a sufficiently long interaction time. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The scattering of atoms by a resonance standing light wave has been studied in many papers (see, for example, monographs [1–3] and references therein). Of special interest are papers [4–9] devoted to the study of a peculiar type of scattered atoms, which were initially excited to one of the long-lived metastable states. The corresponding metastable level E_m is the lower level of a resonance two-level system, whereas the upper level E_e of this system has a large width Γ caused mainly by the spontaneous radiative decay to states different from the metastable state (Fig. 1). In experiments described in papers [6, 9], metastable Ar* atoms were used [$E_m \equiv 1s_5$ (J = 2)]. The field of a 801-nm standing wave performed the resonance interaction of the E_m level with the $E_e \equiv 2p_8$ level ($J = 2$). About 72% of the width of the excited level E_e was caused by radiative transitions to the levels different from E_m.

The assumption about a large width Γ of the E_e level means that $\Gamma \gg \omega_r$, where $\omega_r = k^2/2M$ is the recoil frequency or energy ($\hbar = 1$), M is the atom mass, $k = 2\pi/\lambda = \omega/c$, and ω is the field frequency. Typically, $\omega_r/\Gamma \sim 10^{-3} \ll 1$.

Another important parameter characterizing the interaction of atoms with the field is the Rabi frequency $\Omega = 2\mathbf{d}_{me} \cdot \mathbf{E}_0$, where $\mathbf{d}_{me} = \langle m|\mathbf{d}|e\rangle$ is the matrix element of the dipole moment of the transition between reso-

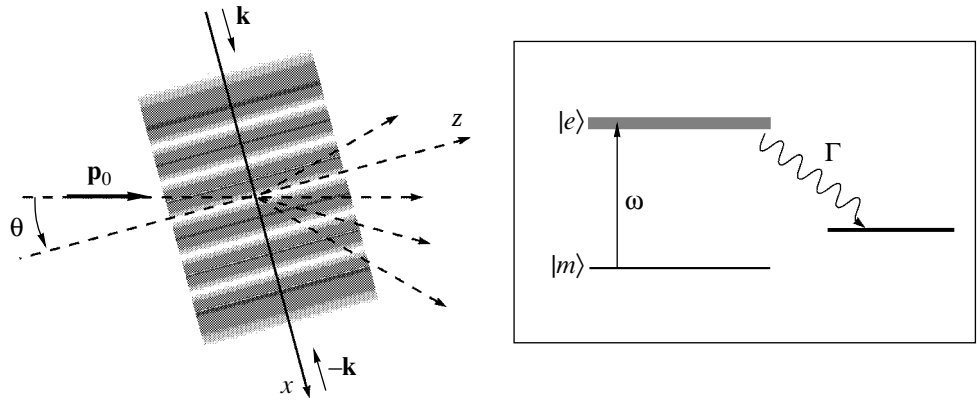

Fig. 1. Scheme of scattering of an atom by a standing light wave and the energy level diagram.

1063-7761/03/9703-0522$24.00 © 2003 MAIK "Nauka/Interperiodica"

nance levels and $\mathbf{E}_0$ is the amplitude of the field strength of each of the counterpropagating light waves forming a standing wave. Depending on the relation between Ω and $4\sqrt{\Gamma\omega_r}$, either Bragg {for $\Omega \ll 4\sqrt{\Gamma\omega_r}$} or diffraction {for $\Omega \gg 4\sqrt{\Gamma\omega_r}$} scattering can occur (see Section 3). Bragg scattering is characterized by a strong dependence of the scattering efficiency on the direction of the initial atomic momentum $\mathbf{p}_0$: scattering is efficient only if the angle θ between the vector $\mathbf{p}_0$ and perpendicular to the direction of light wave vectors $\pm\mathbf{k}$ (the z axis in Fig. 1) is close to the Bragg angle $\pm\theta_{Br}$, where $\theta_{Br} = k/p_0 \ll 1$. In this case, the direction of propagation of scattered atoms approximately corresponds to specular reflection from the plane perpendicular to $\mathbf{k}$; i.e., the angle between the direction of the momentum of scattered electrons and the z axis is close to $\mp\theta_{Br}$. The efficiency of diffraction scattering is almost independent of the angle θ. In addition, many diffraction maxima appear due to a more intense interaction upon scattering. The directions of propagation of metastable atoms in them are determined by the angles $\theta_n = \theta + 2n\theta_{Br}$, where $n = 0, \pm1, \pm2, \ldots$ In papers [4–9], Bragg scattering was mainly considered. In this paper, we studied diffraction scattering, which assumes greater values of the Rabi frequency Ω. We showed that these two cases of scattering substantially differ from each other. In particular, while an atom always decays exponentially in the case of Bragg scattering, in the case of diffraction scattering, as is shown below, the nonexponential decay can occur when the time-dependent probabilities of finding an atom at levels E_m and E_e decrease according to a power law. We also showed that the region of the nonexponential decay is restricted in time. The decay becomes exponential at very long times. However, the decay rate in this case substantially differs from the decay rate Γ for a free atom. The modified decay rate $\tilde{\Gamma}$ decreases with increasing Γ and depends on the value and sign of the resonance detuning $\Delta = E_e - E_m - \omega$, as well as on the Rabi frequency Ω and the recoil frequency ω_r. Finally, the dynamics of the population decay proves to be closely related to the formation of narrow wave packets of the "center of mass" of an atom, which are in the metastable and excited states. We will show below that there exists a limiting, minimal packet size, which can be achieved at very long interaction times and can be important for practical purposes.

2. BASIC EQUATIONS

The wave function of an atom interacting with a light field depends on the radius vector $\mathbf{r}$ of the center of mass of the atom, the intra-atomic variables, and on time t, satisfying the Schrödinger equation

$$i\frac{\partial}{\partial t}\Psi = \left\{-\frac{1}{2M}\nabla^2 + H_{at} - \mathbf{d}\cdot\mathbf{E}(\mathbf{r}, t)\right\}\Psi, \qquad (1)$$

where H_{at} is the Hamiltonian of intra-atomic motion, $\nabla = \partial/\partial\mathbf{r}$, and $\mathbf{E}(\mathbf{r}, t)$ is the electric-field strength, which has the form

$$\mathbf{E}(\mathbf{r}, t) = 2\mathbf{E}_0\cos(\omega t)\cos(kx) \qquad (2)$$

for a standing light wave (the x axis is directed along $\mathbf{k}$) (Fig. 1).

In the rotating wave approximation, the wave function of an atom can be written in the form

$$\Psi(\mathbf{r}, t) = \frac{1}{(2\pi)^{3/2}}\exp\left(-iE_m t + i\mathbf{p}_0\cdot\mathbf{r} - i\frac{p_0^2 t}{2M}\right) \times [\varphi_m(x, t)|m\rangle + \exp(-i\omega t)\varphi_e(x, t)|e\rangle], \qquad (3)$$

where $\varphi_m(x, t)$ and $\varphi_e(x, t)$ are the wave functions of the perturbed motion of the center of mass of the atom in the states $|m\rangle$ and $|e\rangle$, respectively. Equations for $\varphi_m(x, t)$ and $\varphi_e(x, t)$ follow directly from the Schrödinger equation (1):

$$i\frac{\partial}{\partial t}\varphi_m = -\frac{1}{2M}\left(\frac{\partial^2}{\partial x^2} + 2ip_{0x}\frac{\partial}{\partial x}\right)\varphi_m - \frac{\Omega}{2}\cos(kx)\varphi_e, \qquad (4)$$

$$i\frac{\partial}{\partial t}\varphi_e = -\left\{\frac{1}{2M}\left(\frac{\partial^2}{\partial x^2} + 2ip_{0x}\frac{\partial}{\partial x}\right) + \Delta + i\frac{\Gamma}{2}\right\}\varphi_e - \frac{\Omega^*}{2}\cos(kx)\varphi_m. \qquad (5)$$

We assume that the radiative width Γ of the excited level E_e is mainly caused by transitions to levels different from E_m, spontaneous radiation due to the $E_e \longrightarrow E_m$ transitions being neglected within the framework of the model.

We assume that the interaction of the atom with the field is switched on instantly at the moment $t = 0$, while for $t < 0$, the atom is in the metastable state $|m\rangle$ and the wave function of its center of mass is a purely plane wave with the momentum $\mathbf{p}_0$, which corresponds to the initial conditions

$$\varphi_m(x, t = 0) = 1, \quad \varphi_e(x, t = 0) = 0 \qquad (6)$$

in system of equations (4), (5).

The squares of the modulus of functions $\varphi_{m,e}(x, t)$ determine the probability densities of finding atoms in states $|m\rangle$ or $|e\rangle$ at instant t in the vicinity of point x:

$$\frac{dW_{m,e}(x,t)}{dx} = \frac{k}{\pi}|\varphi_{m,e}(x,t)|^2. \quad (7)$$

The probability densities $dW_{m,e}(x, t)/dx$ (7) integrated over the entire interval of the change in variable x from 0 to π/k give the total probabilities of finding the atom at instant t in the metastable or excited states:

$$W_{\text{tot}}^{(m,e)}(t) = \int_0^{\pi/k} dx \frac{dW_{m,e}(x,t)}{dx} = \frac{k}{\pi}\int_0^{\pi/k} dx|\varphi_{m,e}(x,t)|^2. \quad (8)$$

An alternative method for describing the interaction of scattered atoms with the field is based on the use of the periodic dependence of the field $\mathbf{E}(\mathbf{r}, t)$ (2) on variable x, which makes it possible to expand the function $\varphi_{m,e}(x, t)$ in a Fourier series with time-dependent coefficients:

$$\varphi_{m,e}(x,t) = \sum_n a_n^{(m,e)}(t)\exp(inkx),$$
$$a_n^{(m,e)}(t) = \frac{k}{\pi}\int_0^{\pi/k} dx\exp(-inkx)\varphi_{m,e}(x,t). \quad (9)$$

The physical meaning of coefficients $a_n^{(m,e)}(t)$ is that they are the partial amplitudes of the probability of finding atoms in the metastable and excited states at instant t in beams deflected from the initial propagation direction by the angle $n\theta_{Br}$, $n = 0, \pm1, \pm2, \ldots$ Functions $a_n^{(m,e)}(t)$ satisfy equations that follows from Eqs. (4), (5), and (9):

$$i\dot{a}_n^{(m)}(t) = (n^2\omega_r + n\delta)a_n^{(m)} - \frac{\Omega}{4}(a_{n-1}^{(e)} + a_{n+1}^{(e)}), \quad (10)$$

$$i\dot{a}_n^{(e)}(t) = \left(n^2\omega_r + n\delta - \Delta - \frac{i\Gamma}{2}\right)a_n^{(e)} - \frac{\Omega^*}{4}(a_{n-1}^{(m)} + a_{n+1}^{(m)}), \quad (11)$$

where

$$\delta = \frac{kp_{0x}}{M} = 2\omega_r\frac{\theta}{\theta_{Br}}.$$

It is obvious that functions $a_n^{(m,e)}(t)$ should satisfy the initial conditions

$$a_n^{(m)}(t=0) = \delta_{n,0}, \quad a_n^{(e)}(t=0) = 0. \quad (12)$$

It follows from Eqs. (10)–(12), in particular, that only even or odd values of n are possible for atoms in the metastable or excited states, respectively.

The squares of the modulus of functions $a_n^{(m,e)}(t)$ are the partial probabilities of finding atoms at instant t in beams deflected by angles $n\theta_{Br}$:

$$W_n^{(m,e)}(t) = \left|a_n^{(m,e)}(t)\right|^2. \quad (13)$$

Sums of partial probabilities $W_n^{(m)}(t)$ and $W_n^{(e)}(t)$ determine the total probabilities of finding atoms in the metastable and excited states after scattering,

$$W_{\text{tot}}^{(m,e)}(t) = \sum_n W_n^{(m,e)}(t). \quad (14)$$

Definitions (14) and (8) are equivalent, as well as the groups of Eqs. (4)–(6) and (10)–(12), corresponding to two different but equivalent approaches to the solution of the problem in the coordinate and momentum representation.

3. BRAGG SCATTERING AND THE BORRMANN EFFECT

Consider first Bragg scattering when the Rabi frequency $|\Omega|$ is sufficiently low, while the excited-level width Γ is large. We assume that interaction time t is also large compared to $1/\Gamma$, so that

$$\Gamma \gg n^2\omega_r, |n\delta|, \frac{1}{t}, |\Omega|, |\Delta|. \quad (15)$$

Such a scattering regime was theoretically studied in papers [7, 8], and it corresponds to the experimental conditions described in [6, 9]. Here, we will briefly discuss this regime to distinguish it distinctly from diffraction scattering.

According to conditions (15), we can omit the terms $n^2\omega_r a_n^{(e)}$, $n\delta a_n^{(e)}$, $\Delta a_n^{(e)}$, and $i\dot{a}_n^{(e)}$ in Eqs. (11) to obtain

$$a_n^{(e)}(t) \approx \frac{i\Omega^*}{2\Gamma}(a_{n+1}^{(m)}(t) + a_{n-1}^{(m)}(t)). \quad (16)$$

By substituting $a_n^{(e)}(t)$ (16) into Eqs. (10), we obtain

$$i\dot{a}_n^{(m)}(t) = (n^2\omega_r + n\delta)a_n^{(m)}(t) - i\frac{|\Omega|^2}{8\Gamma}(a_{n-2}^{(m)}(t) + 2a_n^{(m)}(t) + a_{n+2}^{(m)}(t)). \quad (17)$$

For $\delta \approx \pm 2\omega_r$ or $|\delta \mp 2\omega_r| \ll \omega_r$ (i.e., when the Wolf–Bragg condition $\theta \approx \pm\theta_{Br}$ is fulfilled), the sum $n^2\omega_r + n\delta$ is zero for $n = 0$ and is small for $n = \mp 2$. The smallness condition $|\delta \mp 2\omega_r|$ can be considered as the Bragg resonance condition at which the angle of incidence of an atomic beam is close to the Wolf–Bragg angle. For a low Rabi frequency $|\Omega|$ and a small Bragg resonance detuning $\delta \mp 2\omega_r$, only the probability amplitudes $a_n^{(m)}(t)$ with $n = 0$ and ∓ 2 are not small in Eqs. (17). Retaining only these terms, we reduce the entire infinite-dimensional system of equations (17) to only two equations

$$i\dot{a}_0^{(m)}(t) = -i\frac{|\Omega|^2}{8\Gamma}(a_{\mp 2}^{(m)}(t) + 2a_0^{(m)}(t)), \quad (18)$$

$$i\dot{a}_{\mp 2}^{(m)}(t) = (4\omega_r \mp 2\delta)a_{\mp 2}^{(m)}(t) - i\frac{|\Omega|^2}{8\Gamma}(2a_{\mp 2}^{(m)}(t) + a_0^{(m)}(t)), \quad (19)$$

where the probability amplitudes $a_0^{(m)}(t)$ and $a_{\mp 2}^{(m)}(t)$ are related to the unscattered and scattered atomic beams, respectively, which are directed approximately at angles $\pm\theta_{Br}$ and $\mp\theta_{Br}$ to the z axis (Fig. 1).

Equations (18) and (19) have stationary solutions of the type

$$a_n^{(m)}(t) \propto \exp(-i\gamma t),$$

where γ is the quasi-energy determined from the condition that the determinant of the system of equations (18) and (19) is zero:

$$\begin{vmatrix} \gamma + i\frac{|\Omega|^2}{4\Gamma} & i\frac{|\Omega|^2}{8\Gamma} \\ i\frac{|\Omega|^2}{8\Gamma} & \gamma + i\frac{|\Omega|^2}{4\Gamma} - (4\omega_r \mp 2\delta) \end{vmatrix} = 0, \quad (20)$$

which gives

$$\gamma_\pm = -i\frac{|\Omega|^2}{4\Gamma} + (2\omega_r \mp \delta) \pm \sqrt{(2\omega_r \mp \delta)^2 - \frac{|\Omega|^4}{64\Gamma^2}}. \quad (21)$$

As follows from these equations, far from the resonance, for

$$|2\omega_r \pm \delta| > \frac{|\Omega|^2}{8\Gamma},$$

the decay rates of two quasi-energy levels are identical:

$$\mathrm{Im}\,\gamma_+ = \mathrm{Im}\,\gamma_- = -\frac{|\Omega|^2}{4\Gamma}. \quad (22)$$

On the contrary, for the exact resonance, when $2\omega_r \mp \delta = 0$, we have

$$\mathrm{Im}\gamma_+ = -\frac{|\Omega|^2}{8\Gamma}, \quad \mathrm{Im}\gamma_- = -3\frac{|\Omega|^2}{8\Gamma}. \quad (23)$$

The first of these solutions decays slower ($\propto \exp[-t|\Omega|^2/8\Gamma]$), while the second solution decays faster ($\propto \exp[-3t|\Omega|^2/8\Gamma]$), than solutions far from the resonance ($\propto \exp[-t|\Omega|^2/4\Gamma]$). Therefore, a solution appearing at the resonance (for $\theta = \pm 2\theta_{Br}$) has a longer lifetime than that beyond the resonance. If $t|\Omega|^2/4\Gamma \gtrsim 1$, then a fraction of atoms remaining at the metastable level in the case of resonance will be greater than that beyond the resonance. It is this effect that was experimentally observed in papers [6, 9] and was interpreted as an anomalous propagation of metastable atoms through a standing light wave in the case of Bragg resonance or as an optical analog of the Borrmann effect (anomalous penetration of X-rays into a crystal incident at the Wolf–Bragg angle on the crystal lattice plane) [10].

In the problem of scattering of atoms under the conditions considered, the decay remains exponential, in accordance with experimental data [6, 9]. In this case, the product of the decay constant $|\Omega|^2/4\Gamma$ by interaction time t cannot be too large. Otherwise, the number of metastable atoms remaining at the metastable level will be so small that it will be impossible to detect them both beyond and in the resonance. The resonance width with respect to parameter δ is $|\Omega|^2/8\Gamma$, while the width of the population peak $W_{\mathrm{tot}}^{(m)}$ depending on θ/θ_{Br} is $|\Omega|^2/16\Gamma\omega_r$. If $|\Omega|^2/16\Gamma \ll \omega_r$, Bragg scattering takes place. In the opposite case, when $|\Omega|^2/16\Gamma \gg \omega_r$, diffraction scattering occurs.

4. ADIABATIC APPROXIMATION

Let us now analyze the case of high Rabi frequencies $|\Omega| \gg 4\sqrt{\Gamma\omega_r}$ corresponding to diffraction scattering. Because in this case the dependence of the scattering pattern on the angle of incidence θ is weak, we consider the case of normal incidence $\theta = 0$. In addition, at the first stage, we will treat the scattering and radiative decay of atoms in the adiabatic approximation. In this

approximation, the kinetic energy of the motion of the center of mass of an atom in the direction parallel to **k** is assumed so small,

$$\left\langle \frac{1}{2M}\frac{\partial^2}{\partial x^2} \right\rangle \ll \left|\Delta + i\frac{\Gamma}{2}\right|, \quad \frac{1}{t}, \tag{24}$$

that the corresponding terms can be omitted in both Eqs. (4) and (5). As a result, these equations take the form

$$i\frac{d}{dt}\varphi_m(x, t) = -\frac{\Omega}{2}\cos(kx)\varphi_e, \tag{25}$$

$$i\frac{d}{dt}\varphi_e(x, t) = -\left(\Delta + i\frac{\Gamma}{2}\right)\varphi_e - \frac{\Omega^*}{2}\cos(kx)\varphi_m. \tag{26}$$

Moreover, these two equations can easily be reduced to one second-order equation:

$$\left\{\frac{d^2}{dt^2} - i\left(\Delta + i\frac{\Gamma}{2}\right)\frac{d}{dt} + \frac{|\Omega|^2}{4}\cos^2(kx)\right\} \times \varphi_m(x, t) = 0. \tag{27}$$

Finally, Eqs. (25) and (26) can be written as one matrix equation of the Schrödinger type,

$$i\frac{d}{dt}\Phi(x, t) = H_{ad}\Phi(x, t), \tag{28}$$

for the two-component wave function

$$\Phi = \begin{pmatrix} \varphi_m(x, t) \\ \varphi_e(x, t) \end{pmatrix} \tag{29}$$

with the Hamiltonian

$$H_{ad} \approx \begin{pmatrix} 0 & -\frac{\Omega}{2}\cos(kx) \\ -\frac{\Omega^*}{2}\cos(kx) & -\Delta - \frac{i}{2}\Gamma \end{pmatrix}. \tag{30}$$

In terms of the partial probability amplitudes $a_n^{(m, e)}(t)$ (9), the adiabatic approximation is equivalent to neglecting the terms proportional to $n^2\omega_r$ in Eqs. (10) and (11), which take the form

$$i\dot{a}_n^{(m)}(t) = -\frac{\Omega}{4}(a_{n-1}^{(e)} + a_{n+1}^{(e)}), \tag{31}$$

$$i\dot{a}_n^{(e)}(t) = -\left(\Delta + i\frac{\Gamma}{2}\right)a_n^{(e)} - \frac{\Omega^*}{4}(a_{n-1}^{(m)} + a_{n+1}^{(m)}) \tag{32}$$

in the case of normal incidence ($\theta = 0$).

Similarly to passage from (25), (26) to (27), two groups of first-order equations (31) and (32) can be reduced to one group of second-order equations:

$$\left(i\frac{d}{dt} + \Delta + i\frac{\Gamma}{2}\right)i\frac{d}{dt}a_n^{(m)}(t) = \frac{|\Omega|^2}{16}(a_{n-2}^{(m)} + 2a_n^{(m)} + a_{n+2}^{(m)}). \tag{33}$$

5. NONEXPONENTIAL DECAY AND NARROWING DOWN OF ATOMIC WAVE PACKETS

Consider first of all a region of sufficiently long interaction times *t*. The second derivative with respect to time in Eq. (27) is small compared to the term proportional to the first derivative:

$$\left|\frac{d^2\varphi_m}{dt^2}\right| \ll \left|\left(\Delta + i\frac{\Gamma}{2}\right)\frac{d\varphi_m}{dt}\right|. \tag{34}$$

The real meaning of this assumption is explained below. Here, using condition (34), we omit in Eq. (27) the term containing the second derivative to obtain the equation of the type

$$i\frac{d}{dt}\varphi_m(x, t) = V_{CY}(x)\varphi_m(x, t), \tag{35}$$

where $V_{CY}(x)$ is the complex Chudesnikov–Yakovlev potential [4]

$$V_{CY}(x) = \frac{1}{4}\frac{|\Omega|^2}{\Delta + i\Gamma/2}\cos^2(kx). \tag{36}$$

It is obvious that the solution of Eq. (35) has the form (see also [5])

$$\varphi_m(x, t) = \exp\{-itV_{CY}(x)\} = \exp\left\{-\frac{it}{4}\frac{|\Omega|^2}{\Delta + i\Gamma/2}\cos^2(kx)\right\}. \tag{37}$$

Now, the function $\varphi_e(x, t)$ also can be easily found using Eq. (25):

$$\varphi_e(x, t) = -\frac{1}{2}\frac{\Omega^*}{\Delta + i\Gamma/2}\cos(kx) \times \exp\left\{-\frac{it}{4}\frac{|\Omega|^2}{\Delta + i\Gamma/2}\cos^2(kx)\right\}. \tag{38}$$

According to (7), the probability densities of finding atoms in the metastable and excited states in the vicin-

ity of the point x are determined by the squares of absolute values of functions $\varphi_m(x, t)$ (37) and $\varphi_e(x, t)$ (38):

$$\frac{dW_m(x, t)}{dx} = \frac{k}{\pi}\exp\left\{-\frac{|\Omega|^2\Gamma t}{\Gamma^2 + 4\Delta^2}\cos^2(kx)\right\} \quad (39)$$

and

$$\frac{dW_e(x, t)}{dx} = \frac{k}{\pi}\frac{|\Omega|^2}{\Gamma^2 + 4\Delta^2}\cos^2(kx) \times \exp\left\{-\frac{|\Omega|^2\Gamma t}{\Gamma^2 + 4\Delta^2}\cos^2(kx)\right\}. \quad (40)$$

The dependences

$$|\varphi_m(x, t)|^2 \propto \frac{dW_{m, e}(x, t)}{dx}$$

of (39) and (40) on the x coordinate for a fixed interaction time t are presented in Fig. 2. The width of the distributions $dW_{m, e}(x, t)/dx$ (39) and (40) is on the order of

$$\Delta x = \frac{2}{k|\Omega|}\sqrt{\frac{\Delta^2 + \Gamma^2/4}{\Gamma t}} \equiv \frac{1}{k}\sqrt{\frac{t_0}{2t}}, \quad (41)$$

where

$$t_0 = 2\frac{\Gamma^2 + 4\Delta^2}{|\Omega|^2\Gamma}. \quad (42)$$

Equation (41) shows that for $t > t_0$, the width $\Delta x(t)$ decreases with increasing interaction time t; i.e., the wave packets of atoms appearing in the metastable and excited states narrow down and their widths become smaller than the wavelength of light, $\lambda = 2\pi/k$.

Solution (37) of Eq. (35) and the width of wave packets (41) make it possible to specify condition (34) for the passage from second-order Eq. (27) to first-order Eq. (35). Indeed, the first and second derivatives of function $\varphi_m(x, t)$ entering Eq. (34) can be easily found from (37). In the case of narrow wave packets, the quantity $\cos(kx)$ can be replaced by the quantity $-k\Delta x$ determined by expression (41), which finally reduces condition (34) to the form

$$t > t_0\frac{|\Omega|^2}{\Delta^2 + \Gamma^2/4} \sim \frac{1}{\Gamma}. \quad (43)$$

Equation (43) determines the region of applicability of first-order equation approximation (35). The characteristic time t_0 (42) introduced above can lie both outside and inside this region depending on whether the Rabi frequency $|\Omega|$ is higher or lower than $|\Delta + i\Gamma/2|$. Therefore, in the general case, instant t, beginning from which the narrowing of the wave packets appears, can be defined as

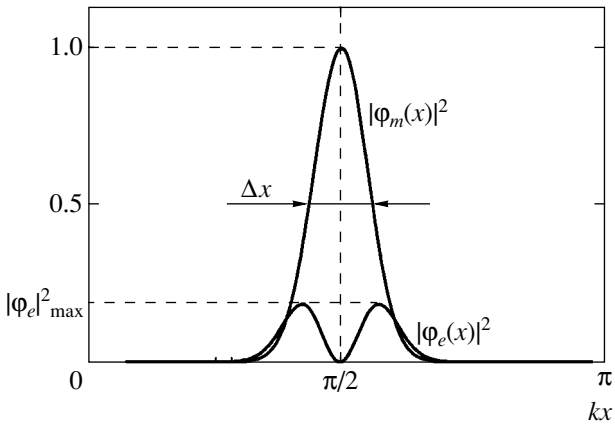

Fig. 2. Squares of the modulus of functions $\varphi_m(x, t)$ and $\varphi_e(x, t)$ defined by expressions (37) and (38).

$$t_0^* = \max(t_0, 1/\Gamma). \quad (44)$$

According to (41), we formally have $\Delta x(t) \longrightarrow 0$ for $t \longrightarrow \infty$. It is clear that, indeed, the narrowing of the wave packets should be restricted in some way. We will show below that the narrowing process becomes limited outside the region of applicability of the adiabatic approximation (see Section 7).

The total probabilities of finding an atom in the metastable and excited states at instant t are determined by the integrals over x from the probability density $dW_{m, e}(x, t)/dx$ (8), which can easily be calculated:

$$W_{\text{tot}}^{(m)}(t) = \exp\left(-\frac{t}{t_0}\right)I_0\left(\frac{t}{t_0}\right), \quad (45)$$

$$W_{\text{tot}}^{(e)}(t) = \frac{1}{\Gamma t_0}\exp\left(-\frac{t}{t_0}\right)\left\{I_0\left(\frac{t}{t_0}\right) - I_1\left(\frac{t}{t_0}\right)\right\}, \quad (46)$$

where I_0 and I_1 are the modified Bessel functions [11, 12]. In the limit $t \gg t_0$. Eqs. (45) and (46) give

$$W_{\text{tot}}^{(m)}(t) \approx \sqrt{\frac{t_0}{2\pi t}} = \frac{\sqrt{\Gamma}}{|\Omega|\sqrt{\pi t}}\left(1 + 4\frac{\Delta^2}{\Gamma^2}\right)^{1/2}, \quad (47)$$

$$W_{\text{tot}}^{(e)}(t) \approx \frac{\sqrt{t_0}}{\Gamma(2t)^{3/2}\sqrt{\pi}} = \frac{1}{2|\Omega|\sqrt{\pi\Gamma}t^{3/2}}\left(1 + 4\frac{\Delta^2}{\Gamma^2}\right)^{1/2}. \quad (48)$$

Therefore, under the conditions considered, a slow nonexponential decay of atoms occurs [13]. The physi-

cal reason for the slow decay is very simple. This is explained by the fact that the atom at the metastable E_m level does not decay directly but should first undergo the transition to an excited state, and only then does it spontaneously emit a photon during a transition to some other levels. The probability of such a process is determined by the product of the probabilities of excitation of the E_e level and its spontaneous decay. The excitation probability is determined by the Rabi frequency. However, even if the Rabi frequency Ω is large compared to the inverse interaction time $1/t$, the effective Rabi frequency $\Omega_{\text{eff}}(x) = \Omega\cos(kx)$ can locally be small for the atoms having an x coordinate (where $\cos(kx)$ is small, i.e., in the vicinity of the nodes of a standing wave). Therefore, the effective decay rate $\Gamma_{\text{eff}}(x)$, defined according to (37) and (38) as

$$\Gamma_{\text{eff}}(x) = -2\text{Im}V_{CY}(x) = \frac{1}{4}\frac{|\Omega|^2\Gamma}{\Delta^2 + \Gamma^2/4}\cos^2(kx), \quad (49)$$

can also be small. Hence, the slowing down of the decay compared to the law $\exp(-\Gamma t)$ is determined by the contribution from the atoms (more exactly, parts of the atomic wave function) for which $\Omega_{\text{eff}}(x)$ and $\Gamma_{\text{eff}}(x)$ are relatively small, i.e., for which the value of $\cos(kx)$ is small, which takes place near the nodes $x \approx \pi/2k$ of the standing wave.

The appearance of the nonexponential, power decay can be also explained quite simply. According to (37) and (38), the decay law for each given x is exponential. Therefore, if an unperturbed state of an atom were characterized by a nonspreading localized wave packet of a size much smaller than Δx, the decay would be purely exponential and the decay rate for this state would be equal to $\Gamma_{\text{eff}}(x)$. However, if the initial state of the atom is a plane wave or a broad wave packet (of a size much greater than Δx), then the decay of this state is determined as a whole by a superposition of exponentials with different decay rates. It is obvious that a sum of exponentials is not an exponential. Therefore, the decay dynamics of the atomic state is characterized by a power rather than an exponential law.

Expression (41) describing the atomic wave-packet width Δx during scattering can be also used to specify a general formal condition for the applicability of the adiabatic approximation (24). Assuming $\partial/\partial x \sim 1/\Delta x$, we can reduce condition (24) to the form

$$\frac{t}{t_0} \ll \frac{|\Delta + i\Gamma/2|}{\omega_r}, \quad \frac{1}{\sqrt{\omega_r t_0}}, \quad (50)$$

where the expressions in the right-hand side of inequalities are large because of a small recoil energy.

Note finally that, similarly to passage from Eqs. (27) to (35), we can also simplify Eq. (33) for the probability amplitudes $a_n^{(m)}(t)$ by omitting the second derivative with respect to time in (33) in the limit of large t. Then, using the replacement

$$a_n^{(m)}(t) = b_{n'}(t)i^{n'}\exp\left(-i\frac{|\Omega|^2}{8(\Delta + i\Gamma/2)}t\right),$$

where $n' = n/2$ and n is an even number, $n = 2n'$, we find from (33) that functions $b_{n'}(t)$ satisfy the well-known Raman–Nath equation [14]

$$\frac{d}{dt}b_{n'}(t) = -\frac{|\Omega|^2}{16(\Delta + i\Gamma/2)}(b_{n'-1} - b_{n'+1}), \quad (51)$$

which has the solution

$$b_{n'}(t) = J_{n'}\left(-\frac{|\Omega|^2 t}{8(\Delta + i\Gamma/2)}\right). \quad (52)$$

As a result, the nonzero probability amplitudes $a_{2n}^{(m)}(t)$ take the form (primes at n' are omitted)

$$a_{2n}^{(m)}(t) = (-i)^n\exp\left(-i\frac{|\Omega|^2}{8(\Delta + i\Gamma/2)}t\right) \times J_n\left(\frac{|\Omega|^2 t}{8(\Delta + i\Gamma/2)}\right). \quad (53)$$

Obviously, the same result can be obtained using the Fourier transform of the function $\varphi_m(x, t)$ (37). The probability amplitudes $a_{2n+1}^{(e)}(t)$ can be found similarly (for example, using the Fourier transform of the function $\varphi_e(x, t)$ (38)):

$$a_{2n+1}^{(e)}(t) = -\frac{\Omega^*}{4(\Delta + i\Gamma/2)}\exp\left(-i\frac{|\Omega|^2}{8(\Delta + i\Gamma/2)}t\right) \times (-i)^n\left[J_n\left(\frac{|\Omega|^2 t}{8(\Delta + i\Gamma/2)}\right) + iJ_{n+1}\left(\frac{|\Omega|^2 t}{8(\Delta + i\Gamma/2)}\right)\right]. \quad (54)$$

Sums of the squares of the modulus of functions $a_{2n}^{(m)}(t)$ (53) and $a_{2n+1}^{(e)}(t)$ (54) give the same expressions for the total scattering probabilities as (45) and (46).

Therefore, a general conclusion concerning the Raman–Nath approximation used in the problem under study is that this approximation is not identical and is more particular with respect to the adiabatic approximation. For the Raman–Nath approximation to be valid, it is necessary, along with fulfillment of the adiabatic approximation, to neglect the second derivatives of functions $\varphi_m(x, t)$ and $a_n^{(m)}(t)$ with respect to time in Eqs. (27) and (33), respectively. As follows from the above analysis, the latter approximation is valid only when the interaction time is long enough, $t > 1/\Gamma$ (43). We will show in the next section that for $t \lesssim 1/\Gamma$, the

expressions for $W_n^{(m)}(t)$ and $W_n^{(e)}(t)$ are much more complicated than (45) and (46).

6. QUASI-ENERGIES AND SUPPRESSION OF RABI OSCILLATIONS

Thus, the treatment based on the solution of simplified first-order equation (35) is quite satisfactory in the limit of sufficiently long interaction times ($t > 1/\Gamma$); however, it cannot be used to analyze the dynamics of the population of levels E_m and E_e in the range of shorter times t. This time range is of independent interest, and to study it, second-order equation (27) or the equivalent two-component equation (28) with matrix Hamiltonian (30) should be solved. To solve these equations with the same initial conditions as before, it is necessary to preliminarily find the eigenvalues and the eigenfunctions for these equations. In physical terms, the eigenvalues and eigenfunctions are the quasi-energies and quasi-energy wave functions of the system. Further, because Hamiltonian (30) does not contain, in the adiabatic approximation, derivatives with respect to x, the x coordinate is a parameter or a quantum number of the problem. Taking into account these remarks, we will write the expressions for quasi-energies $\gamma_\pm(x)$ defined as the eigenvalues of Hamiltonian (30):

$$\gamma_\pm(x) = -\frac{\Delta + i\Gamma/2}{2} \pm \frac{1}{2}\sqrt{(\Delta + i\Gamma/2)^2 + |\Omega|^2\cos^2(kx)}. \quad (55)$$

The quasi-energies can be conveniently and distinctly represented graphically as broadened levels with the boundaries

$$\{\mathrm{Re}\gamma_\pm(x) + \mathrm{Im}\gamma_\pm(x), \mathrm{Re}\gamma_\pm(x) - \mathrm{Im}\gamma_\pm(x)\}.$$

The distance between the upper and lower boundaries of each band is equal to the width of the corresponding quasi-energy level $\Gamma_\pm(x) = -2\mathrm{Im}\gamma_\pm(x)$, while the "centers of gravity" $\mathrm{Re}_\pm(x)$ of the bands characterize the shift of the levels in the field with respect to their position $\Omega = 0$ in the absence of field. A typical structure of quasi-energy bands is shown in Fig. 3.

As follows from Eq. (55), one of the quasi-energies, $\gamma_+(x)$, is transformed to the potential $V_{CY}(x)$ (36) when

$$|\Omega\cos(kx)| \ll \sqrt{\Delta^2 + \frac{\Gamma^2}{4}}.$$

In this case, the approximate equality $\gamma_+(x) \approx V_{CY}(x)$ is valid over the entire range of variation of x only if the Rabi frequency is small, $|\Omega| \ll \sqrt{\Delta^2 + \Gamma^2/4}$. If, however, the Rabi frequency is large, $|\Omega| \gg \sqrt{\Delta^2 + \Gamma^2/4}$, then we have $\gamma_+(x) \approx V_{CY}(x)$ only in a relatively small vicinity of a node of the standing wave, where the effective Rabi frequency $\Omega_{\mathrm{eff}} = \Omega\cos(kx)$ is small. In this case the approximation of $\gamma_+(x)$ by the potential $V_{CY}(x)$ is satisfactory for describing the behavior of the system at long interaction times when the dynamic evolution of atomic populations is determined by the region of the most long-lived quasi-energy states. At shorter times, the approximation $\gamma_+(x) \approx V_{CY}(x)$ is incorrect, and the problem should be solved taking into account all the quasi-energy states and correct expressions for quasi-energies and quasi-energy wave functions. The latter are defined as the eigenfunctions of Hamiltonian (30) and, as can be easily verified, have the form

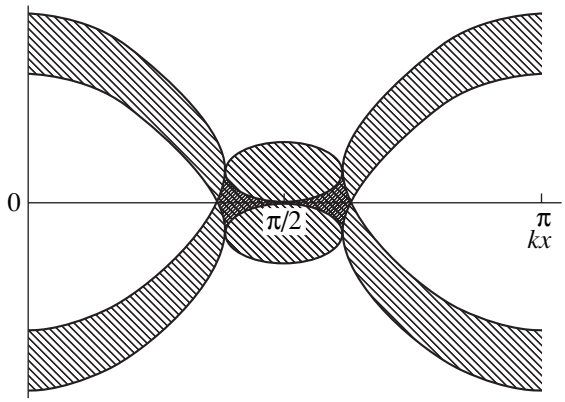

Fig. 3. Quasi-energy bands of an atom in the field of a standing light wave for $\Delta = 0$ and $|\Omega| = \sqrt{2}\Gamma$.

$$\Phi_{x',\pm}(x) = \frac{1}{[\gamma_\pm(x')(2\gamma_\pm(x') + \Delta + i\Gamma/2)]^{1/2}} \times \begin{pmatrix} -\frac{1}{2}\Omega\cos(kx') \\ \gamma_\pm(x') \end{pmatrix} \delta(x - x'), \quad (56)$$

where x' is a quantum number and x is a dynamic variable.

Note that, because Hamiltonian H_{ad} (30) is non-Hermitian, its eigenfunctions (56) are not orthogonal and do not form a complete system of functions. To formulate the conditions of orthogonality and completeness, it is necessary to additionally introduce a system of eigenfunctions the Hermitian-conjugated Hamiltonian $H_{ad}^\dagger$. The eigenvalues of $H_{ad}^\dagger$ are equal to $\gamma_\pm^*(x)$, and the corresponding eigenfunctions $\tilde{\Phi}_{x',\pm}(x)$ differ from

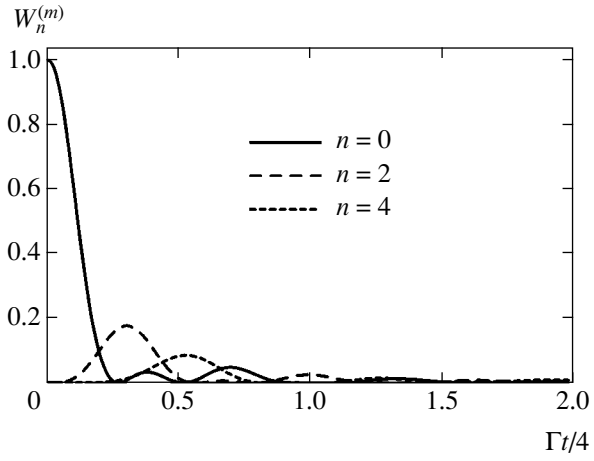

Fig. 4. Time-dependent partial probabilities $W_n^{(m)}(t)$ of scattering into different diffraction maxima of atoms in a metastable state. $|\Omega|/\Gamma = 5$, $\Delta = 0$.

$\Phi_{x',\pm}(x)$ (56) only by the replacement of Γ by $-\Gamma$. For this reason, the complex conjugate functions $\tilde{\Phi}^*_{x',\pm}(x)$ differ from $\Phi_{x',\pm}(x)$ (56) only by the replacement of Ω by Ω^*. The functions $\Phi_{x',\pm}(x)$ and $\tilde{\Phi}_{x',\pm}(x)$ form a system of biorthogonal functions [15], and the conditions of their orthogonality and completeness are written in the form

$$\int_0^{\pi/k} dx \tilde{\Phi}^*_{x',\mp}(x)\Phi_{x'',\pm}(x) = 0,$$

$$\int_0^{\pi/k} dx \tilde{\Phi}^*_{x',\pm}(x)\Phi_{x'',\pm}(x) = \delta(x'-x''), \tag{57}$$

$$\int_0^{\pi/k} dx' \sum_{\pm} \Phi_{x',\pm}(x_1)\tilde{\Phi}^*_{x',\pm}(x_2) = \delta(x_1-x_2). \tag{58}$$

Completeness condition (58) makes it possible at once to write the solution to the initial problem:

$$\Phi(x,t) = \int_0^{\pi/k} dx' \int_0^{\pi/k} dx'' \exp\{-it\gamma_\pm(x')\} \times \sum_{\pm} \Phi_{x',\pm}(x)\tilde{\Phi}^*_{x',\pm}(x'')\Phi(x'',0). \tag{59}$$

For the initial conditions (6) under study, we have $\Phi(x'',0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and, finally, Eq. (59) gives

$$\varphi_m(x,t) = \sum_{\pm} \mp \frac{\gamma_\mp(x)\exp\{-i\gamma_\pm(x)t\}}{\sqrt{(\Delta + i\Gamma/2)^2 + |\Omega|^2\cos^2(kx)}}, \tag{60}$$

$$\varphi_e(x,t) = \sum_{\pm} \mp \frac{\Omega^*\cos(kx)\exp\{-i\gamma_\pm(x)t\}}{2\sqrt{(\Delta + i\Gamma/2)^2 + |\Omega|^2\cos^2(kx)}}, \tag{61}$$

$$W_{2n}^{(m)}(t) = \exp\left(-\frac{\Gamma t}{2}\right)\left| J_{2n}\left(\frac{|\Omega|t}{2}\right) - it\frac{\Delta + i\Gamma/2}{2} \times \int_0^1 dz J_{2n}\left(\frac{|\Omega|t}{2}z\right)\left[J_0\left(t\frac{\Delta + i\Gamma/2}{2}\sqrt{1-z^2}\right) - iJ_1\left(t\frac{\Delta + i\Gamma/2}{2}\sqrt{1-z^2}\right)\frac{1}{\sqrt{1-z^2}}\right]\right|^2, \tag{62}$$

$$W_{2n+1}^{(e)} = \left(\frac{|\Omega|t}{4}\right)^2 \exp\left(-\frac{\Gamma t}{2}\right) \times \left|\int_0^1 dz J_0\left(t\frac{\Delta + i\Gamma/2}{2}\sqrt{1-z^2}\right) \times \left[J_{2n+2}\left(\frac{|\Omega|t}{2}z\right) - J_{2n}\left(\frac{|\Omega|t}{2}z\right)\right]\right|^2, \tag{63}$$

where J_n is the Bessel function [11, 12] (see the Appendix). The dependences of functions $W_{2n}^{(m)}(t)$ (62) on interaction time t are shown in Fig. 4.

The partial probabilities $W_n^{(m)}(t)$ summed over n determine the interaction, time-dependent, total residual probability $W_{\text{tot}}^{(m)}(t)$ of finding an atom in a metastable state. The dependence $W_{\text{tot}}^{(m)}(t)$ is shown in Fig. 5. One can easily see that, for $\Gamma t \lesssim 1$, the function $W_{\text{tot}}^{(m)}(t)$ oscillates at a frequency equal to the Rabi frequency. For a purely two-level system in a homogeneous resonance electric field with the strength amplitude $2\mathbf{E}_0$, the dependence of the lower-level population $W^{(m)}(t)$ is shown by the dashed curve. One can see from Fig. 5 that the Rabi oscillations are strongly suppressed in the case of scattering atoms compared to the oscillations of the population for a purely two-level system [13]. This is explained by the fact that the oscillations of partial probabilities are not harmonic and the positions of minima and maxima of the functions $W_{2n}^{(m)}(t)$ (62) depend

on the number n (Fig. 4). This leads to the suppression and smoothing of oscillations upon summation over n [see (14)]. One can see from Fig. 5 that the degree of smoothing of Rabi oscillations increases with interaction time t, and for $\Gamma t \gg 1$, oscillations are absent altogether. In addition, the curves in Fig. 5 clearly demonstrate that, according to the results of analysis performed in the previous section, the decay of the population $W_{\text{tot}}^{(m)}(t)$ noticeably slows down in the limit of large t compared to the purely exponential decay in a two-level system in a homogeneous resonance field (dashed curve). The dependences $W_{2n}^{(m)}(t)$ calculated numerically agree well with analytic asymptotic expressions (47).

The partial, $W_{2n+1}^{(e)}(t)$, and total, $W_{\text{tot}}^{(e)}(t)$, probabilities of finding an atom in the excited state are described by curves similar to those presented in Figs. 4 and 5.

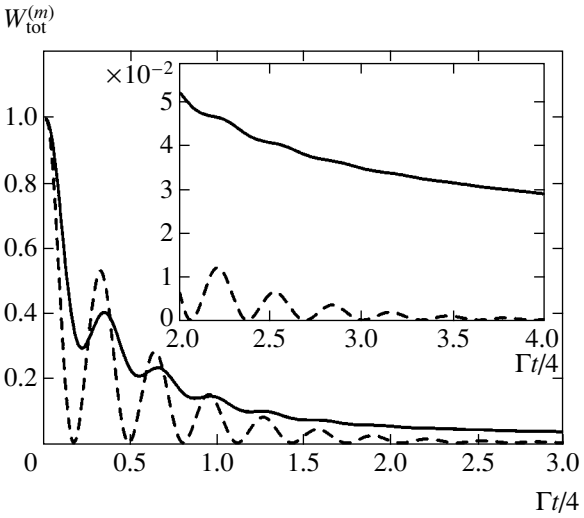

Fig. 5. Dependence of the total population $W_{\text{tot}}^{(m)}(t)$ of a metastable level on the interaction time t. $|\Omega|/\Gamma = 5$, $\Delta = 0$. The dashed curve is the population of the lower level of a two-level system in a homogeneous resonance field.

7. BEYOND THE SCOPE OF THE ADIABATIC APPROXIMATION

7.1. Basic Approximations and Equations

According to the conclusion made in Section 4, the kinetic energy operator $(-1/2M)\partial^2/\partial x^2$ is omitted in the adiabatic approximation both in Eq. (4) for $\varphi_m(x, t)$ and Eq. (5) for $\varphi_e(x, t)$. A general condition for the applicability of these simplifications is written in the form (24). On the other hand, Eqs. (4) and (5) are not symmetrical: Eq. (5) contains a large term proportional to $|\Delta + i\Gamma/2|$, whereas Eq. (4) does not contain such a term. In addition, when the interaction time is long, $t|\Delta + i\Gamma/2| \gg 1$, two parameters in the right-hand side of Eq. (24) are of substantially different orders of magnitude. For this reason, a range of parameters exists where the kinetic energy is small compared to $|\Delta + i\Gamma/2|$ but not small compared to $1/t$,

$$\frac{1}{t} \lesssim \left\langle \frac{1}{2M}\frac{\partial^2}{\partial x^2} \right\rangle \ll \left|\Delta + i\frac{\Gamma}{2}\right|. \tag{64}$$

Under these conditions, it is reasonable to use in the formulation given in Section 4 an alternative approximation, in which the kinetic energy operator is omitted in Eq. (5) but retained in Eq. (4). Further, since, according the analysis performed in Section 4, the main contribution to the long-lived states of atoms comes from the region near the nodes of a standing wave, where $|\cos(kx)| \ll 1$, we will approximate the cosine by a linear function, $\cos(kx) \approx -\xi$, where $\xi = kx - \pi/2$. Finally, as before, we will only consider the case of normal incidence, $p_{0x} = 0$. In this case, Eqs. (4) and (5) take the form

$$i\frac{\partial}{\partial t}\varphi_m(\xi, t) = -\omega_r\frac{\partial^2}{\partial \xi^2}\varphi_m(\xi, t) + \frac{1}{2}\Omega\xi\varphi_e(\xi, t) \tag{65}$$

and

$$\left\{i\frac{\partial}{\partial t} + \exp(i\alpha)\left(\Delta^2 + \frac{\Gamma^2}{4}\right)^{1/2}\right\}\varphi_e(\xi, t) = \frac{1}{2}\Omega^*\xi\varphi_m(\xi, t), \tag{66}$$

where

$$\alpha = \frac{\pi}{2} - \arctan\frac{2\Delta}{\Gamma}, \quad 0 \le \alpha < \pi. \tag{67}$$

Similarly to passage from Eqs. (25), (26) to Eq. (27), Eqs. (65) and (66) can be replaced by one equivalent higher order equation:

$$\left\{i\frac{\partial}{\partial t} + \exp(i\alpha)\left(\Delta^2 + \frac{\Gamma^2}{4}\right)^{1/2}\right\}i\frac{\partial}{\partial t}\varphi_m = \frac{1}{4}|\Omega|^2\xi^2\varphi_m - \left\{i\frac{\partial}{\partial t} + \exp(i\alpha)\left(\Delta^2 + \frac{\Gamma^2}{4}\right)^{1/2}\right\}\omega_r\frac{\partial^2}{\partial \xi^2}\varphi_m. \tag{68}$$

Finally, similarly to passage from Eqs. (27) to (35), assuming that $\partial/\partial t \sim 1/t$ and taking into account the

smallness of $1/t$ compared to $|\Delta + \Gamma/2|$ (64), we omit the corresponding small terms in Eq. (68) and reduce it to the equation for the wave function of a quantum harmonic oscillator with a complex potential:

$$i\frac{\partial}{\partial t}\varphi_m(\xi, t) = \left\{-\omega_r\frac{\partial^2}{\partial\xi^2} + \frac{|\Omega|^2\exp(-i\alpha)}{4\sqrt{\Delta^2 + \Gamma^2/4}}\xi^2\right\}\varphi_m(\xi, t) \quad (69)$$

for

$$\varphi_e(\xi, t) \approx \frac{1}{2}\frac{\Omega^*\exp(-i\alpha)}{\sqrt{\Delta^2 + \Gamma^2/4}}\xi\varphi_m(\xi, t). \quad (70)$$

Equation (69) has stationary or quasi-energy solutions:

$$\varphi_m(\xi, t) = \exp(-i\gamma t)\varphi_m(\xi).$$

In the range $|\alpha| < \pi$, the stationary solutions vanishing for $|\xi| \longrightarrow \infty$ are usual wave functions of a harmonic oscillator, but with a complex argument

$$\varphi_n^{(m)}(\xi) = \left(\frac{d\eta/d\xi}{2^n n!\sqrt{\pi}}\right)^{1/2}\exp\left(-\frac{\eta^2}{2}\right)H_n(\eta), \quad (71)$$

where $n = 0,1, 2, \ldots$, H_n are Hermitian polynomials and

$$\eta = \left(\frac{|\Omega|^2}{4\omega_r\sqrt{\Delta^2 + \Gamma^2/4}}\right)^{1/4}\exp\left(-i\frac{\alpha}{4}\right)\xi. \quad (72)$$

The eigenvalues of Eq. (69), i.e., the complex quasi-energies γ_n, are

$$\gamma_n = (2n + 1)\gamma_0, \quad (73)$$

where γ_0 is the complex energy of the ground state of the oscillator with the complex potential

$$\gamma_0 = \left(\frac{|\Omega|^2\omega_r}{4\sqrt{\Delta^2 + \Gamma^2/4}}\right)^{1/2}\exp\left(-i\frac{\alpha}{2}\right). \quad (74)$$

The functions $\varphi_n^{(e)}(\xi)$, corresponding to the found functions $\varphi_n^{(m)}(\xi)$ (71), are determined from Eq. (70).

Unlike the adiabatic approximation (Section 4), the eigenvalue equation in the case of condition (64) is the equation for one-component functions $\varphi_m(\xi)$ (69), for which the conditions of normalization and completeness should be formulated. As before, the system of functions $\varphi_n^{(m)}(\xi)$ (71) belongs to the class of biorthogonal functions. The eigenvalues of the Hermitian-conjugated Hamiltonian are γ_n^*, while the eigenfunctions $\tilde{\varphi}_n^{(m)}(\xi)$ differ from $\varphi_n^{(m)}(\xi)$ only by the replacement of α by $-\alpha$ (67). This means that functions $[\tilde{\varphi}_n^{(m)}(\xi)]^*$ coincide with $\varphi_n^{(m)}(\xi)$. The conditions of normalization and completeness are determined by equations of type (57) and (58) and have the form

$$\int d\xi[\tilde{\varphi}_n^{(m)}(\xi)]^*\varphi_{n'}^{(m)}(\xi) \equiv \int d\xi\varphi_n^{(m)}(\xi)\varphi_{n'}^{(m)}(\xi) = \delta_{n, n'}, \quad (75)$$

$$\sum_{n=0}^{\infty}\varphi_n^{(m)}(\xi)[\tilde{\varphi}_n^{(m)}(\xi')]^* \equiv \sum_{n=0}^{\infty}\varphi_n^{(m)}(\xi)\varphi_n^{(m)}(\xi') = \delta(\xi - \xi'). \quad (76)$$

7.2. Solution of the Initial Problem

Using the completeness condition (76) for the initial condition $\varphi_m(\xi, t = 0) \equiv 1$ (6), the solution of the initial problem is written in the form

$$\varphi_m(\xi, t) = \sum_{n=0}^{\infty}\exp(-i\gamma_n t)\varphi_n^{(m)}(\xi)\int_{-\infty}^{+\infty}d\xi'\varphi_n^{(m)}(\xi'). \quad (77)$$

The sum over n can be calculated explicitly using the Möller formula [11] as

$$\varphi_m(\xi, t) = \left\{\cos\left[\tau\exp\left(-i\frac{\alpha}{2}\right)\right]\right\}^{1/2} \times \exp\left\{-i\frac{\xi^2}{2}\left(\frac{|\Omega|^2\sin\alpha}{2\omega_r\Gamma}\right)^{1/2}\exp\left(-i\frac{\alpha}{2}\right)\tan\left[\tau\exp\left(-i\frac{\alpha}{2}\right)\right]\right\}, \quad (78)$$

where

$$\tau = \sqrt{2}|\Omega|t\sqrt{\frac{\omega_r}{\Gamma}\sin\alpha} = \frac{|\Omega|t\sqrt{\omega_r}}{(\Delta^2 + \Gamma^2/4)^{1/4}} \equiv \frac{t}{t_1} \quad (79)$$

is the dimensionless time and t_1 is a new characteristic value of the interaction time:

$$t_1 = \left(\frac{\Gamma}{2|\Omega|^2\omega_r\sin\alpha}\right)^{1/2} = \frac{(\Delta^2 + \Gamma^2/4)^{1/4}}{|\Omega|\sqrt{\omega_r}}. \quad (80)$$

Because the recoil frequency ω_r is small, as a rule, $t_1 > t_0^*$ (44), although the interval between t_0^* and $t = t_1$ cannot be too large. It is interesting that when the detuning is very large, $|\Delta| \gg \Gamma$, the value of the parameter

$$t_0^* = t_0 \approx \frac{8\Delta^2}{\Gamma|\Omega|^2}$$

can be comparable to

$$t_1 \approx \frac{\sqrt{|\Delta|}}{|\Omega|\sqrt{\omega_r}}$$

(for $|\Delta| \approx (\Gamma|\Omega|)^{2/3}/4\omega_r^{1/3}$).

The probability density (7) determined by $\varphi_m(\xi, t)$ (78) is

$$\frac{dW_m(\xi, t)}{d\xi} = \frac{1}{\pi}|\varphi_m(\xi, t)|^2 = \frac{1}{\pi\sqrt{D}}\exp\left(-\frac{\xi^2}{\Delta\xi^2(t)}\right), \quad (81)$$

where

$$D = \cos^2\left(\tau\cos\frac{\alpha}{2}\right)\cosh^2\left(\tau\sin\frac{\alpha}{2}\right) + \sin^2\left(\tau\cos\frac{\alpha}{2}\right)\sinh^2\left(\tau\sin\frac{\alpha}{2}\right); \quad (82)$$

$\Delta\xi(t)$ is the time-dependent width of the distribution $dW_m(\xi, t)/d\xi$,

$$\Delta\xi^2 = 2D\left(\frac{2\omega_r\Gamma}{|\Omega|^2\sin\alpha}\right)^{1/2} \times\left\{\sin\frac{\alpha}{2}\sin\left(2\tau\cos\frac{\alpha}{2}\right) + \cos\frac{\alpha}{2}\sinh\left(2\tau\sin\frac{\alpha}{2}\right)\right\}^{-1}. \quad (83)$$

The time-dependent total probability $W_{\text{tot}}^{(m)}(t)$ (8) of finding an atom in the metastable state is obtained from (81) by integrating over ξ:

$$W_{\text{tot}}^{(m)}(t) = \sqrt{\frac{2}{\pi}}\left(\frac{2\omega_r\Gamma}{|\Omega|^2\sin\alpha}\right)^{1/4} \times\left\{\sin\frac{\alpha}{2}\sin\left(2\tau\cos\frac{\alpha}{2}\right) + \cos\frac{\alpha}{2}\sinh\left(2\tau\sin\frac{\alpha}{2}\right)\right\}^{-1/2}. \quad (84)$$

According to Eq. (70), the distribution of the probability density of finding atoms at the excited level is related to $dW_m(\xi, t)/d\xi$ (81) by the expression

$$\frac{dW_e(\xi, t)}{d\xi} = \frac{|\Omega|^2}{\Delta^2 + \Gamma^2/4}\frac{\xi^2}{4}\frac{dW_m(\xi, t)}{d\xi}, \quad (85)$$

which, after integration with respect to ξ, gives

$$W_{\text{tot}}^{(e)}(t) = \frac{|\Omega|^2}{\Delta^2 + \Gamma^2/4}\frac{\Delta\xi^2}{8}W_{\text{tot}}^{(m)}(t). \quad (86)$$

7.3. Limiting Cases and Discussion of Results

7.3.1 Adiabatic limit. In the limiting case of small τ, $\tau \ll 1$, general formulas (81)–(86) are transformed to the expressions for distributions of the probability density, the width of the distributions (41) and total probabilities (45) and (46) found earlier. Therefore, according to definitions (79) and (80), the region of existence of a power decay described in the adiabatic approximation is restricted by the condition $t_0^* < t < t_1$. According to the above discussion, t_0^* can become greater than t_1 at large detunings, resulting in the disappearance of the region of existence of a power decay.

To compare in more detail the results obtained in the adiabatic approximation in Sections 4–6 with the results of this section, it is interesting to compare the corresponding quasi-energy functions and quasi-energies. It is clear that in passing from the quasi-energy problem to the initial problem, the quasi-energy γ is replaced by the time derivative $\partial/\partial t$, which is on the order of $1/t$, i.e., $|\gamma| \sim 1/t$. It follows from this relation and Eqs. (73), (74) that the characteristic values of n contributing to the solution of the initial problem (77) are on the order of $n_{\text{eff}} \sim 1/\tau$. In the adiabatic limit, for $\tau \ll 1$, we have $n_{\text{eff}} \gg 1$. This means that a great number of terms with $n \sim n_{\text{eff}} \gg 1$ make a substantial contribution to the sum over n in (77), which makes it possible to replace the summation over n in (77) by integration. Further, the wave functions of a harmonic oscillator with the complex arguments of type (71) possess quite interesting properties and can noticeably differ from the wave functions of an oscillator with a real potential. Figure 6 shows an example of the function $\left|\varphi_{64}^{(m)}(\xi)\right|$ (71) (for $\Delta = 0$), which has one distinct maximum. Analysis of such curves for arbitrary n shows that the position of their maxima is determined by the condition $|\eta|_0(n) = 2^{1/4}\sqrt{n}$ or

$$\xi_0(n) = 2^{3/4}\sqrt{n}\frac{(\Delta^2 + \Gamma^2/4)^{1/8}\omega_r^{1/4}}{|\Omega|^{1/2}}. \quad (87)$$

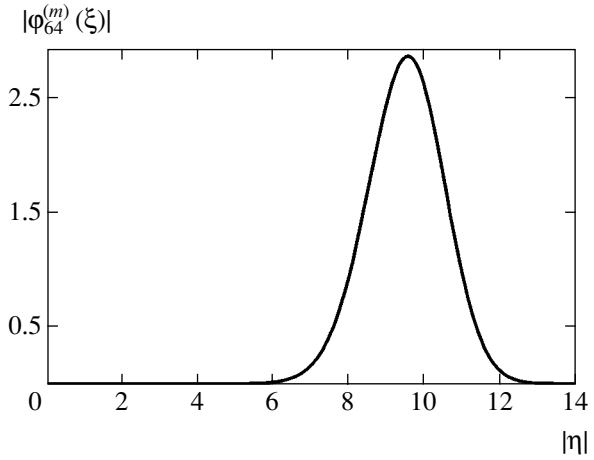

Fig. 6. Function $\left|\varphi_{64}^{(m)}(\xi)\right|$ (71) in units $10^{11}|\Omega|^{1/4}/(2\omega_r\Gamma)^{1/8}$.

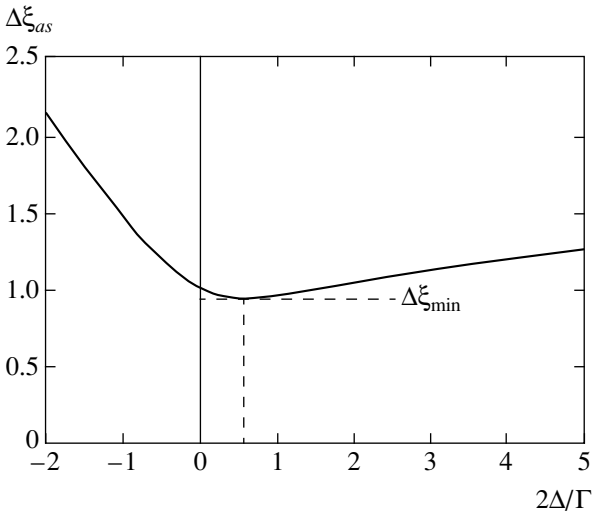

Fig. 7. Asymptotic width $\Delta\xi_{as}$ (89) of the wave packet of atoms in a metastable state as a function of the resonance detuning Δ; $\Delta\xi_{as}$ is expressed in units $(4\omega_r\Gamma)^{1/4}|\Omega|^{-1/2}$ Δ—in units $\Gamma/2$.

By using ξ_0 (87) instead of n as a new quantum number, we reduce the sum over n in Eq. (77), transformed to the integral with respect to dn, to the integral with respect to $d\xi_0$, in accordance with Eq. (59) obtained in the adiabatic approximation. The power decay law appearing in this case is explained by the fact that, although each term in the sum over n in (77) depends on exponentially time, their sum (or integral) is not an exponential.

In terms of ξ_0 (87), for $n \gg 1$, the expressions for quasi-energies (73) and (74) take the form

$$\gamma(\xi_0) = \frac{\xi_0^2}{2\sqrt{2}}\frac{|\Omega|^2}{\sqrt{\Delta^2+\Gamma^2/4}}\exp\left(-i\frac{\alpha}{2}\right). \qquad (88)$$

For $\Delta = 0$, Eq. (88) gives

$$\gamma(\xi_0) = (1+i)\gamma_+(x) = (1+i)V_{CY}(x),$$

where $\gamma_+(x)$ and $V_{CY}(x)$ are described by Eqs. (55) and (36), $kx \equiv \xi_0$, and $|\xi_0| \ll 1$. This result shows that, in the limit $\tau \ll 1$ for $\Delta = 0$, the imaginary part of the quasi-energy is transformed to expressions obtained in the adiabatic approximation. However, there also exists the real part $\gamma(\xi_0)$, equal to the modulus of the imaginary part, which does not appear in the adiabatic approximation. Therefore, there is no, in a certain sense, complete passage to the adiabatic limit at $\tau \longrightarrow 0$ or $\omega_r \longrightarrow 0$, and point $\omega_r = 0$ for a quantum harmonic oscillator with a complex potential [Eq. (69)] is an essential singularity.

7.3.2. Nonadiabatic scattering. As follows from the above discussion, nonadiabatic scattering occurs in the case opposite to that considered above, when $\tau \gg 1$. This is the asymptotics of a long interaction time. According to the estimate $n_{\text{eff}} \sim 1/\tau$ obtained earlier, when $\tau \gg 1$, only the term with $n = 0$ makes the main contribution to the sum over n in (77). This is explained by the fact that in the limit $\tau \gg 1$, all terms with $n \geq 1$ become exponentially small compared to the term with $n = 0$. Therefore, for $\tau \gg 1$, the width of the distribution $dW_m(\xi, t)/d\xi$ tends to its asymptotic value $\Delta\xi_{as}$, which is determined by the width of the region of localization of the wave function of the ground state of the oscillator with the complex potential $\varphi_0^{(m)}(\xi)$ (71), which, in turn, is determined by the condition $\mathrm{Re}(\eta^2) \approx 1$, giving

$$\Delta\xi_{as}^2 = \left(\frac{2\omega_r\Gamma}{|\Omega|^2\sin(\alpha)}\right)^{1/2}\frac{1}{\cos(\alpha/2)} = \frac{2}{|\Omega|}\left[\frac{2\omega_r(\Delta^2+\Gamma^2/4)}{\Delta+\sqrt{\Delta^2+\Gamma^2/4}}\right]^{1/2}. \qquad (89)$$

Of course, the same result follows from general expression (83) in the limit of large τ. The width $\Delta\xi_{as}$ (89) is independent of time t. Therefore, the wave packet of atoms in the metastable state narrows down with increasing interacting time t until the power decay begins ($t < t_1$), and the wave packet ceases to narrow down for $t \gg t_1$. The asymptotic behavior of the wave packet width at long interaction times (see below) and the asymptotic value of the width (89) depend on the value and sign of the resonance detuning Δ. The dependence of the asymptotic width $\Delta\xi_{as}$ (89) on the resonance detuning Δ is shown in Fig. 7. One can easily see that this dependence is asymmetric with respect to the sign of the detuning. For large negative detunings Δ, the

asymptotic width of the distribution $\Delta\xi$ rapidly increases,

$$\Delta\xi_{as} \approx \frac{2\sqrt{2}\omega_r^{1/4}|\Delta|^{3/4}}{(\Gamma|\Omega|)^{1/2}}.$$

The condition of smallness of the asymptotic width of the wave packet compared to the wavelength of light, $\Delta\xi_{as} \ll 1$, restricts in this case the value of detuning,

$$|\Delta| \ll \frac{(\Gamma|\Omega|)^{2/3}}{\omega_r^{1/3}}.$$

This restriction coincides with the above-discussed condition of the existence of the nonzero region of adiabatic nonexponential decay $t_0 \ll t_1$. The value of the asymptotic width $\Delta\xi_{as}$ (89) of the wave packet minimized by the resonance detuning Δ is achieved for $2\Delta/\Gamma = 1/\sqrt{3}$ and is equal to

$$\Delta\xi_{\min} = \left(\frac{16}{3\sqrt{3}}\frac{\omega_r\Gamma}{|\Omega|^2}\right)^{1/4} \approx 1.32\left(\frac{\omega_r\Gamma}{|\Omega|^2}\right)^{1/4}. \qquad (90)$$

Using the minimal width $\Delta\xi_{\min}$ (90) of the wave packet, we can specify the restriction on the value of the Rabi frequency $|\Omega|$ at which the condition of smallness of the kinetic energy of atoms compared to $|\Delta + i\Gamma/2|$ (64) is fulfilled. Assuming that $\partial/\partial x \sim k/\Delta\xi_{\min}$ and $|\Delta| \sim \Gamma$, we find the restriction on $|\Omega|$:

$$|\Omega| \ll \Gamma\sqrt{\frac{\Gamma}{\omega_r}}, \qquad (91)$$

which follows from (64). Because $\Gamma \gg \omega_r$ (usually $\Gamma \sim 10^3\omega_r$), restriction (91) does not contradict the assumption on the realization of diffraction rather than Bragg scattering, $|\Omega| \geq \sqrt{\Gamma\omega_r}$.

As for the asymptotic behavior of the packet width $\Delta\xi(t)$ tending to its asymptotic value (89), it is determined in the general case by Eq. (83). Figure 8 shows the dependences $\Delta\xi(t)$ for different values and signs of the resonance detuning Δ. One can see that the specific feature of the case of positive detunings $\Delta > 0$ is that the width $\Delta\xi(t)$ of the atomic wave packet tends to its asymptotic value (89), exhibiting distinct oscillations (curve 1 in Fig. 8). The oscillations appear due to the presence of terms with different n in superposition (77). The oscillation frequency is equal to the eigenfrequency of an oscillator with the complex potential $\mathrm{Re}\gamma_0$, where γ_0 is determined by Eq. (74). The specific

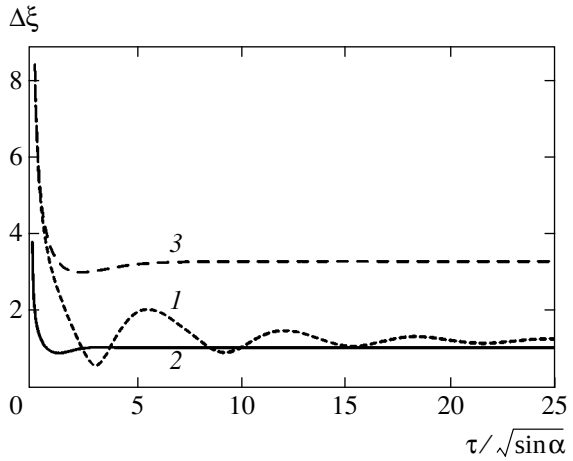

Fig. 8. Width $\Delta\xi$ (83) of an atomic wave packet as a function of interaction time t for $2\Delta/\Gamma = 2 + \sqrt{3}$ (*1*), $\Delta = 0$ (*2*), $2\Delta/\Gamma = -2 - \sqrt{3}$ (*3*); $\Delta\xi$ is expressed in units of $(8\omega_r\Gamma)^{1/4}|\Omega|^{-1/2}$; the scale of t is independent of detuning Δ.

feature of the case of sufficiently large positive detunings is determined by the fact that we have

$$\frac{\mathrm{Re}\gamma_n}{|\mathrm{Im}\gamma_n|} \approx \frac{\cot(\alpha/2)}{n + 1/2} \approx \frac{4\Delta}{\Gamma(n + 1/2)} \gg 1$$

(for not very large values of n). For this reason, both superposition (77) and the width of the corresponding distribution over ξ have time to perform a few oscillations for the time ($\sim|\mathrm{Im}\gamma_0|^{-1}$) when the total probability $W_{\mathrm{tot}}^{(m)}(t)$ takes its asymptotic value. Because only the term with $n = 0$ remains in the superposition of quasienergy wave functions (77) in the long-time interaction asymptotics $\tau \gg 1$, the asymptotic dependence of $W_{\mathrm{tot}}^{(m)}(t)$ is determined by the factor

$$\exp(-2\mathrm{Im}\gamma_0 t) \equiv \exp(-\tilde{\Gamma}t).$$

This means that, in the nonadiabatic regime (when $t \gg t_1$), the power decay occurring in the interval $t_0^* \lesssim t \lesssim t_1$ passes to the exponential decay with the modified decay constant

$$\tilde{\Gamma} = 2\mathrm{Im}\gamma_0 = |\Omega|\left(\frac{\omega_r}{2}\frac{\sqrt{\Delta^2 + \Gamma^2/4} - \Delta}{\Delta^2 + \Gamma^2/4}\right)^{1/2}, \qquad (92)$$

which differs substantially from the decay constant Γ of a free atom.

Because, as was shown, the wave packets of atoms in the metastable and excited states do not narrow down in the region $t > t_1$ and the absolute value of populations in these states decreases, it is useless in practice to

increase the interaction time above t_1. Let us present some estimates. For $t = d/v_0$, where d and v_0 are the diameter of a laser beam and the velocity of atoms, respectively; $v_0 \sim 10^5$ cm s^{-1}; $|\Omega| = 10\Gamma$; $\omega_r = 10^{-3}\Gamma$; and $\Gamma \sim 10^8$ s^{-1}, condition $t = t_1$ and Eqs. (80) and (90) give $d \sim 3 \times 10^{-3}$ cm and $\Delta x_{\min} \sim 10^{-6}$ cm.

8. CONCLUSIONS

In this paper we have studied the dynamics of spontaneous radiation of atoms scattered by a resonance standing wave under the conditions when (i) diffraction rather than Bragg scattering occurs; (ii) the lower of the two resonance levels of an atom is narrow (metastable), while the upper level is broad; (iii) the width of the upper (excited) level is mainly determined by spontaneous radiative transitions to the levels other than the metastable level; and (iv) the Rabi frequency Ω, the decay rate Γ of the excited level, and the recoil frequency ω_r satisfy the conditions $|\Omega| \geq \Gamma \gg \omega_r$.

We have shown that in this case, substantially different regimes of adiabatic and nonadiabatic scattering can be realized. The conditions of the existence of these regimes are $t < t_1$ and $t > t_1$, respectively, where t is the interaction time and t_1 is determined from Eq. (80).

We have found that, in the adiabatic regime when the recoil of an atom can be completely neglected, the metastable and excited levels decay according to a power law rather than an exponential one. We have also shown that, at small times, $t < 1/\Gamma < t_1$, the populations of working levels oscillate at the Rabi frequency $|\Omega|$, their oscillations being strongly suppressed compared to the case of an atom in a homogeneous resonance field.

In addition, in this regime, the spatial region of localization of the wave packets of atoms remaining in the metastable and excited states narrows down, the appearing narrow wave packets being localized in the vicinity of the nodes of a standing wave.

We also have shown that the narrowing down of the wave packets ceases in the nonadiabatic regime and their width achieved its minimal value determined by Eq. (89). In this case, the power decay of atomic levels transforms to the exponential decay with the decay constant $\tilde{\Gamma}$ (92), which is lower than the decay rate Γ of a free atom, decreases with increasing Γ, and depends both on the value and sign of the resonance detuning Δ and on the Rabi frequency $|\Omega|$. We believe that the effects discovered and described in this paper are quite interesting and can be observed experimentally.

ACKNOWLEDGMENTS

This work was partially supported by the Russian Foundation for Basic Research (project nos. 01-02-17643, 02-02-16400, and 03-02-06145), the Alexander Humboldt Foundation (M.V.F.), and the DFG German Research Center (W.P.Shch.).

APPENDIX

Derivation of Expressions (62) and (63)

Consider in more detail the derivation of expressions (62) and (63) for partial probabilities $W_n^{(m,e)}(t)$. This can be most simply demonstrated for the function $W_n^{(e)}(t)$. For this purpose, we represent the function $\varphi_e(x, t)$ (61) in the form

$$\varphi_e(x, t) = i\Omega^* \cos(kx) \exp\left\{ i\frac{\Delta + i\Gamma/2}{2} t \right\} \times \frac{\sin\left(\frac{t}{2} \sqrt{\left(\Delta + \frac{i\Gamma}{2} \right)^2 + |\Omega|^2 \cos^2(kx)} \right)}{\sqrt{\left(\Delta + \frac{i\Gamma}{2} \right)^2 + |\Omega|^2 \cos^2(kx)}}. \quad \text{(A.1)}$$

Using the expression

$$\frac{\sin\sqrt{a^2 + b^2}}{\sqrt{a^2 + b^2}} = \int_0^1 dz J_0(a\sqrt{1 - z^2}) \cos(bz), \quad \text{(A.2)}$$

which is valid for any complex a and b [16], we represent Eq. (A.1) in the form

$$\varphi_e(x, t) = i\frac{\Omega^* t}{2} \cos(kx) \exp\left\{ i\frac{\Delta + i\Gamma/2}{2} t \right\} \times \int_0^1 dz J_0\left(t\frac{\Delta + i\Gamma/2}{2} \sqrt{1 - z^2} \right) \cos\left[\frac{|\Omega| t}{2} \cos(kx) z \right]. \quad \text{(A.3)}$$

Thus, the calculation of the amplitude of the partial probability $a_n^{(e)}(t)$ (9) is reduced to the calculation of the Fourier transform of the product $\cos(kx)\cos[(|\Omega|tz/2)\cos(kx)]$ over x:

$$a_n^{(e)}(t) = i\frac{\Omega^* t}{2} \exp\left\{ i\frac{\Delta + i\Gamma/2}{2} t \right\} \times \int_0^1 dz J_0\left(t\frac{\Delta + i\Gamma/2}{2} \sqrt{1 - z^2} \right) \times \int_0^\pi \frac{d\varphi}{\pi} \cos\varphi \cos\left(\frac{|\Omega| t}{2} z \cos\varphi \right) \exp(-in\varphi). \quad \text{(A.4)}$$

The integral with respect to the variable φ in the right-hand side of Eq. (A.4) can be calculated by sub-

stituting the well-known expression for a sum containing Bessel functions J_{2n} [12, 16]

$$\sum_{n=-\infty}^{+\infty}(-1)^n J_{2n}(\eta)\cos(2n\varphi) = \cos(\eta\cos\varphi) \quad \text{(A.5)}$$

into the integrand, which gives

$$\frac{1}{\pi}\int_0^\pi d\varphi\cos\varphi\cos\left(\frac{|\Omega|t}{2}z\cos\varphi\right)\exp(-in\varphi)$$
$$= \frac{(-1)^{m+1}}{2}\left\{J_{2m+2}\left(\frac{|\Omega|t}{2}z\right) - J_{2m}\left(\frac{|\Omega|t}{2}z\right)\right\}\delta_{n,2m+1}. \quad \text{(A.6)}$$

Here, $\delta_{n,2m+1}$ is the Kronecker delta and $m = 0, \pm1, \pm2, \ldots$

By substituting expression (A.6) into Eq. (A.4) for the amplitude $a_n^{(e)}(t)$, we obtain expression (63) for the probability $W_{2n+1}^{(e)}(t)$.

To calculate the Fourier components for the function $\varphi_m(x, t)$, we convert Eq. (60) to the form

$$\varphi_m(x,t) = \exp\left\{i\frac{\Delta + i\Gamma/2}{2}t\right\}$$
$$\times\left\{\cos\left(\frac{t}{2}\sqrt{\left(\Delta + \frac{i\Gamma}{2}\right)^2 + |\Omega|^2\cos^2(kx)}\right)\right.$$
$$-i\left(\Delta + \frac{i\Gamma}{2}\right)\sin\left(\frac{t}{2}\sqrt{\left(\Delta + \frac{i\Gamma}{2}\right)^2 + |\Omega|^2\cos^2(kx)}\right)$$
$$\left.\times\frac{1}{\sqrt{\left(\Delta + \frac{i\Gamma}{2}\right)^2 + |\Omega|^2\cos^2(kx)}}\right\}. \quad \text{(A.7)}$$

The Fourier transform of the first term in the right-hand side of Eq. (A.7) is calculated using a similar formula [16]:

$$\cos\sqrt{a^2 + b^2} = \cos b$$
$$-a\int_0^1 dz J_1(az)\frac{\cos(b\sqrt{1-z^2})}{\sqrt{1-z^2}}, \quad \text{(A.8)}$$

where the quantities $t(\Delta + i\Gamma/2)/2$ and $|\Omega|t\cos(kx)/2$ are used, as before, as parameters a and b, respectively. Using Eqs. (A.2) and (A.8), we represent the function $\varphi_m(x, t)$ (A.7) in the form

$$\varphi_m(x,t) = \exp\left\{i\frac{\Delta + i\Gamma/2}{2}t\right\}$$
$$\times\left\{\cos\left[\frac{|\Omega|t}{2}\cos(kx)\right] - it\frac{\Delta + i\Gamma/2}{2}\right.$$
$$\left.\times\int_0^1 dz\cos\left[\frac{|\Omega|t}{2}\cos(kx)z\right]\left[J_0\left(t\frac{\Delta + i\Gamma/2}{2}\sqrt{1-z^2}\right)\right.\right.$$
$$\left.\left. - i\frac{1}{\sqrt{1-z^2}}J_1\left(t\frac{\Delta + i\Gamma/2}{2}\sqrt{1-z^2}\right)\right]\right\}. \quad \text{(A.9)}$$

Therefore, the problem of calculating the amplitude of the partial probability $a_n^{(m)}(t)$ (9) is reduced to the calculating the Fourier transform of the function $\cos[(|\Omega|tz/2)\cos(kx)]$ over x:

$$\frac{1}{\pi}\int_0^\pi d\varphi\cos\left(\frac{|\Omega|t}{2}z\cos\varphi\right)\exp(-in\varphi). \quad \text{(A.10)}$$

Expression (A.10) can be calculated by substituting Eq. (A.5) into the integrand, which gives

$$\frac{1}{\pi}\int_0^\pi d\varphi\cos\left(\frac{|\Omega|t}{2}z\cos\varphi\right)\exp(-in\varphi)$$
$$= (-1)^m J_{2m}\left(\frac{|\Omega|t}{2}z\right)\delta_{n,2m}. \quad \text{(A.11)}$$

By substituting expression (A.9) into expression (9) for the partial amplitude $a_n^{(m)}(t)$ and taking into account Eq. (11), we obtain expression (62) for the probability $W_{2n}^{(m)}(t)$.

REFERENCES

1. A. P. Kazantsev, G. I. Surdutovich, and V. P. Yakovlev, *The Mechanical Action of Light on Atoms* (Nauka, Moscow, 1991).
2. V. I. Balykin and V. S. Letokhov, *Atom Optics with Laser Light* (Harwood Academic, Chur, 1995).
3. W. P. Schleich, *Quantum Optics in Phase Space* (Wiley, New York, 2001).
4. D. O. Chudesnikov and V. P. Yakovlev, Laser Phys. **1**, 110 (1991).

5. D. S. Krähmer, A. M. Herkommer, E. Mayr, *et al.*, in *Quantum Optics VI*, Ed. by D. F. Walls and J. D. Harvey (Springer, Berlin, 1994), Springer Proc. Phys., Vol. 77, p. 87.
6. M. K. Oberthaler, R. Abfalterer, S. Bernet, *et al.*, Phys. Rev. Lett. **77**, 4980 (1996).
7. H. Batelaan, E. M. Rasel, M. K. Oberthaler, *et al.*, J. Mod. Opt. **44**, 2629 (1997).
8. M. V. Berry and D. H. J. O'Dell, J. Phys. A: Math. Gen. **31**, 2093 (1998).
9. M. K. Oberthaler, R. Abfalterer, S. Bernet, *et al.*, Phys. Rev. A **60**, 456 (1999).
10. G. Borrmann, Phys. Z. **42**, 157 (1941).
11. *Higher Transcendental Functions (Bateman Manuscript Project)*, Ed. by A. Erdelyi (McGraw-Hill, New York, 1953; Nauka, Moscow, 1974), Vol. 2.
12. *Handbook of Mathematical Functions*, 2nd ed., Ed. by M. Abramowitz and I. A. Stegun (Dover, New York, 1971; Nauka, Moscow, 1979).
13. M. A. Efremov, M. V. Fedorov, V. P. Yakovlev, and W. P. Schleich, Laser Phys. **13**, 995 (2003).
14. M. Born and E. Wolf, *Principles of Optics*, 4th ed. (Pergamon Press, Oxford, 1969; Nauka, Moscow, 1970).
15. F. R. Gantmakher, *The Theory of Matrices*, 3rd ed. (Fizmatgiz, Moscow, 1966; Chelsea, New York, 1959), Chap. 9, Para. 8.
16. A. P. Prudnikov, Yu. A. Brychkov, and O. I. Marichev, *Integrals and Series. Special Functions* (Nauka, Moscow, 1983; Gordon and Breach, New York, 1986).

Translated by M. Sapozhnikov

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 539–553.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 596–611.

ATOMS, SPECTRA, RADIATION

Necessary Stationary Condition for Polarized Radiation in Scattering Media

I. A. Vasilieva

Institute of High Energy Densities, Institutes of High Temperatures Scientific Association (IVTAN), Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 127412 Russia

e-mail: vasilieva@mtu-net.ru

Received November 19, 2002

Abstract—The stationary condition is derived taking into account the polarization of radiation in the general case of a scattering inhomogeneous medium in an arbitrary-shape emitter. The necessary stationary condition for an emitter in which radiation is emitted and extinguished simultaneously is complete extinction of the entire emitted radiation. Radiation extinction as a result of absorption by the medium and the emergence of radiation from the emitter is analyzed. The stationary condition is an analytical form of writing that extinction of radiation is a sure event whose probability is equal to unity. The passage of radiation through the medium is described on the basis of the linear transport theory with the help of the matrices of the Green functions. The stationary condition includes the characteristics of polarized radiation extinction of which is analyzed, the absorption coefficients of the medium, and the elements of the matrices of the Green functions, which are determined by optical and geometrical parameters of the emitter. The stationary condition obtained is used for deriving the relations between the components of scalar intensity observed in an arbitrary region of the emitter. These relations include, in addition to the absorption coefficients and the matrix elements of the Green functions, the powers of the primary radiation. Possible applications of the stationary condition and the relations between intensity components in computations and experimental studies are considered. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In this study, stationary radiation in scattering media is investigated taking polarization into account. Radiation is characterized by four Stokes parameters or, which is the same, by the vector intensity [1, 2]. The vector intensity is usually described by the stationary transport equation similar to the scalar equation used in the absence of polarization effects. The results of solution of this equation are diversified and are determined by specific conditions in the emitter. The solutions of stationary transport equations are treated in a large number of publications in this field.

In the following, a different approach will be developed to describe the emission of objects with scattering media. This study is aimed at deriving relations connecting the characteristics of radiation, which are valid under various conditions, i.e., independent of the shape, inhomogeneities, and other specific features of the emitter. Such relations cannot be obtained from an analysis or solution of equations describing the process of radiation transport in question. In order to derive these relations, general conditions that are not contained in the equations are required. The necessary condition for stationary radiation will be analyzed and used here as such a condition.

In a stationary emitter, radiation appears and vanishes continuously. Radiation is usually emitted as a result of transformation of the energy of a substance in the emitter into the radiant energy or as a result of the arrival of radiant flux from without. Radiation may vanish as a result of opposite or other processes. It is important that the entire radiation of a preset frequency, appearing over a certain time interval, disappears over the same time period. Extinction might not occur in the same parts of the emitter where radiation appears; i.e., a detailed balancing that takes place under equilibrium conditions alone is not required. The condition of complete extinction of the emerging radiation is a necessary stationary condition. Radiation forming any stationary flux must vanish since all fluxes result from the passage of primary radiation through the medium.

The necessary stationary condition was derived earlier [3–6] without taking into account the polarization of radiation. The stationary condition was written as the condition that extinction of the emerging radiation is a sure event whose probability is equal to unity. It was proved that the stationary condition provides new ways for describing radiation and may be helpful in experiments and computations. The stationary condition was derived both in the case when the radiation transport theory is valid and when this theory cannot be applied to the interior of a scattering medium.

In this study, the necessary stationary condition is treated as the condition of complete extinction of radiation in the case when polarization has to be taken into account (including the polarization of primary radiation

and the effect of dispersion in the medium on the polarization. The passage of radiation through the medium is described using the matrices of the Green functions. The analysis is carried out under the assumption that the linear theory of radiation transport is applicable for describing both the observed radiant fluxes as well as extinction of radiation. The transport theory can be used under the assumption that the processes of radiation emergence and extinction in the emitter are compensated instantaneously since the radiation averaged over time periods much longer than the times of radiation emergence, extinction, and propagation is usually considered.

In a quite general case, the emitter occupies a certain volume containing a heterogeneous emitting, absorbing, and dissipative medium. The volume is surrounded by a limiting surface. Primary radiation can emerge in the medium in the bulk or penetrate the volume from the surface. The preset intensity of radiation arriving from outside plays the role of the boundary condition of the problem involving the solution of the transport equations. Generally speaking, the boundary surface can be chosen arbitrarily. It is only important that the radiation entering the volume be specified on the boundary surface. Often, a real surface such as the surface of the walls of a gas-discharge or plasma emitter is chosen as the boundary surface. The radiation intensity in any region of the emitter is determined by solving the transport equation if primary sources of radiation in the volume, the characteristics of interaction between radiation and the substance in the volume, and the radiation entering in the volume are known.

Let us consider the conditions for the applicability of the linear transport theory in emitters with considerable dispersion. The linear transport theory can be used when the following basic requirements are satisfied.

First, it is necessary that geometrical optics be applicable in the description of propagation of radiation in the medium in the emitter volume [7, 8]. In a scattering medium, the requirement of a large distance between strong inhomogeneity regions from which scattering takes place is important. Scatterers must be separated by distances larger than the distance between a scatterer and the wave region. At such distances, electromagnetic waves can be regarded as quasi-plane, and geometrical optics is applicable [8]. Geometrical optics must be applicable everywhere within the boundary surface.

Second, the linear theory must be valid. All characteristics of the interaction between radiation and the substance in the volume must be preset; i.e., these characteristics must be independent of the radiation under study. The flux entering the emitter from outside through the boundary surface must also be independent of the radiation in question. This is taken into account in the above boundary condition, according to which the intensity of incoming radiation must be preset. However, the requirement that the incoming flux must be independent of the radiation being analyzed means that the flux falling on the same boundary surface from inside should not return to the emitting volume. In other words, radiation falling on the boundary surface from the interior must disappear completely from the emitter, as assumed in this study.

Thus, radiation in the given problem emerges and disappears in the same region, namely, in a volume surrounded by a bounding surface and on this surface. The linear transport theory must be applicable precisely in this region.

Extinction of radiation emerging at a certain real surface is ensured when the surface absorbs this radiation completely or completely transmits radiation to the surroundings. Cases of reflection or scattering from the boundary surface will not be considered in the subsequent analysis. It appears at first glance that this considerably limits the range of possible application of the results. In fact, the problem can be solved easily in many cases since the boundary surface is chosen, as mentioned above, quite at random and can often be chosen in such a way that the condition for complete extinction of the emerging radiation is satisfied.

In the derivation of the necessary stationary condition, we will consider the extinction of any radiant flux in the emitter. Extinction of radiation is determined by absorption in the substance in the volume of the emitter and complete extinction of radiation falling on the bounding surface.

The stationary condition forms the basis for deriving a relation between the radiation intensity components. The relation contains quantities which have to be investigated both theoretically and experimentally. Since the stationary condition and the relation between the components are valid in a quite general case, these can be used in theoretical and experimental studies of various emitters.

2. CHARACTERISTICS OF RADIATION AND ITS INTERACTION WITH A MEDIUM

We will consider below elliptically polarized radiation since radiation with an arbitrary degree of polarization can be represented as a superposition of two independent elliptically polarized radiations [1].

Polarized radiation will be described using the Stokes parameters I, Q, U, and V. Here, I is the total scalar intensity defined as a specific radiant flux calculated for unit intervals of time, radiation frequency, solid angle, and area element perpendicular to the flux. This flux includes the entire radiation irrespective of the polarization and does not differ from the scalar intensity used in the case when polarization is not taken into account. The polarization state is described by parameters Q, U, and V. Parameters Q and U depend on the ellipticity (the ratio of the major and minor axes of the polarization ellipse) and the ellipse orientation, while parameter V is determined by the ellipticity and the

direction of motion of electric vector **E** (Fig. 1). In Fig. 1, the ellipticity is determined by angle β and orientation in the coordinate system *x*, *y*, *z* is determined by angle χ. Parameters *Q*, *U*, and *V* can be defined as [1]

$$\begin{aligned} Q &= I\cos(2\beta)\cos(2\chi), \\ U &= I\cos(2\beta)\sin(2\chi), \\ V &= I\sin(2\beta). \end{aligned} \tag{2.1}$$

All the Stokes parameters are proportional to the total radiation intensity *I* and have the same dimension. In the case of nonpolarized radiation, $Q = U = V = 0$, while for linearly polarized light, for $\chi = 0$, we have $U = V = 0$.

We introduce the following notation for the trigonometric factors appearing in Eqs. (2.1):

$$\begin{aligned} q \equiv \cos(2\beta)\cos(2\chi), \quad u \equiv \cos(2\beta)\sin(2\chi), \\ v \equiv \sin(2\beta). \end{aligned} \tag{2.2}$$

All the Stokes parameters are usually written as intensity vector **I**, i.e., in the form of a 1×4 matrix. Using notation (2.2), we can represent the intensity vector in one of the following forms:

$$\mathbf{I} = \begin{bmatrix} I \\ Q \\ U \\ V \end{bmatrix} = I \times \begin{bmatrix} 1 \\ q \\ u \\ v \end{bmatrix} = I \times [1, q, u, v]^T. \tag{2.3}$$

Here and below, the symbol ^{T} denotes transposition.

In the vector equation for radiation transport, the change in the Stokes parameters in each act of scattering is determined by the phase matrix, which can be written in the general case in the form

$$\mathbf{Z}[(\mathbf{u}' \longrightarrow \mathbf{u})] = \begin{bmatrix} Z_{II} & Z_{IQ} & Z_{IU} & Z_{IV} \\ Z_{QI} & Z_{QQ} & Z_{QU} & Z_{QV} \\ Z_{UI} & Z_{UQ} & Z_{UU} & Z_{UV} \\ Z_{VI} & Z_{VQ} & Z_{VU} & Z_{VV} \end{bmatrix}. \tag{2.4}$$

Here, $\mathbf{u}'$ and $\mathbf{u}$ are the unit vectors determining the direction of radiation before and after scattering. The phase matrix and its elements depend on these directions.

Individual elements of matrix (2.4) describe the transformation of each of the Stokes parameters as a result of a single act of scattering. In accordance with the generally accepted notation, the second subscript on each element indicates the parameter being transferred, while the first subscript indicates the resultant parameter. For example, Z_{QI} describes the variation of parameter *Q* as a result of variation of parameter *I*.

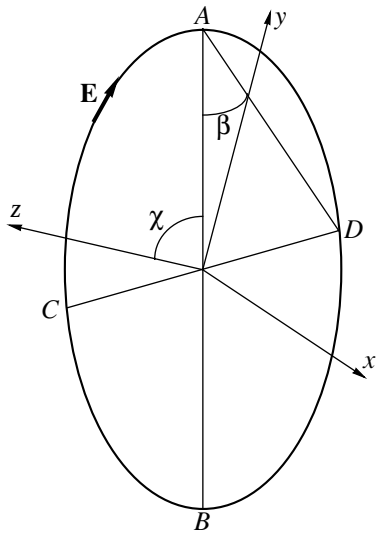

Fig. 1. Polarization ellipse: *x*, *y*, and *z* represent an arbitrary system of coordinates; *AB* and *CD* are the major and minor axes of the ellipse; χ is the angle determining the orientation of the ellipse in the *x*, *y*, *z* system; and β is the angle determining the ellipticity of polarization.

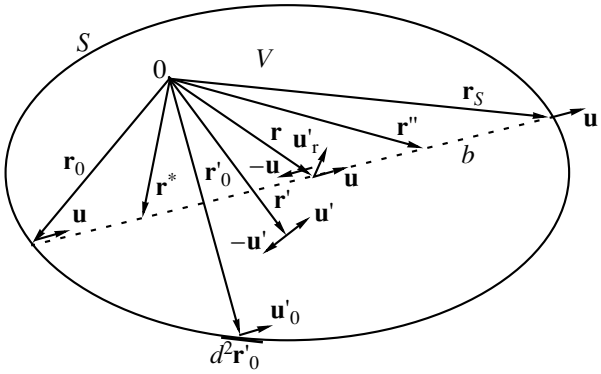

Fig. 2. Schematic diagram illustrating the passage of radiation through a scattering medium: *b* is the light ray, **r**, **r**', **r**'', and **r*** are radius vectors of points in the bulk of volume *V*; $\mathbf{r}_0$, $\mathbf{r}'_0$, and $\mathbf{r}_S$ are the radius vectors of points on surface *S*; **u**, –**u**, **u**', –**u**', $\mathbf{u}'_\mathbf{r}$, and $\mathbf{u}'_0$ are the unit vectors determining the direction of radiation.

Phase matrix $\mathbf{Z}[\mathbf{r}, (\mathbf{u}' \longrightarrow \mathbf{u})]$ determines the transformation of intensity vector $\mathbf{I}(\mathbf{r}, \mathbf{u}')$ into vector $\mathbf{I}_{\text{scat1}}(\mathbf{r}, \mathbf{u})$ upon single scattering via the relation

$$\mathbf{I}_{\text{scat1}}(\mathbf{r}, \mathbf{u}) = \mathbf{Z}[\mathbf{r}, (\mathbf{u}' \longrightarrow \mathbf{u})] \times \mathbf{I}(\mathbf{r}, \mathbf{u}'). \tag{2.5}$$

Here, it is noted that phase matrix **Z** in the general case may be different in different regions of the emitter; i.e., it can be a function of radius vector **r** (Fig. 2). When scatterers possess symmetry, the characteristics of scattering may satisfy certain symmetry relations [9]. In the subsequent analysis, the reciprocity relation connecting

the phase functions in forward and backward scattering will be required. It was shown in [10] that the reciprocity relation in the case of scattering from arbitrarily oriented particles possessing a symmetry plane has the form

$$\mathbf{Z}(-\mathbf{u} \longrightarrow -\mathbf{u}') = \mathbf{e}_3\mathbf{Z}^T(\mathbf{u}' \longrightarrow \mathbf{u})\mathbf{e}_3. \quad (2.6)$$

Here, $(\mathbf{u}' \longrightarrow \mathbf{u})$ and $(-\mathbf{u} \longrightarrow -\mathbf{u}')$ determine the change in the direction of emission as a result of single direct and backward acts of scattering, respectively, and $\mathbf{e}_3 \equiv$ diag[1, 1, –1, 1] is a diagonal matrix. In accordance with Eq. (2.6), multiplication by this matrix reverses the signs of nondiagonal elements of the third column and the third row of matrix $\mathbf{Z}^T$, which describes the sign reversal of Stokes parameter U [10, 11]. This is required since the reversal of the direction of the radiant flux changes the sign of angle χ (see Fig. 1) and, in accordance with formulas (2.1), the sign of U also.

When scatterers have an arbitrary shape, reciprocity relation (2.6) holds only in certain special cases, for example, in the case of Rayleigh scattering, when randomly oriented scatterers are small as compared to the radiation wavelength.

In this study, the interaction of radiation with a substance is described not only by the phase matrix, but also by the coefficients of absorption (k_{abs}), scattering (k_{scat}), and extinction ($k_{ext} = k_{abs} + k_{scat}$). It is assumed that the coefficients of absorption and scattering (as well as of extinction) are independent of the direction of radiation and its polarization, but may depend on the position of the region in the emitter under investigation. The refractive index of the medium, which determines the propagation of radiation between acts of scattering, is assumed to be equal to unity.

In this study, we consider only elastic scattering, when the radiation frequency does not change as a result of scattering. All the characteristics of radiation and its interaction with the medium introduced above are usually functions of frequency; for the sake of brevity, the notation for frequency will be omitted.

3. MATRICES OF THE GREEN FUNCTIONS AND EXPRESSION OF INTENSITY IN TERMS OF PRIMARY RADIATION SOURCES

In this section, we will write the transport equation in the general case, introduce the matrices of the Green functions, and represent the solution to the equation with the help of these matrices. The Green function will be used in Section 4 for deriving the stationary condition. The solution to the transport equation will be required for deriving the relation between the radiation intensity components in Section 5.

We will write the steady-state transport equation for polarized radiation. Suppose that volume V contains an inhomogeneous scattering, absorbing, and emitting medium (see Fig. 2). Primary radiation emitted by the medium can be polarized. Volume V is surrounded by an arbitrary nonconcave surface S from which radiation with a preset vector intensity $\mathbf{I}_S(\mathbf{r}_0, \mathbf{u}_0)$ can enter the volume. This determines the boundary conditions of the problem. The transport equation for the intensity vector has the form [1, 12]

$$\frac{\partial \mathbf{I}(\mathbf{r}, \mathbf{u})}{\partial b} = \mathbf{j}(\mathbf{r}, \mathbf{u}) - k_{ext}\mathbf{I}(\mathbf{r}, \mathbf{u}). \quad (3.1)$$

Here, $\partial/\partial b$ denotes differentiation along ray b passing through point $\mathbf{r}$ in direction $\mathbf{u}$. Vector $\mathbf{j}$ is a function of sources, which describes the emission of radiation by a unit volume of the medium in unit intervals of solid angle, frequency, and time. In the case considered here, we can write

$$\mathbf{j}(\mathbf{r}, \mathbf{u}) = k_{scat}(\mathbf{r})\int_{4\pi} d\mathbf{u}'_\mathbf{r}\mathbf{Z}(\mathbf{r}, \mathbf{u}'_\mathbf{r} \longrightarrow \mathbf{u})\mathbf{I}(\mathbf{r}, \mathbf{u}'_\mathbf{r}) + \mathbf{p}_V(\mathbf{r}, \mathbf{u}) + \mathbf{p}_S(\mathbf{r}, \mathbf{u}), \quad (3.2)$$

where the first term on the right-hand side is the source determined by scattering of radiation arriving from all directions, $\mathbf{u}'_\mathbf{r}$ is the direction of arriving radiation, $d\mathbf{u}'_\mathbf{r}$ is an element of the solid angle, $\mathbf{p}_V(\mathbf{r}, \mathbf{u})$ is the primary source of radiation in volume V, and $\mathbf{p}_S(\mathbf{r}, \mathbf{u})$ is the primary stepwise radiation source determined by the expression:

$$\mathbf{p}_S(\mathbf{r}, \mathbf{u}) = \mathbf{I}_S(\mathbf{r}_0, \mathbf{u}_0)\delta(\mathbf{r} - \mathbf{r}_0)\delta(\mathbf{u} - \mathbf{u}_0). \quad (3.3)$$

The boundary conditions specified on surface S are introduced into the transport equation with the help of such a surface source, as was done in [13] for nonpolarized radiation. The vectors $\mathbf{p}_V$ and $\mathbf{I}_S$ of primary volume sources and intensities of radiation entering the volume are assumed to be preset.

Vectors $\mathbf{j}$ and $\mathbf{p}$ of sources introduced above are determined by the Stokes parameters analogously to the intensity vector defined by formulas (2.3). For example, the vector of primary sources can be represented in the form

$$\mathbf{p}_V(\mathbf{r}, \mathbf{u}) = p_V(\mathbf{r}, \mathbf{u})[1, q_p, u_p, v_p]^T. \quad (3.4)$$

Here, $p_V(\mathbf{r}, \mathbf{u})$ is the total power of the primary source, including radiation of any polarization. The polarization of the primary source is determined by vector $[1, q_p, u_p, v_p]^T$; for the sake of brevity, arguments $(\mathbf{r}, \mathbf{u})$ have been omitted in the notation of this vector. Subscripts p denote the trigonometric characteristics of polarization of the primary sources. It can be seen from

Eq. (3.3) that polarization $\mathbf{p}_S$ of the source does not differ from polarization of intensity $\mathbf{I}_S(\mathbf{r}_0, \mathbf{u}_0)$.

The solution to the linear transport equation, i.e., the expression for intensity vector $\mathbf{I}$ in terms of primary sources $\mathbf{p}_V$ and $\mathbf{p}_S$, can be written in terms of the matrix of the Green functions $\mathbf{G}$ [14–17]. In the general case, a matrix of the Green function has the form

$$\mathbf{G}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] = \begin{bmatrix} G_{II} & G_{IQ} & G_{IU} & G_{IV} \\ G_{QI} & G_{QQ} & G_{QU} & G_{QV} \\ G_{UI} & G_{UQ} & G_{UU} & G_{UV} \\ G_{VI} & G_{VQ} & G_{VU} & G_{VV} \end{bmatrix}. \quad (3.5)$$

This matrix describes the variation of Stokes parameters as a result of passage of radiation from point $(\mathbf{r}', \mathbf{u}')$ to point $(\mathbf{r}, \mathbf{u})$ in the emitter volume. The meaning of the indices of the matrix elements is similar to that of the indices of the phase matrix in the sense that the second subscript indicates the Stokes parameter being transformed and the first subscript denotes the resultant parameter. The significant difference is that the parameter being transformed characterizes radiation at the initial point $(\mathbf{r}', \mathbf{u}')$ of radiation transport through the medium, while the resultant parameter characterizes radiation at the final point $(\mathbf{r}, \mathbf{u})$. Each matrix element describes this transformation of the initial parameter to the final one. For example, G_{UI} describes the transformation of parameter I characterizing radiation at point $(\mathbf{r}', \mathbf{u}')$ into parameter U corresponding to radiation at final point $(\mathbf{r}, \mathbf{u})$.

Function $\mathbf{G}$ has the meaning of the response function. To be more precise, function $\mathbf{G}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]$ describes the part of the radiant flux having direction $\mathbf{u}'$ at point $\mathbf{r}'$, which arrives at point $\mathbf{r}$ in direction $\mathbf{u}$. The polarization state of radiation changes in this case. When radiation passes through the emitter, scattering may occur in any region of volume V any number of times, the last scattering occurring at any point $\mathbf{r}^*$ of ray b between points $\mathbf{r}_0$ and $\mathbf{r}$ (see Fig. 2). If radiation gets from volume V to surface S, it is excluded from the subsequent analysis since the intensity of radiation entering volume V from surface S has already been taken into account in the given problem by specifying the boundary intensity $\mathbf{I}_S(\mathbf{r}_0, \mathbf{u}_0)$.

Function $\mathbf{G}$ is calculated for unit area elements and unit intervals of the solid angle at the final point of radiation trajectory, i.e., at $(\mathbf{r}, \mathbf{u})$. Since radiation transport is a statistical process, functions $\mathbf{G}$ have a probabilistic meaning. For example, $\mathbf{G}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]$ is the probability density that radiation passes through the medium along arbitrary paths in volume V from point $(\mathbf{r}', \mathbf{u}')$ to point $(\mathbf{r}, \mathbf{u})$.

The equation determining the matrix of the response function can be written on the basis of probabilistic meaning, taking into account all possible paths of radiation in the scattering medium from one point to another without taking into account the polarization as was done in Section 3.2 of [3]. We will not write this equation here to avoid repetition.

Matrix $\mathbf{G}$ and all its elements are determined by the set of optical parameters of the interaction of radiation with the substance introduced above; consequently, the matrix and its elements can be functions of radiation frequency. The response functions also depend on the geometrical parameters of the emitter. On the other hand, the matrices of the response function describe the propagation of an arbitrary radiation of the given frequency in the medium and are independent of the characteristics of this radiation and on whether or not the radiation in question is primary radiation.

It can be proved that a similar relation holds for the matrices of the Green functions in the case when the reciprocity relation holds for phase functions (2.6) and the coefficients of absorption and extinction are independent of the direction:

$$\mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = \mathbf{e}_3\mathbf{G}^T[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]\mathbf{e}_3. \quad (3.6)$$

The correctness of this relation can be proved in the same way as in [13] for nonpolarized radiation on the basis of analysis of equations for the Green function. Similar relations are successfully used in various applications (see, for example, [14, 15]).

The expression for the intensity vector in terms of primary sources with the help of the matrices of the Green functions at an arbitrary point of the emitter can be written in the form

$$\mathbf{I}(\mathbf{r}, \mathbf{u}) = \int_V d^3\mathbf{r}' \int_{4\pi} d\mathbf{u}' \mathbf{G}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]\mathbf{p}_V(\mathbf{r}', \mathbf{u}') + \int_S d^3\mathbf{r}'_0 \int_{2\pi} d\mathbf{u}'_0 \mathbf{G}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]\mathbf{I}_S(\mathbf{r}'_0, \mathbf{u}'_0). \quad (3.7)$$

Here, $d^3\mathbf{r}'$ and $d^2\mathbf{r}'_0$ are elements of the volume and the surface and $d\mathbf{u}'$ and $d\mathbf{u}'_0$ are elements of solid angles. Here and below, the representation of the intensity vector in the form of sums and integrals is a consequence of the linearity of the transport equation and of the lack of correlation for primary sources of radiation.

The first term in Eq. (3.7) takes into account the arrival of radiation from all primary volume sources $p_V(\mathbf{r}', \mathbf{u}')$ at point $(\mathbf{r}, \mathbf{u})$, while the second term takes into account all surface sources $\mathbf{I}_S(\mathbf{r}'_0, \mathbf{u}'_0)$. We denote by $\mathbf{I}_V(\mathbf{r}, \mathbf{u})$ and $\mathbf{I}_S(\mathbf{r}, \mathbf{u})$ the first and second terms on the right-hand side of Eq. (3.7). In addition, we introduce instead of the integrands the partial intensity components $\mathbf{i}$ of radiation observed at point $(\mathbf{r}, \mathbf{u})$. Each partial component of intensity $\mathbf{i}$ is due to primary radiation emitted in a certain region of the volume or the surface,

the primary radiation being calculated per unit volume or surface per unit solid angle:

$$\mathbf{I}(\mathbf{r}, \mathbf{u}) = \mathbf{I}_V(\mathbf{r}, \mathbf{u}) + \mathbf{I}_S(\mathbf{r}, \mathbf{u}), \tag{3.8}$$

$$\mathbf{I}_V(\mathbf{r}, \mathbf{u}) = \int_V d^3\mathbf{r}' \int_{4\pi} d\mathbf{u}' \mathbf{i}_\mathbf{V}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})], \tag{3.9}$$

$$\mathbf{I}_S(\mathbf{r}, \mathbf{u}) = \int_S d^2\mathbf{r}'_0 \int_{2\pi} d\mathbf{u}'_0 \mathbf{i}_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]. \tag{3.10}$$

Let us write in greater detail the partial scalar intensities, i.e., partial components of the Stokes parameter I. We assume that i_V and i_S are determined by the integrands from Eq. (3.7), where the Green functions are described by matrices (3.5), while primary sources are described by vectors (3.4) for ($\mathbf{r} = \mathbf{r}'$, $\mathbf{u} = \mathbf{u}'$) or by Eq. (2.3) for $I = I_S(\mathbf{r}'_0, \mathbf{u}'_0)$. Then the partial scalar intensities are determined by the product of the first row of matrix (3.5) and a column of vector (3.4) or (2.3). We also take into account the fact that the trigonometric factors of primary sources depend on the radius vectors and directions of these sources. This gives

$$\begin{aligned} &i_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] \\ &= p_V(\mathbf{r}', \mathbf{u}')\{G_{II}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] \\ &\quad + G_{IQ}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] q_p(\mathbf{r}', \mathbf{u}') \\ &\quad + G_{IU}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] u_p(\mathbf{r}', \mathbf{u}') \\ &\quad + G_{IV}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] v_p(\mathbf{r}', \mathbf{u}')\}, \end{aligned} \tag{3.11}$$

$$\begin{aligned} &i_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] \\ &= I_S(\mathbf{r}'_0, \mathbf{u}'_0)\{G_{II}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] \\ &\quad + G_{IQ}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] q(\mathbf{r}'_0, \mathbf{u}'_0) \\ &\quad + G_{IU}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] u(\mathbf{r}'_0, \mathbf{u}'_0) \\ &\quad + G_{IV}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] v(\mathbf{r}'_0, \mathbf{u}'_0)\}. \end{aligned} \tag{3.12}$$

It follows hence that the partial scalar intensities depend to a considerable extent on the scalar and trigonometric characteristics of primary sources as well as on all geometrical and optical parameters of the object (via the elements of the Green function).

4. NECESSARY STATIONARY CONDITION

The statement that the emergence of radiation in a steady-state emitter must be compensated by its extinction refers to any radiant flux. Let us analyze all possibilities for the extinction of a steady-state flux having direction $\mathbf{u}$ at point $\mathbf{r}$ and characterized by intensity $\mathbf{I}(\mathbf{r}, \mathbf{u})$.

In the problem considered here, radiation extinction can be due only to the absorption by the medium or emergence beyond the emitter, i.e., to the surrounding surface S. In contrast to scattering, the extinction of radiation affects only scalar parameters of fluxes and does not influence their polarization since absorption by the medium, which is determined by absorption coefficient k_{abs}, as well as extinction on the surrounding surface, is independent of polarization.

The flux in question passes through the emitter and can, generally speaking, reach any point in volume V or surface S as a result of scattering. If the passage of radiation is described with the help of the Green functions, the arrival from point ($\mathbf{r}$, $\mathbf{u}$) to any point ($\mathbf{r}'$, $\mathbf{u}'$) can be represented in the form

$$\begin{aligned} &\mathbf{i}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] \\ &= G[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')]\mathbf{I}(\mathbf{r}, \mathbf{u}). \end{aligned} \tag{4.1}$$

Analogously to expressions (3.11) and (3.12), we can write expressions for the partial scalar intensities of radiation arriving from point ($\mathbf{r}$, $\mathbf{u}$) to various parts of the emitter volume and surface:

$$\begin{aligned} &i_V[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] \\ &= I(\mathbf{r}, \mathbf{u})\{G_{II}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] \\ &\quad + G_{IQ}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] q(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IU}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] u(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IV}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] v(\mathbf{r}, \mathbf{u})\}, \end{aligned} \tag{4.2}$$

$$\begin{aligned} &i_S[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}'_0, \mathbf{u}'_0)] \\ &= I(\mathbf{r}, \mathbf{u})\{G_{II}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}'_0, \mathbf{u}'_0)] \\ &\quad + G_{IQ}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}'_0, \mathbf{u}'_0)] q(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IU}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}'_0, \mathbf{u}'_0)] u(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IV}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}'_0, \mathbf{u}'_0)] v(\mathbf{r}, \mathbf{u})\}. \end{aligned} \tag{4.3}$$

These fluxes were calculated per unit area element perpendicular to directions $\mathbf{u}'$ and $\mathbf{u}'_0$, respectively. In order to determine the flux absorbed by the medium in a unit volume surrounding $\mathbf{r}'$, we must multiply the obtained expression by the absorption coefficient $k_{abs}(\mathbf{r}')$. To find the fraction of scalar flux $I(\mathbf{r}, \mathbf{u})$ absorbed in the vicinity of point $\mathbf{r}'$ in the same unit volume, we must divide the result by the flux. As a result, we find that a unit volume in the vicinity of ($\mathbf{r}'$, $\mathbf{u}'$) absorbs the following fraction of the radiation that had direction $\mathbf{u}$ at point $\mathbf{r}$:

$$\begin{aligned} &D_V[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] \\ &= k_{abs}(\mathbf{r}')\{G_{II}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] \\ &\quad + G_{IQ}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] q(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IU}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] u(\mathbf{r}, \mathbf{u}) \\ &\quad + G_{IV}[(\mathbf{r}, \mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')] v(\mathbf{r}, \mathbf{u})\}. \end{aligned} \tag{4.4}$$

Similarly, we can write the expression for the fraction of the same radiation emitted at point (**r**, **u**) and arriving at the arbitrary region of bounding surface *S*, in which radiation vanishes completely according to the condition of the problem:

$$\begin{aligned} D_S[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')] &= G_{II}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')] \\ &+ G_{IQ}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')]q(\mathbf{r},\mathbf{u}) \\ &+ G_{IU}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')]u(\mathbf{r},\mathbf{u}) \\ &+ G_{IV}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')]v(\mathbf{r},\mathbf{u})\}. \end{aligned} \tag{4.5}$$

The fractions of the flux introduced above have a probabilistic meaning as functions of the response; namely, each such fraction is the probability density that radiation that had direction **u** at point **r** passes arbitrarily to point (**r**', **u**') or ($\mathbf{r}_0'$, $\mathbf{u}_0'$), where it vanishes in a unit volume or at a unit surface area. In order to derive the expression for the complete extinction of radiation $I(\mathbf{r}, \mathbf{u})$ in an arbitrary region, we must integrate D_V and D_S over the volume and the surface, respectively. In addition, we must integrate with respect to the solid angle to take into account the fact that radiation can arrive in the region of its extinction from different directions:

$$\begin{aligned} &\int_V d^3\mathbf{r}' \int_{4\pi} d\mathbf{u}' D_V[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}',\mathbf{u}')] \\ &+ \int_S d^3\mathbf{r}_0' \int_{2\pi} d\mathbf{u}_0' D_S[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')] = 1. \end{aligned} \tag{4.6}$$

This is the necessary stationary condition, i.e., the condition that the entire radiation in question, which had direction **u** at point **r**, disappears completely in the bulk or at the surface of the emitter. In other words, equality (4.6) indicates that the extinction of the radiation observed in any region of the emitter is a sure event. Condition (4.6) includes the characteristics of polarization of primary radiation, absorption coefficients, and the elements of the first row of the Green matrix (3.5). It is important to note that integration and summation in Eq. (4.6) account for all possibilities of extinction of the chosen radiation both in the bulk and at the surface.

When condition (4.6) is not satisfied, radiation cannot be steady-state radiation.

In the case of natural radiation at point (**r**, **u**), for $q = u = v = 0$, expressions (4.4) and (4.5) for the fractions of vanishing radiation are simplified:

$$\begin{aligned} D_V[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}',\mathbf{u}')] &= k_{\text{abs}}(\mathbf{r}')G_{II}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}',\mathbf{u}')], \\ D_S[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')] &= G_{II}[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}_0',\mathbf{u}_0')]. \end{aligned} \tag{4.7}$$

In the absence of scattering ($k_{\text{ext}} = k_{\text{abs}}$), radiation is absorbed in the medium only along the path from point **r** in direction **u** along ray *b* (at points **r**''), right up to point $\mathbf{r}_S$, at which it vanishes at the surface (see Fig. 2). In this case, for the Green function, we have

$$G[(\mathbf{r},\mathbf{u}) \longrightarrow (\mathbf{r}'',\mathbf{u})] = \exp[-t(\mathbf{r} \longrightarrow \mathbf{r}'')], \tag{4.8}$$

where

$$t(\mathbf{r} \longrightarrow \mathbf{r}'') = \left| \int_{\mathbf{r}}^{\mathbf{r}''} k_{\text{ext}}(\mathbf{r}) d\mathbf{r} \right| \tag{4.9}$$

is the optical density between points **r** and **r**''. Integration is carried out along ray *b* passing through these points. The exponential function in Eq. (4.8) describes the probability that radiation passes from **r** to **r**'' without interacting with the medium.

Instead of Eq. (4.6), in this case we obtain

$$\int_{\mathbf{r}}^{\mathbf{r}_S} k_{\text{abs}}(\mathbf{r}'')\exp[-t(\mathbf{r} \longrightarrow \mathbf{r}'')]d\mathbf{r}'' + \exp[-t(\mathbf{r} \longrightarrow \mathbf{r}_S)] = 1.$$

Let us now consider the extinction of radiation emerging primarily at point **r** in direction –**u**. The flux of such radiation emerging from a unit volume is determined by the vector $\mathbf{p}_V(\mathbf{r}, -\mathbf{u})$ of the primary source in accordance with formula (3.4). This radiation forms at point **r** a flux whose direction is opposite to that of the flux described by formulas (3.7)–(3.12) and used for deriving the stationary condition (4.6). In this case, the signs of trigonometric parameter *u* and, hence, the Stokes parameter *U* are opposite to the signs of the same parameters of radiation observed at (**r**, **u**).

Let us write the primary source vector at point (**r**, –**u**) using formula (3.4):

$$\begin{aligned} \mathbf{p}_V(\mathbf{r},-\mathbf{u}) &= p_V(\mathbf{r},-\mathbf{u}) \\ &\times [1, q_p(\mathbf{r},-\mathbf{u}), u_p(\mathbf{r},-\mathbf{u}), v_p(\mathbf{r},-\mathbf{u})]^T. \end{aligned} \tag{4.10}$$

Let us consider all possibilities of extinction of the chosen radiant flux in the bulk of volume *V* and at surface *S* with the help of the Green functions $\mathbf{G}_{\uparrow\downarrow}$ describing the arrival of radiation from (**r**, –**u**) to an arbitrary

region in the bulk or on the surface of the emitter. For example, the partial intensity of radiation passing through a unit area element in a unit solid angle in the vicinity of point $(\mathbf{r}', -\mathbf{u}')$ in the bulk of volume V is given by

$$\begin{aligned}&\mathbf{i}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &= \mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\mathbf{p}_V(\mathbf{r}, -\mathbf{u})\\ &= \mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]p_V(\mathbf{r}, -\mathbf{u})\\ &\times [1, q_p(\mathbf{r}, -\mathbf{u}), u_p(\mathbf{r}, -\mathbf{u}), v_p(\mathbf{r}, -\mathbf{u})]^T.\end{aligned} \tag{4.11}$$

We can write matrix $\mathbf{G}_{\uparrow\downarrow}$ in a form similar to Eq. (3.5):

$$\begin{aligned}&\mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &= \begin{bmatrix} G_{\uparrow\downarrow II} & G_{\uparrow\downarrow IQ} & G_{\uparrow\downarrow IU} & G_{\uparrow\downarrow IV} \\ G_{\uparrow\downarrow QI} & G_{\uparrow\downarrow QQ} & G_{\uparrow\downarrow QU} & G_{\uparrow\downarrow QV} \\ G_{\uparrow\downarrow UI} & G_{\uparrow\downarrow UQ} & G_{\uparrow\downarrow UU} & G_{\uparrow\downarrow UV} \\ G_{\uparrow\downarrow VI} & G_{\uparrow\downarrow VQ} & G_{\uparrow\downarrow VU} & G_{\uparrow\downarrow VV} \end{bmatrix}.\end{aligned} \tag{4.12}$$

After multiplication, in accordance with Eq. (4.11), we obtain the following expression for the scalar intensity of the flux arriving from point $(\mathbf{r}, -\mathbf{u})$ to point $(\mathbf{r}', -\mathbf{u}')$:

$$\begin{aligned}i_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] &= [G_{\uparrow\downarrow II} + G_{\uparrow\downarrow IQ}q_p(\mathbf{r}, -\mathbf{u})\\ &+ G_{\uparrow\downarrow IU}u_p(\mathbf{r}, -\mathbf{u}) + G_{\uparrow\downarrow IV}v_p(\mathbf{r}, -\mathbf{u})]p_V(\mathbf{r}, -\mathbf{u}).\end{aligned} \tag{4.13}$$

This flux is calculated for a unit area element perpendicular to direction $-\mathbf{u}'$. Now, we must completely repeat the arguments concerning radiation extinction, which led to formulas (4.4)–(4.6). Instead of these formulas, we obtain

$$\begin{aligned}&D_V[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &= k_{\text{abs}}(\mathbf{r}')\{G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &+ G_{\uparrow\downarrow IQ}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]q_p(\mathbf{r}, -\mathbf{u})\\ &+ G_{\uparrow\downarrow IU}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]u_p(\mathbf{r}, -\mathbf{u})\\ &+ G_{\uparrow\downarrow IV}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]v_p(\mathbf{r}, -\mathbf{u})\}.\end{aligned} \tag{4.14}$$

$$\begin{aligned}&D_S[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]\\ &= G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]\\ &+ G_{\uparrow\downarrow IQ}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]q_p(\mathbf{r}, -\mathbf{u})\\ &+ G_{\uparrow\downarrow IU}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]u_p(\mathbf{r}, -\mathbf{u})\\ &+ G_{\uparrow\downarrow IV}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]v_p(\mathbf{r}, -\mathbf{u}).\end{aligned} \tag{4.15}$$

$$\begin{aligned}&\int_V d^3\mathbf{r}' \int_{4\pi} d(-\mathbf{u}')D_V[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &+ \int_V d^2\mathbf{r}'_0 \int_{2\pi} d(-\mathbf{u}'_0)D_S[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)] = 1.\end{aligned} \tag{4.16}$$

The latter expression is also a necessary stationary condition; i.e., the entire radiation emerging initially at point $\mathbf{r}$ in direction $-\mathbf{u}$ vanishes completely in the bulk or on the surface of the emitter. The only difference from Eq. (4.6) is that Eq. (4.16) is written for the extinction of primary radiation and that its direction is $-\mathbf{u}$ and not $\mathbf{u}$. The fractions D_V and D_S include the characteristics of polarization of radiation whose absorption is analyzed.

If the radiation analyzed at point $(\mathbf{r}, -\mathbf{u})$ is not polarized ($q_p = u_p = v_p = 0$), we obtain, analogously to Eq. (4.7),

$$\begin{aligned}&D_V[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]\\ &= k_{\text{abs}}(\mathbf{r}')G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')],\\ &D_S[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]\\ &= G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)].\end{aligned} \tag{4.17}$$

These fractions D_V and D_S do not differ from the probabilities of extinction of steady-state radiation in the bulk and on the boundary surface obtained earlier [3].

In the absence of scattering, the radiation in question is absorbed only on its path from point $\mathbf{r}$ in direction $-\mathbf{u}$ along ray b to point $\mathbf{r}_0$, while the remaining radiation disappears at the surface at point $\mathbf{r}_0$.

Stationary conditions (4.4)–(4.6) and (4.14)–(4.16) derived above are necessary but not sufficient. In order to obtain a sufficient condition, we must impose a constraint on radiation sources. The sources must not change with time. It is clear that if radiation sources decay, the resultant radiation will be damped.

The necessary condition (4.4)–(4.6) or (4.14)–(4.16) for the extinction of emerging radiation can be written in the form of relations between the matrix elements of the Green functions with preset coefficients of medium absorption and polarization characteristics of the radiation being analyzed. The Green functions are determined by all optical and geometrical parameters of the object; consequently, the stationary condition obtained above is the implicit relation between these parameters, which is required for stationarity. Let us consider a simple example of violation of the stationary condition because of geometrical features of the emitter. Suppose that a steady-state source of radiation is placed in a purely scattering medium surrounded by a closed absolutely reflecting medium. In this case, we can choose the bounding surface S with a given intensity of radiation entering the emitter only outside the mirror. However, the radiation emitted by the object does not reach this surface; i.e., the corresponding response functions and the fractions of arriving radiation are equal to zero. Since, in addition, radiation is not absorbed inside the emitter, $D_V = D_S = 0$, and condition (4.6) or (4.16) cannot be satisfied. Obviously, radiation in this case is accumulated and cannot be steady-state radiation.

5. RELATION BETWEEN RADIATION COMPONENTS

The stationary conditions derived above make it possible to obtain relations between the radiation intensity components. In order to derive the required relations, we will use the steady-state condition (4.16) and the solutions to the steady-state transport equation in the form (3.11), (312). We will connect the partial intensity of radiation arriving at point $(\mathbf{r}, \mathbf{u})$ from point $(\mathbf{r}', \mathbf{u}')$ or from point $(\mathbf{r}'_0, \mathbf{u}'_0)$ with the fraction of radiation propagating in the opposite direction from $(\mathbf{r}, -\mathbf{u})$ and absorbed at these points.

Using Eqs. (3.11) and (4.14), we obtain

$$D_V[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = \frac{k_{\text{abs}}(\mathbf{r}')i_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]}{p_V(\mathbf{r}', \mathbf{u}')} g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})], \tag{5.1}$$

where the following notation is introduced:

$$g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] \equiv \frac{G_{\uparrow\downarrow II} + G_{\uparrow\downarrow IQ} q_p(\mathbf{r}, -\mathbf{u}) + G_{\uparrow\downarrow IU} u_p(\mathbf{r}, -\mathbf{u}) + G_{\uparrow\downarrow IV} v_p(\mathbf{r}, -\mathbf{u})}{G_{II} + G_{IQ} q_p(\mathbf{r}', \mathbf{u}') + G_{IU} u_p(\mathbf{r}', \mathbf{u}') + G_{IV} v_p(\mathbf{r}', \mathbf{u}')}. \tag{5.2}$$

Similarly, we obtain from Eqs. (3.12) and (4.15)

$$D_S[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)] = \frac{i_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]}{I_S(\mathbf{r}'_0, \mathbf{u}'_0)} g_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})], \tag{5.3}$$

where

$$g_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})] \equiv \frac{G_{\uparrow\downarrow II} + G_{\uparrow\downarrow IQ} q_p(\mathbf{r}, -\mathbf{u}) + G_{\uparrow\downarrow IU} u_p(\mathbf{r}, -\mathbf{u}) + G_{\uparrow\downarrow IV} v_p(\mathbf{r}, -\mathbf{u})}{G_{II} + G_{IQ} q(\mathbf{r}'_0, \mathbf{u}'_0) + G_{IU} u(\mathbf{r}'_0, \mathbf{u}'_0) + G_{IV} v(\mathbf{r}'_0, \mathbf{u}'_0)}. \tag{5.4}$$

The arguments of functions G_{mk} are $[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]$ in the expressions for g_V and $[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]$ in the expressions for g_S. The arguments of functions $G_{\uparrow\downarrow mk}$ are $[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]$ in the expressions for g_V and $[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}'_0, -\mathbf{u}'_0)]$ in the expressions for g_S.

Functions g_V and g_S depend on the first rows of the Green matrices, which determine the scalar intensities at the final point of the path of polarized radiation between two points. In addition, functions g_V and g_S depend on the polarization of radiation at the initial point of each such path.

The following important feature of expressions (5.1) and (5.3) is worth noting. It can be seen from formulas (3.11) and (3.12) that the ratios

$$\frac{i_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]}{p_V(\mathbf{r}', \mathbf{u}')}, \quad \frac{i_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]}{I_S(\mathbf{r}'_0, \mathbf{u}'_0)}$$

appearing in these expressions are determined by the polarization of the primary radiation and by the Green functions, but are independent of the intensities p_V and I_S of scalar primary sources. Consequently, expressions (5.1) and (5.3) make sense for infinitely small values of these intensities also. When $p_V \longrightarrow 0$ and $I_S \longrightarrow 0$, indeterminacies of the 0/0 type can be expanded with the help of formulas (3.11) and (3.12) and are equal to the denominators of the expressions for g_V and g_S.

Using formulas (5.1) and (5.3), we can write the stationary condition (4.16) in the form

$$D_V(\mathbf{r}, -\mathbf{u}) + D_S(\mathbf{r}, -\mathbf{u}) = 1, \tag{5.5}$$

where

$$D_V(\mathbf{r}, -\mathbf{u}) = \int_V d^3\mathbf{r}' \int_{4\pi} d(-\mathbf{u}') \frac{k_{\text{abs}}(\mathbf{r}')i_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]}{p_V(\mathbf{r}', \mathbf{u}')} \times g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})], \tag{5.6}$$

$$D_S(\mathbf{r}, -\mathbf{u}) = \int_S d^2\mathbf{r}'_0 \int_{2\pi} d(-\mathbf{u}'_0) \frac{i_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]}{I_S(\mathbf{r}'_0, \mathbf{u}'_0)} \times g_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]. \tag{5.7}$$

Stationary condition (5.5) combined with relations (5.6) and (5.7) reflects the relation between the partial terms i_V and i_S of the scalar intensity of radiation. This relation includes the quantities i_V, k_{abs}/p_V, i_S, and I_S which have been investigated in spectroscopic experiments and computations.

The following two remarks should be made concerning the above relations. First, relation (5.5) is valid for any radiation frequency, although the quantities k_{abs}, i_V, i_S, p_V, and I_S appearing in it can change radically with the frequency. Second, the relations for the necessary stationary condition were derived using the solution of the steady-state transport equation which presumes the invariability of the primary sources of radiation. Consequently, we can assume that this relation is not only necessary, but also sufficient.

From the integrals in the expressions for D_V and D_S, one can separate individual components. This is convenient in the cases when these components can be determined directly from experiments. For instance, at least two components of the integral in Eq. (5.7) can be determined experimentally in different cases.

First, the relative partial intensity of radiation arriving directly from point $\mathbf{r}_0$ of surface S to point $\mathbf{r}$ in direction $\mathbf{u}$ in the absence of interaction with the medium is determined by the exponential function $\exp[-t(\mathbf{r} \longrightarrow \mathbf{r}_0)]$ of the optical density. Integral from formula (5.7) for D_S implicitly contains the corresponding exponential component. If we separate this component from integral D_S, the remaining partial intensity describes radiation arriving from surface S to point $(\mathbf{r}, \mathbf{u})$ only after scattering. In order to emphasize this, we can denote the corresponding partial intensities as $i_{S,\,scat}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]$. The optical density can be measured experimentally from the damping of external radiation. The measurement can be made most easily when the point of observation $\mathbf{r}$ also lies on the surface and is opposite to point $\mathbf{r}_0$ of the emitter; i.e., $\mathbf{r} = \mathbf{r}_S$ (see Fig. 2).

Second, if surface S consists of a radiating part S_{rad}, where $I_S \neq 0$, and a nonradiating part S_{nonrad} that is not illuminated from outside, where $I_S = 0$, we can write

$$S = S_{rad} + S_{nonrad}.$$

It is possible to measure the fraction of primary radiation emerging from surface S_{nonrad} (see Section 5.3). We denote this fraction of primary radiation by $D_{S,\,ex}$. The measurement of $D_{S,\,ex}$ is especially simple when the point of observation lies on the surface, i.e., $\mathbf{r} = \mathbf{r}_S$.

Separating both of the above terms from $D_S(\mathbf{r}, -\mathbf{u})$, we can write

$$D_S(\mathbf{r}, -\mathbf{u}) = D_{S_{rad},\,scat}(\mathbf{r}, -\mathbf{u}) + D_{S,\,ex}(\mathbf{r}, -\mathbf{u}) + \exp[-t(\mathbf{r} \longrightarrow \mathbf{r}_0)]. \tag{5.8}$$

Here,

$$D_{S_{rad},\,scat}(\mathbf{r}, -\mathbf{u}) = \int_{S_{rad}} d^2\mathbf{r}'_0 \int_{2\pi} d(-\mathbf{u}'_0) \times \frac{i_{S,\,scat}[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]}{I_S(\mathbf{r}'_0, \mathbf{u}'_0)} g_S[(\mathbf{r}'_0, \mathbf{u}'_0) \longrightarrow (\mathbf{r}, \mathbf{u})]. \tag{5.9}$$

The quantity $D_{S_{rad},\,scat}(\mathbf{r}, -\mathbf{u})$ differs from $D_S(\mathbf{r}, -\mathbf{u})$ in the domain of integration over the surface and in the fact that it takes into account the emergence of radiation on surface S only after scattering. Substituting Eqs. (5.8) and (5.9) into relation (5.5), we obtain a relation between radiation components, which is similar to that obtained earlier in [3] for the case when polarization was disregarded. Clearly, other components can also be separated in $D_S(\mathbf{r}, -\mathbf{u})$ when required.

Thus, the radiation extinction characteristics $D(\mathbf{r}, -\mathbf{u})$ appearing in formulas (5.5)–(5.9) depend on the polarization of the radiation under study at point $(\mathbf{r}, -\mathbf{u})$ and the matrix elements of the Green functions $\mathbf{G}_{\uparrow\downarrow}$, describing the passage of this radiation in various parts of the object. In addition, the application of the intensity of radiation having direction $\mathbf{u}$ at point $\mathbf{r}$ (see Eqs. (5.1) and (5.3)) resulted in the characteristics of polarization of primary radiation and matrix elements of the Green functions, which describe the arrival of this radiation at point $(\mathbf{r}, \mathbf{u})$. These quantities appear in the denominators of functions g defined by formulas (5.2) and (5.4).

5.1. Functions g_V and g_S

Functions g_V and g_S in relations (5.5)–(5.7) directly take into account the polarization of radiation. Let us consider these functions in greater detail in some special cases.

1. We will use reciprocity relation (3.6) between functions $\mathbf{G}$ and $\mathbf{G}_{\uparrow\downarrow}$ in order to replace the elements of matrix $\mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', \mathbf{u}')]$ by elements of matrix $\mathbf{G}_{\uparrow\downarrow}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]$. For this purpose, we must carry out the multiplication of matrices on the right-hand side of Eq. (3.6) and compare, element by element, the result with the matrix on the left-hand side. Such a comparison of matrix elements leads to the quite expected result for the elements appearing in expression (5.2):

$$G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = G_{II}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})],$$

$$G_{\uparrow\downarrow IQ}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = G_{QI}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})],$$

$$G_{\uparrow\downarrow IU}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = -G_{UI}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})],$$

$$G_{\uparrow\downarrow IV}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = G_{VI}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]. \quad (5.10)$$

Taking into account these expressions in relation (5.2), we obtain

$$g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] = \frac{G_{II} + G_{QI}q_p(\mathbf{r}, -\mathbf{u}) - G_{UI}u_p(\mathbf{r}, -\mathbf{u}) + G_{VI}v_p(\mathbf{r}, -\mathbf{u})}{G_{II} + G_{IQ}q_p(\mathbf{r}', \mathbf{u}') + G_{IU}u_p(\mathbf{r}', \mathbf{u}') + G_{IV}v_p(\mathbf{r}', \mathbf{u}')}. \quad (5.11)$$

Here, all elements of the Green matrix depend on the initial and final points of the radiation path, i.e., on $[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]$.

The elements of the Green matrices appearing in expression (5.4) for g_S satisfy relations similar to (5.10). In order to write these relations, we must replace $(\mathbf{r}', \mathbf{u}')$ by $(\mathbf{r}_0', \mathbf{u}_0')$ and $(\mathbf{r}', -\mathbf{u}')$ by $(\mathbf{r}_0', -\mathbf{u}_0')$.

Thus, when relation (3.6) holds, functions g_V and g_S are determined by the polarization of primary sources, as well as by the first row and column of the matrix of the response function describing the passage of radiation from the points where it emerges to the point $(\mathbf{r}, \mathbf{u})$ where it is observed.

2. Let us now consider function g in particular cases of different polarization of primary radiation.

Let us suppose that the radiation initially emerging in the emitter is nonpolarized natural radiation typical of thermal sources. In this case, at points $(\mathbf{r}', \mathbf{u}')$ and $(\mathbf{r}, -\mathbf{u})$, we have $q_p = u_p = v_p = 0$. If radiation entering the emitter from surface S is also natural, we also have $q = u = v = 0$ at points $(\mathbf{r}_0', \mathbf{u}_0')$ of the surface.

In this case, we obtain from expressions (5.2) and (5.4)

$$g_V = \frac{G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')]}{G_{II}[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]},$$

$$g_S = \frac{G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}_0', -\mathbf{u}_0')]}{G_{II}[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})]}. \quad (5.12)$$

If the reciprocity relation (3.6) and, hence, (5.10), is also valid, we have

$$g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})] = g_S[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})] = 1. \quad (5.13)$$

Thus, the relation between radiation components derived from stationary condition (5.5) combined with relations (5.8) and (5.9) exactly coincides with that obtained for the case when polarization was not taken initially into account, i.e., radiation was regarded as natural everywhere, and the reciprocity relation was satisfied in each scattering act [3].

It is well known that natural radiation passing through a scattering medium can be polarized in it [18]. This also follows from formula (2.5) describing single scattering. Indeed, if at least one nondiagonal element of the first column of phase matrix $\mathbf{Z}$ differs from zero, expression (2.5) for $\mathbf{I}_{\text{scat1}}$ acquires, after multiplications on the right-hand side, nonzero characteristics of polarization even for $q' = u' = v' = 0$. The first column of matrix $\mathbf{Z}$ usually contains nonzero nondiagonal elements. In spite of polarization emerging during scattering of natural radiation, the value of the general (scalar) intensity of scattered radiation is determined only by the scalar quantity of primary flux and the first element of the response function G_{II} (or $G_{\uparrow\downarrow II}$).

When the radiation emerging initially in the bulk of the emitter or on its surface is plane-polarized, we have $u_p = v_p = 0$ and $u = v = 0$ in the regions of its emergence, and the corresponding terms in the numerators and denominators of functions g_V and g_S are absent. Various combinations of the type of primary radiation are possible. Let us suppose, for example, that only natural radiation emerges in the bulk of the emitter and the outer radiation is plane-polarized. This can be realized when a thermal emitter is illuminated from outside by a laser. In this case, $g_V = 1$ everywhere in the volume, while for g_S, we find from (5.4)

$$g_S[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})] = \frac{G_{\uparrow\downarrow II}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}_0, -\mathbf{u}_0')]}{G_{II}[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})] + G_{IQ}[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})]q(\mathbf{r}_0', \mathbf{u}_0')}. \quad (5.14)$$

If relations (5.10) hold, from (5.4) we obtain

$$g_S[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})] = \frac{1}{1 + G_{IQ}[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})]q(\mathbf{r}_0', \mathbf{u}_0')/G_{II}[(\mathbf{r}_0', \mathbf{u}_0') \longrightarrow (\mathbf{r}, \mathbf{u})]}. \quad (5.15)$$

Thus, in various specific cases of polarization, functions g are simplified.

5.2. Relations between Radiation Components in Special Cases

Relation (5.5) taking into account Eqs. (5.6) and (5.7) has been obtained in the general form and is not commonly used for practical applications, However, relations for various special cases can be easily derived from it. Let us consider three such cases.

1. It is often assumed in experimental and theoretical studies of radiation that the main characteristics of the medium, as well as primary sources in the bulk of the emitter and on its surface, remain unchanged. Let us suppose that ratio p_V/k_{abs} in the bulk of the volume, intensity I_S on the boundary surface, and functions g_V and g_S are constant, i.e., independent of coordinates $\mathbf{r}$, $\mathbf{r}_0$ and directions $\mathbf{u}$, $\mathbf{u}_0$. Removing these functions from the integrands in formulas (5.6) and (5.7) and taking into account expressions (3.9) and (3.10) for intensities, we can write stationary condition (5.5) in the form

$$\frac{I_V(\mathbf{r}, \mathbf{u})}{p_V/k_{\text{abs}}} g_V(\mathbf{r}, \mathbf{u}) + \frac{I_S(\mathbf{r}, \mathbf{u})}{I_S} g_S(\mathbf{r}, \mathbf{u}) = 1. \quad (5.16)$$

This relation connects the components of scalar intensity I.

2. Let I_S differ from zero and be constant only on radiating surface S_{rad}. In this case, we can use relations (5.8) and (5.9) for $D_S(\mathbf{r}, -\mathbf{u})$ and obtain, instead of relation (5.16),

$$\frac{I_V(\mathbf{r}, \mathbf{u})}{p_V/k_{\text{abs}}} g_V(\mathbf{r}, \mathbf{u}) + \frac{I_{S_{\text{rad}}, \text{scat}}(\mathbf{r}, \mathbf{u})}{I_S} g_S(\mathbf{r}, \mathbf{u}) + D_{S, \text{ex}}(\mathbf{r}, -\mathbf{u}) + \exp[-t(\mathbf{r} \longrightarrow \mathbf{r}_0)] = 1. \quad (5.17)$$

The meaning of the terms and corresponding subscripts is the same as in relations (5.8) and (5.9).

3. Let a scattering and radiating medium be a mixture of k different components, and let the following relation hold:

$$p_V(\mathbf{r}, \mathbf{u}) = \sum_k \mathbf{p}_k(\mathbf{r}, \mathbf{u}),$$
$$k_{\text{abs}}(\mathbf{r}) = \sum_k k_{\text{abs}, k}(\mathbf{r}), \quad (5.18)$$
$$k_{\text{ext}}(\mathbf{r}) = \sum_k k_{\text{ext}, k}(\mathbf{r}).$$

Analogously to Eq. (3.4), we can write the vector of the primary source of the kth component, assuming that the polarization of primary radiation is the same for all the components:

$$\mathbf{p}_k(\mathbf{r}, \mathbf{u}) = p_k(\mathbf{r}, \mathbf{u})[1, q_p(\mathbf{r}, \mathbf{u}), u_p(\mathbf{r}, \mathbf{u}), v_p(\mathbf{r}, \mathbf{u})]^T. \quad (5.19)$$

The scalar partial intensities i_k and $i_{k\uparrow\downarrow}$ for each component can be written analogously to expressions (3.11) and (4.13), where p_V is replaced by p_k. The absorption of radiation in the vicinity of point $(\mathbf{r}', -\mathbf{u}')$ at individual components is summed in accordance with relations (5.18). Consequently, instead of relation (4.14), we obtain the following expression for the relative fraction of the primary radiation of the kth component of the mixture, which emerges at point $(\mathbf{r}, -\mathbf{u})$ and is absorbed at point $(\mathbf{r}', -\mathbf{u}')$:

$$D_{V, k}[(\mathbf{r}, -\mathbf{u}) \longrightarrow (\mathbf{r}', -\mathbf{u}')] = \frac{i_k[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]}{p_k(\mathbf{r}', \mathbf{u}')} \times \sum_k k_{\text{abs}, k}(\mathbf{r}') g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]. \quad (5.20)$$

This expression is written analogously to Eq. (5.1). Here, g_V is defined by formula (5.2).

Integrating Eq. (5.20) over volume and directions, we obtain

$$D_{V, k}(\mathbf{r}, -\mathbf{u}) = \int_V d^3\mathbf{r}' \int_{4\pi} d(-\mathbf{u}) \frac{i_k[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]}{p_k(\mathbf{r}', \mathbf{u}')} \times \sum_k k_{\text{abs}, k}(\mathbf{r}') g_V[(\mathbf{r}', \mathbf{u}') \longrightarrow (\mathbf{r}, \mathbf{u})]. \quad (5.21)$$

This expression can be used instead of D_V in the relation (5.5) connecting the intensity components.

Radiation passing through the emitter interacts with all components of the medium; consequently, the elements of the Green matrices, as well as functions g_V and g_S, depend on the optical properties of all these components.

5.3. Possible Applications of Stationary Conditions and Relations

The stationary condition (4.6) (or (4.16)), including all elements of the Green matrices, is independent of the shape or size of the emitter. This condition remains valid irrespective of the radiation frequency, characteristics of radiation scattering and absorption, or their variation in the medium. On the other hand, the Green functions depend considerably on all these quantities. The stationary condition is exact when the linear theory of radiation transport is applicable and the main assumptions concerning the interaction of radiation with the medium hold (scattering is elastic, the absorption coefficients are independent of the direction and polarization of radiation, and the refractive medium of the medium remains equal to unity between acts of scattering). It follows hence that condition (4.6) can be useful for determining the Green functions. Indeed, the matrices of the Green functions are calculated for various specific emitters (see, for example, [14–17, 19]). In spite of numerous specific circumstances, the necessary stationary condition must hold. This corresponds to any region and direction of radiation in the emitter; preset primary sources of radiation can also be arbitrary and may be located in the bulk of volume V or on surface S. If the characteristics of radiation polarization are specified in a region of the object as well as the absorption coefficients in the emitter, expression (4.6) combined with (4.4) and (4.5) describes the relation between matrix elements of the Green functions. Since these elements are determined only by the optical parameters of the medium and affect the polarization, remaining independent of the polarization of the radiation being analyzed, one can specify various characteristics of polarization at point $(\mathbf{r}, \mathbf{u})$, thus separating individual matrix elements in accordance with formulas (4.4) and (4.5). For example, for natural radiation, only elements G_{II} will be left in the expressions for D_V and D_S, while term G_{IQ} is added in the case of plane-polarized radiation. The stationary condition can be used in different ways in different problems; in particular, the correctness of computation of the Green function can be verified with the help of this condition analogously to [3], where computations of radiation were verified without taking polarization into account.

Relations (5.5)–(5.7), (5.16), and (5.17) include quantities that have been investigated in many experiments. In spectroscopic experiments, the total intensities of radiation and some of their components are measured at different frequencies. The values of $D_{S,\,\mathrm{ex}}$ can also be measured in some cases, e.g., when radiation in a medium illuminated from outside is considered. In order to determine the value of $D_{S,\,\mathrm{ex}}$ in this case, it is necessary to obtain the ratio of the flux emerging in all directions from the medium to the incident flux [3]. It was mentioned above that optical density t also must be measured.

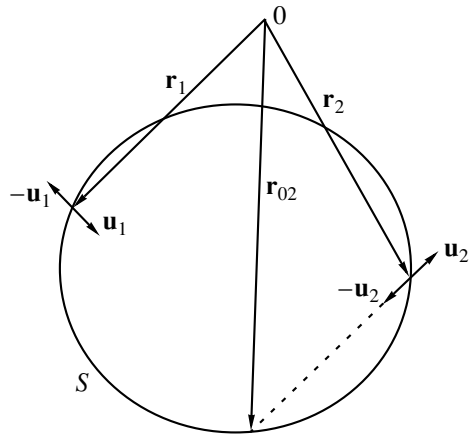

Fig. 3. Schematic diagram of experiments studying the effect of polarization on the passage of radiation through a scattering medium: $\mathbf{r}_1$ and $\mathbf{r}_2$ are the radius vectors of the regions of external illumination and measurements on surface S; $\mathbf{u}_1$, $-\mathbf{u}_1$, $\mathbf{u}_2$, and $-\mathbf{u}_2$ are the unit vectors determining the directions of illumination and measurements; $\mathbf{r}_{02}$ is the radius vector of the point at which radiation from $(\mathbf{r}_2, -\mathbf{u}_2)$ can arrive without being scattered.

On the other hand, such quantities as ratios p_V/k_{abs} and characteristics of polarization of primary radiation together with the Green functions g_V and g_S are of interest in spectroscopic experiments, but cannot be measured easily. The stationary conditions may help in determining the values of these quantities.

It was mentioned above that relations (5.5) and (5.16) do not differ from those in which polarization is not taken into account if primary radiation is not polarized and $g = 1$ (see relations (5.2) and (5.4). It follows hence that all the applications of the relations connecting intensity components [3] can also be used here. Among such applications, we can mention the determination of individual unknown terms to the resultant radiation from other known components and determination of relative populations of the levels or corresponding temperatures of the emitter when all radiation components are known. Individual terms were determined in real experiments where the effect of macroparticles on the profiles of atomic spectral lines in a dusty plasma was studied. The temperature of strongly scattering solid porous materials was measured using the stationary relations between the terms. In this case, the theory of radiation transport is inapplicable, but (see Introduction) steady-state relations between the radiation components were derived in this case also.

Let us consider a possible application of relations connecting the radiation components in the experimental study of the effect of polarized radiation on the passage of external radiation through a purely scattering medium surrounded by a transparent surface. Let us suppose that collimated external radiation can be directed onto the scattering medium alternatively from the neighborhoods of two arbitrarily chosen points $\mathbf{r}_1$ and $\mathbf{r}_2$ (Fig. 3). Let the direction of radiation in the vicinities of points $\mathbf{r}_1$ and $\mathbf{r}_2$ be defined by the vectors $\mathbf{u}_1$ and $-\mathbf{u}_2$, respectively. In the same regions, we can also measure the scalar intensity of radiation emerging from the medium (in particular, in directions $-\mathbf{u}_1$ and $\mathbf{u}_2$, respectively). The regions of illumination and observation can be interchanged.

Let us suppose that a neighborhood of point $(\mathbf{r}_1, \mathbf{u}_1)$ is illuminated by arbitrarily polarized radiation of scalar intensity I_1, the scalar intensity of scattered radiation measured in a neighborhood of point $(\mathbf{r}_2, \mathbf{u}_2)$ being I_2. We assume that scalar intensities and polarization do not change in the regions of illumination and measurements. If the regions of illumination and observation are not very large, function g_S in these regions are also constant. In this case, we can use formula (5.17). For $I_V = 0$, we obtain

$$\frac{I_2}{I_1} g_S[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)] + D_{S,\,\mathrm{ex}}(\mathbf{r}_2, -\mathbf{u}_2) + \exp[-t(\mathbf{r}_2 \longrightarrow \mathbf{r}_{02})] = 1. \tag{5.22}$$

This relation directly contains the quantity $D_{S,\,\mathrm{ex}}$ describing the passage of scattered radiation through the medium. In order to determine the value of this quantity, we must measure intensity ratio I_2/I_1 and optical density $t(\mathbf{r}_2 \longrightarrow \mathbf{r}_{02})$. In addition, we must know the polarization characteristic g_S. In order to calculate the value of this quantity using formula (5.4), it is sufficient to measure the same scalar intensities, changing the type of polarization of external radiation and using illumination at both points $\mathbf{r}_1$ and $\mathbf{r}_2$. Indeed, we can write the expression for the observable scalar intensity using the solution to transport equations (3.10), (3.12) and the above assumptions. In the case when point $(\mathbf{r}_1, \mathbf{u}_1)$ is illuminated, we have

$$\begin{aligned} I_2 = I_1\{&G_{II}[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)] \\ &+ G_{IQ}[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)] q(\mathbf{r}_1, \mathbf{u}_1) \\ &+ G_{IU}[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)] u(\mathbf{r}_1, \mathbf{u}_1) \\ &+ G_{IV}[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)] v(\mathbf{r}_1, \mathbf{u}_1)\}. \end{aligned} \tag{5.23}$$

Let the external radiation not be polarized at point $(\mathbf{r}_1, \mathbf{u}_1)$; in this case, we have

$$\mathbf{I}_1(\mathbf{r}_1, \mathbf{u}_1) = I_0[1, 0, 0, 0]^T. \tag{5.24}$$

Only the first term then remains on the right-hand side of Eq. (5.23), and we can use the measured intensities to determine the element

$$G_{II}[(\mathbf{r}_1, \mathbf{u}_1) \longrightarrow (\mathbf{r}_2, \mathbf{u}_2)].$$

Let the external radiation at point $(\mathbf{r}_1, \mathbf{u}_1)$ be plane-polarized (e.g., produced by a laser) and

$$\mathbf{I}_1(\mathbf{r}_1, \mathbf{u}_1) = I_1(\mathbf{r}_1, \mathbf{u}_1)[1, 1, 0, 0]^T. \tag{5.25}$$

In this case, two terms remain on the right-hand side of Eq. (5.23). Since the first element G_{II} has already been determined, we can find G_{IQ} by measuring scalar intensities. Thus, by complicating the polarization of incident radiation, it is also possible to find the remaining two elements.

In exactly the same way, it is possible to measure the elements of the Green matrix, which describe the passage of radiation in the opposite direction (from point $(\mathbf{r}_2, -\mathbf{u}_2)$ to point $(\mathbf{r}_1, -\mathbf{u}_1)$).

The above procedure for measuring scalar intensities makes it possible to determine the required matrix elements of the Green function using formula (5.23) and to analyze the effect of polarization on the passage of radiation through a scattering object using steady-state relation (5.22).

6. CONCLUSIONS

1. The necessary stationary condition (4.4)–(4.6) or (4.14)–(4.16) for a scattering medium has been derived from an analysis of extinction of radiation as a result of absorption of radiation by the medium and its emergence at the emitter surface. The stationary condition is derived under the assumption that the linear transport theory is valid for describing radiation and its extinction as well as the following assumptions concerning the interaction of radiation with a substance: scattering is elastic, the absorption and extinction coefficients of the medium are independent of the direction of radiation and its polarization, and the refractive index of the medium between scattering centers is equal to unity.

2. The stationary condition describes the relation between the elements in the first rows of the matrices of the Green functions for given polarization characteristics of radiation whose extinction is being analyzed and for the preset absorption coefficients of the medium. The condition is valid for an arbitrary radiation frequency.

3. Relations between the scalar intensity component of steady-state radiation are derived in integral form (5.5)–(5.9). The function taking into account the effect of polarization of radiation is separated as a factor in the integrands.

4. Special cases of the general relation connecting the scalar intensity components have been considered

for different characteristics of the polarization of primary radiation, when the reciprocity relations hold, and in the case of a multicomponent medium (see Sections 5.1 and 5.2). Useful relations have been derived in algebraic form in the case when the primary sources of radiation do not change in the bulk of the emitter and on radiating parts of its surface (see (5.16), (5.17), and (5.22)).

5. It is shown that the stationary conditions and the relations between intensity components in the case when primary radiation is not polarized do not differ from those in the absence of polarization. Consequently, the stationary conditions and the relations can be used in computations and experimental studies in the same way as in the absence of polarization, although radiation can be polarized as a result of scattering. When primary radiation is polarized, additional possibilities appear, in particular, for analyzing the matrix elements of the Green functions.

REFERENCES

1. S. Chandrasekhar, *Radiative Transfer* (Clarendon Press, Oxford, 1950; Inostrannaya Literatura, Moscow, 1953).
2. V. V. Sobolev, *A Treatise on Radiative Transfer* (Gostekhizdat, Moscow, 1956; Van Nostrand, Princeton, N.J., 1963).
3. I. A. Vasil'eva, Usp. Fiz. Nauk **171**, 1317 (2001) [Phys. Usp. **44**, 1255 (2001)].
4. I. A. Vasil'eva, J. Quant. Spectrosc. Radiat. Transf. **66**, 223 (2000).
5. I. A. Vasil'eva, Teplofiz. Vys. Temp. **38**, 274 (2000).
6. I. A. Vasilieva, Radiat. Eff. Defects Solids **154**, 61 (2001).
7. *Radiation Processes in Plasmas*, Ed. by G. Bekefi (Wiley, New York, 1966; Mir, Moscow, 1971).
8. L. A. Apresyan and Yu. A. Kravtsov, *Theory of Radiation Transfer: Statistical and Wave Aspects* (Nauka, Moscow, 1983).
9. H. C. van de Hulst, *Light Scattering by Small Particles* (Wiley, New York, 1957; Inostrannaya Literatura, Moscow, 1961).
10. J. W. Hovenier, J. Atmos. Sci. **26**, 488 (1969).
11. D. W. Mueller, Jr., and A. L. Crosbie, J. Quant. Spectrosc. Radiat. Transf. **57**, 81 (1997).
12. G. C. Pomraning, *The Equations of Radiation Hydrodynamics* (Pergamon, Oxford, 1973).
13. K. M. Case and P. F. Zweifel, *Linear Transport Theory* (Addison-Wesley, Reading, Mass., 1967; Mir, Moscow, 1972).
14. H. Domke and E. G. Yanovitskij, J. Quant. Spectrosc. Radiat. Transf. **26**, 389 (1981).
15. E. P. Zege and L. I. Chaikovskaya, J. Quant. Spectrosc. Radiat. Transf. **55**, 19 (1996).
16. E. P. Zege and L. I. Chaikovskaya, J. Quant. Spectrosc. Radiat. Transf. **64**, 413 (2000).
17. C. E. Siewert, J. Quant. Spectrosc. Radiat. Transf. **64**, 227 (2000).
18. C. F. Bohren and D. R. Nuffman, *Absorption and Scattering of Light by Small Particles* (Wiley, New York, 1983; Mir, Moscow, 1986).
19. N. Domke, J. Quant. Spectrosc. Radiat. Transf. **30**, 119 (1983).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 554–565.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 612–624.

PLASMA, GASES

Third-Harmonic Generation in an Ionized Gas and Its Relation to the Residual Energy of Electrons

N. E. Andreev, M. E. Veisman, and M. V. Chegotov*

High Energy Density Research Center, Joint Institute for High Temperatures, Russian Academy of Sciences, ul. Izhorskaya 13/19, Moscow, 125412 Russia

**e-mail: chegotov@ihed.ras.ru*

Received December 23, 2002

Abstract—The generation of low-order harmonics by a short ionizing laser pulse passing through a gas is investigated in regard to the space–time phase-synchronism conditions. This investigation is based on the results of hydrodynamic calculations for the model of [1] supplemented by taking into account the ionization current [2–5] and on the numerical solution of a one-dimensional time-dependent Schrödinger equation. As applied to the description of the third-harmonic spectrum, the hydrodynamic model with ionization current is shown to be in good agreement with a quantum-mechanical model. In this case, the amplitude of the third harmonic is determined by the intensity of the laser field at the moment of maximal ionization rate; this fact allows one to relate the amplitude of the third harmonic to the residual energy of electrons [5–8] and may provide grounds for the diagnosis of the residual energy by the spectrum of the third harmonic, which is important for the development of X-ray lasers based on ionization nonequilibrium plasma. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

In the last decade, the processes associated with the optical ionization of a gas by a short high-power laser pulse have been intensively studied (see, for example, [1–10] and references cited therein). In addition to its fundamental interest, this fact is related to the prospects of application of laser-frequency harmonics generated during these processes to the fabrication of short-wavelength radiation sources and the generation of attosecond electromagnetic pulses [11, 12], as well as to the possibility of using the information about the spectrum of these harmonics for the diagnosis of processes in an ionized gas.

The harmonics of a high-power laser pulse may be generated due to cooperative processes—bremsstrahlung of electrons as they collide with any ions generated by the plasma ionization [13]—as well as due to the individual nonlinear response of the "ionized electron + parent ion" system to the laser field. An efficient generation of harmonics by the first mechanism requires a high density of a substance, close to the solid-state density [14]; in this case, the efficiency of the bremsstrahlung mechanism can be increased due to the preliminary excitation of a gas ionized by a laser pulse [15]. In the present paper, we consider the generation of low-order harmonics in the field of a femtosecond laser pulse for a relatively low density of the ionized gas when the bremsstrahlung mechanism of the harmonic generation is neglected.

At present, the following viewpoint is widely adopted: the generation of high-order harmonics that manifest themselves as a plateau region in the transverse-field spectrum [10] occurs under the radiative recombination of an electron during its return motion in the field of the parent ion. To describe this process, both quantum-mechanical [10, 16] and semiclassical [17] methods have been applied. At present, there is no generally adopted viewpoint about the mechanisms of low-order harmonic generation whose total energy is much greater than the energy of high-order harmonics in the plateau region. It is assumed that the generation of low-order harmonics is associated with the bound–bound transitions of electrons between energy levels of an atom [18]. Moreover, there exists a simple hydrodynamic model proposed in [1] in which one considers harmonic generation due to the fact that the electron density increases under tunneling ionization two times within a laser period, at the moments corresponding to the maximal intensity of the laser field. This model is very attractive due to its simplicity and allows one to carry out a self-consistent analysis of the harmonic generation in regard to the propagation of a laser pulse to considerable distances. However, until recently, the question has remained unsolved as to how adequately the phenomenological model of [1] may describe a real physical situation.

To clear up this question, in Section 3, we investigate the generation of harmonics using both the hydrodynamic and quantum-mechanical models (which are described in Section 2). Moreover, we investigate the behavior of the amplitudes of the harmonics depending on the number of bound states in a potential well. In contrast to [18], where the authors investigated the

1063-7761/03/9703-0554$24.00

spectrum of the third-order time derivative of the dipole moment of an atom in the field of a given laser pulse, in the present paper, we investigate the spectrum of the generated radiation propagating in an ionized gas with regard to the space–time phase-synchronism conditions. The latter conditions are naturally taken into account in the self-consistent solution of the Schrödinger equation and the Maxwell equations. We show that the third harmonic of the carrier frequency of laser radiation is well described by the hydrodynamic model of [1] supplemented by taking into consideration the ionization current that guarantees the energy and momentum conservation laws of the laser pulse during the gas ionization [2–5]. This means that the amplitude of the third harmonic is fully determined by the parameters of the ionization front and the intensity of the laser field near the ionization front. In this case, one can associate the amplitude of the third harmonic with the residual energy of electrons. The latter is the energy that is irreversibly transferred from the laser field to electrons during their nonadiabatic interaction at the moment of ionization [6]. The magnitude of the residual energy determines the generation efficiency of coherent X-ray radiation in circuits with recombination pumping [8]. Therefore, the possibility to diagnose the residual energy of electrons by the third-harmonic spectrum is important for designing X-ray lasers with an active medium formed when a gas is ionized by an intense optical laser pulse. Note that three-dimensional propagation of the third harmonic in a nonlinear medium was investigated, in particular, in [19].

A model that associates the spectrum of the third harmonic with the residual energy of electrons is constructed in Section 4. The absorption of the residual energy by an electron generated by the gas ionization, just like the harmonic generation, is, strictly speaking, a quantum-mechanical process. Therefore, the comparison presented in Section 4 between the hydrodynamic formulas [6] used for calculating the residual energy and the quantum-mechanical calculation of the residual energy is of interest.

2. BASIC EQUATIONS

In the most general form within classical electrodynamics, the generation of laser-radiation harmonics in an ionized gas can be described by the Maxwell equations

$$\mathrm{rot}\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{B}}{\partial t}, \quad \mathrm{rot}\mathbf{B} = \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{c}\frac{\partial \mathcal{P}}{\partial t} \qquad (1)$$

and the corresponding equations for the dipole moment of electrons in a unit volume of the gas,

$$\mathcal{P} = -e\sum_j \langle \mathbf{r}_j \rangle$$

(where the sum is taken over all electrons in this volume; the angular brackets denote averaging over the wave functions of these electrons, $\langle \hat{f} \rangle = \langle \Psi | \langle \hat{f} \rangle | \Psi \rangle$, for the operator $\langle \hat{f} \rangle$; e is the absolute value of the electron charge; and the contribution of the motion of ions is neglected). In the approximation of a medium consisting of noninteracting one-electron atoms, we have

$$\mathcal{P}(\mathbf{r}, t) = -en_{\mathrm{at}}(\mathbf{r}, t)\int \mathbf{r}_\mathrm{a}|\Psi(\mathbf{r}, t, \mathbf{r}_\mathrm{a})|^2 d\mathbf{r}_\mathrm{a}. \qquad (2)$$

Here, n_{at} is the total concentration of atoms and ions, which is equal to the initial concentration of the gas (henceforth, we assume that it is independent of time); $\mathbf{r}_\mathrm{a}$ are internal atomic coordinates with the origin at an appropriate atomic nucleus; $\mathbf{r}$ is the radius vector of an atomic nucleus; Ψ is a one-electron wave function satisfying the Schrödinger equation

$$i\hbar\frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m}\Delta\Psi + [U(\mathbf{r}_\mathrm{a}) + \hat{V}_f(\mathbf{r}_\mathrm{a}, t)]\Psi, \qquad (3)$$

where the spatial differentiation is carried out with respect to the internal atomic coordinates; $U(\mathbf{r}_\mathrm{a})$ is the atomic potential; and $\hat{V}_f(\mathbf{r}_\mathrm{a}, t)$ is the operator of interaction between electrons and the field.

From formulas (1) and (2), we obtain the following wave equation for the dimensionless high-frequency field $\mathbf{a} = e\mathbf{E}/m\omega_0 c$:

$$\left(\frac{\partial^2}{\partial\eta^2} + \frac{2\partial^2}{\partial\eta\partial\xi} + \Delta_\perp\right)\mathbf{a} = N\mathbf{R}, \qquad (4)$$

$$\mathbf{R} = -a_{\mathrm{a.u.}}\frac{\partial^2\langle \mathbf{r}_\mathrm{a}\rangle}{\partial t_\mathrm{a}^2}. \qquad (5)$$

Here, $\xi = k_0(x - ct)$ and $\eta = k_0 x$ are coordinates tied to the laser pulse (the laser pulse propagates along the x axis),

$$\Delta_\perp = \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

ω_0 and $k_0 = \omega_0/c$ are the frequency and the wave vector of laser radiation in vacuum, $N \equiv n_{\mathrm{at}}/n_\mathrm{c}$,

$$n_\mathrm{c} = \frac{m\omega_0^2}{4\pi e^2}$$

is the critical concentration,

$$a_{\mathrm{a.u.}} \equiv \frac{eE_{\mathrm{a.u.}}}{m\omega_0 c}, \quad E_{\mathrm{a.u.}} = \frac{m^2 e^5}{\hbar^4}$$

is the atomic field strength, t_a is time in atomic units,

$$\frac{\partial}{\partial t_a} = -\frac{\omega_0}{\omega_{a.u.}}\frac{\partial}{\partial \xi} \text{ and } \omega_{a.u.} = \frac{me^4}{\hbar^3}$$

is the atomic frequency, and the source **R** is expressed in atomic units as $e = \hbar = m = 1$, $c = 137$. Using the fact that, according to (3),

$$\frac{\partial \langle \mathbf{r}_a \rangle}{\partial t_a} = \langle \hat{\mathbf{p}} \rangle + i\langle [\hat{V}_f, \mathbf{r}_a] \rangle,$$

where $\hat{\mathbf{p}} = -i\nabla$ is the operator of canonical momentum, and taking into account the explicit form of the operator $\hat{V}_f$ in the length calibration, in the dipole approximation (neglecting the dependence of the field on the coordinate $\mathbf{r}_a$), we obtain

$$\mathbf{R} = -a_{a.u.}\frac{\partial}{\partial t_a}\langle \hat{\mathbf{p}} \rangle. \quad (6)$$

Using Eq. (3), one can easily reduce expression (6) to the form (in the dipole approximation)

$$\mathbf{R} = \mathbf{a} + a_{a.u.}\langle \nabla U(\mathbf{r}_a) \rangle. \quad (7)$$

Thus, according to (5), (6), and (7), to calculate the source **R**, one has to determine any of the three following quantities: $\langle \mathbf{r}_a \rangle$, $\langle \hat{\mathbf{p}} \rangle$, or $\langle \nabla U(\mathbf{r}_a) \rangle$. This problem can be realized in the most consistent form (under a minimal number of assumptions) by numerically solving the Schrödinger equation (3). Recently, various methods have been developed for such a solution [20–23]; however, this problem is still laborious, especially as one has to solve it repeatedly to integrate wave equation (4). Therefore, approximate methods for calculating *R*, such as the numerical solution of a one-dimensional Schrödinger equation for a one-dimensional model atom [24, 25], as well as the approximate analytical calculation of $\langle \mathbf{r}_a \rangle$ [16] and $\partial\langle \hat{\mathbf{p}} \rangle/\partial t_a$ [10] under different assumptions, have become important. Among these assumptions, the basic ones are the following: first, the assumption that the binding potential weakly affects the part of the wave function that corresponds to the continuous spectrum; second, the assumption that all bound states, except the ground state, are neglected; and, third, the assumption that the depletion of the ground state and its Stark broadening are not substantial. Note that the assumption about the insignificant role of the binding potential in the continuous-spectrum region suggests that, strictly speaking, the corresponding analytic theories can hardly be applied to the low-order harmonics considered in this paper, with numbers smaller than $U_1/\hbar w_0$ [16], where U_1 is the energy of the ground state of the binding potential. Another reason for the possible inaccuracy of analytic quantum-mechanical theories as applied to the calculation of the third harmonic associated with the third above assumption is discussed in [26].

The difficulties associated with the calculation of the wave function of an electron and the subsequent calculation of the source **R** in one of the forms (5)–(7) seem inevitable in the case of high-order harmonics that manifest themselves as a plateau region in the transverse-field spectrum [10]. These harmonics are generated due to a radiative recombination of the electron during its return motion in the field of the parent ion [10, 16]. At the same time, we show below that the low-frequency part of the transverse-field spectrum that contains the third harmonic and in which the greater part of the field energy (except for the carrier-frequency energy) is concentrated can be described much more simply. To this end, in the tunneling-ionization regime considered below [27], which is realized in the case of sufficiently short and intense laser pulses [28], it suffices to know the instantaneous ionization frequency, for example, in the form given in [29]. The idea of such a description was first put forward in [1]. Below, we will show that the phenomenological model presented in [1] can well describe the low-frequency part of the spectrum provided that one takes into account—in addition to the free-electron current $\mathbf{J} = -en_e\mathbf{V}$, where n_e is the concentration and **V** is the hydrodynamic velocity of electrons—the "ionization" current associated with the loss of energy and momentum of laser radiation due to the gas ionization [2–4]:

$$\mathbf{J}_{ion} = \frac{\mathbf{E}}{|\mathbf{E}|^2}\sum_{z=1}^{z_n} W_z U_z n_{z-1}. \quad (8)$$

Here, U_z is the ionization potential of ions with the ionization multiplicity $z - 1$, n_{z-1} is the concentration of these ions, W_z is the ionization frequency, and z_n is the charge of the nucleus. In this case, the derivative of the polarizability is defined as a sum of currents $\mathbf{J} + \mathbf{J}_{ion}$. As a result, according to Eqs. (1) and definition (4), the source **R** is calculated as

$$\mathbf{R} = -\frac{1}{cn_{at}e}\frac{\partial(\mathbf{J} + \mathbf{J}_{ion})}{\partial \xi} \quad (9)$$

(in dimensionless coordinates). Thus, to determine **R** within this model (which is called a hydrodynamic model in what follows), it suffices to find the concentrations n_{z-1}, the ionization frequencies W_z, and the hydrodynamic velocity **V**. The electron concentration is defined as

$$n_e = \sum_{z=1}^{z_n} z n_z$$

under the assumption that it does not differ from its quasineutral value. In the weak-relativity approximation, the velocity **V** is determined from the equation [6, 7, 30]

$$\frac{\partial \mathbf{V}}{\partial t} = -\frac{e}{m}\mathbf{E} - \frac{\mathbf{V}S}{n_e}, \tag{10}$$

where $S = n_{at}\partial Z/\partial t$,

$$Z = \frac{\sum_{z=1}^{z_n} z n_z}{n_{at}} \tag{11}$$

is the average charge of ions and S is the ionization rate. In the approximation of successive ionization of electron levels, S satisfies the equations

$$\begin{aligned} S &= \sum_{z=1}^{z_n} S^{\{z\}}, \quad S^{\{z\}} = W_z n_{z-1}, \\ \frac{\partial n_z}{\partial t} &= S^{\{z\}} - S^{\{z+1\}}, \quad z = 1, \ldots, z_n - 1, \\ \frac{\partial n_{z_n}}{\partial t} &= S^{\{z_n\}}. \end{aligned} \tag{12}$$

In the approximation of weak relativity and slowly (within a laser period) varying amplitude of the carrier signal, we can derive from Eqs. (8)–(11) the following final expression for **R** in the hydrodynamic model:

$$\mathbf{R} = Z\mathbf{a} - \frac{\alpha^2}{2}\frac{\partial}{\partial \xi}\left\{\mathbf{a}\sum_{z=1}^{z_n}\frac{U_z}{U_\mathrm{H}}\frac{n_{z-1}}{n_\mathrm{c}}\frac{W_z}{|\mathbf{a}|^2\omega_0}\right\}, \tag{13}$$

where $\alpha \equiv e^2/\hbar c$ and U_H is the ionization potential of hydrogen. The second term on the right-hand side of (13) is attributed to the ionization current. Formula (13) is supplemented by the formula for the ionization frequencies W_z. In the tunneling ionization regime considered here, when the Keldysh parameter [27] is

$$\gamma = \frac{\omega_0\sqrt{2mU_z}}{e|\mathbf{E}|} \ll 1,$$

we can apply the adiabatic approximation by substituting the modulus of the instantaneous field $|\mathbf{E}|$ into the formula for the ionization frequency [31]. In this case, we can represent W_z as

$$W_z = C_z\omega_{\mathrm{a.u.}}\left(\frac{a_{\mathrm{a.u.}}}{|\mathbf{a}|}\right)^{\beta_z} \times \exp\left\{-\frac{2}{3}\frac{a_{\mathrm{a.u.}}}{|\mathbf{a}|}\left(\frac{U_z}{U_\mathrm{H}}\right)^{3/2}\right\}, \tag{14}$$

where C_z and β_z are constants that depend on the shape of the atomic potential. Equations (11)–(14) fully define the problem of finding **R** in the hydrodynamic model.

3. COMPARISON OF HYDRODYNAMIC AND QUANTUM CALCULATIONS OF THE HARMONIC AMPLITUDES

Consider the problem concerning the applicability of the hydrodynamic model considered above by comparing the solutions of the wave equation with the hydrodynamic source (13) with those obtained with a quantum-mechanical source. As was mentioned above, analytical quantum-mechanical models for calculating **R** may not be sufficiently accurate in the low-frequency region of the spectrum. Therefore, it is expedient to compare, first of all, the solutions obtained with the hydrodynamic source (13) with the results of numerical quantum-mechanical calculations. To this end, it is convenient to use **R** in form (7).

Assuming that the transverse dimension σ_r of a laser pulse is sufficiently large and that the distance x travels in an ionized gas is sufficiently small, $x < k_0\sigma_r^2/2$, and taking into account that the radiation patterns of the harmonics are narrow [32], we solve the wave equation (4) in the one-dimensional geometry, neglecting the transverse Laplacian. We assume that the polarization of the laser pulse is linear and directed along the y axis, $\mathbf{a} = a\mathbf{e}_y$. The Schrödinger equation is also solved in the one-dimensional geometry for a model one-electron binding potential; here, we use the length calibration and the dipole approximation:

$$i\frac{\partial \Psi}{\partial t_\mathrm{a}} = -\frac{1}{2}\frac{\partial^2\Psi}{\partial y_\mathrm{a}^2} + \left(\frac{a}{a_{\mathrm{a.u.}}}\right)y_\mathrm{a}\Psi + U(y_\mathrm{a})\Psi. \tag{15}$$

We assume that the potential $U(y_\mathrm{a})$ is short-range,

$$U(y_\mathrm{a}) = -2u\exp\left(-\frac{y_\mathrm{a}^2}{\sigma_y^2}\right), \tag{16}$$

where the quantities $\sigma_y \sim 1$ and $u \sim 1$ are chosen so that the ground state of energy coincides with the corresponding atomic level from which the ionization occurs. For potential (16), the constants β_1 and C_1 in (14) are equal to $\beta_1 = 0$ and

$$C_1 = \sqrt{\frac{U_1}{U_\mathrm{H}}}\left[\lim_{y_\mathrm{a}\to\infty}\Psi_1(y_\mathrm{a})\exp\left\{|y_\mathrm{a}|\sqrt{\frac{U_1}{U_\mathrm{H}}}\right\}\right]^2, \tag{17}$$

respectively, where $\Psi_1(y_\mathrm{a}) = \Psi(y_\mathrm{a}, t = 0)$ is the wave function of the bound state. In addition to the simplicity

arguments, such a choice of the potential is associated with the fact that, for moderately small gas concentrations n_{at}, the oscillation amplitude $2r_E = 2eE/m\omega_0^2$ of an electron in the field of laser radiation may be so large that the effect of a neighboring atomic residual is identical to the effect of the parent atom; to this end, $2r_E$ must be greater than half the distance between neighboring atoms $n_{at}^{-1/3}/2$, which is achieved for

$$0.3\frac{I(\xi_z)}{10^{14}\ \mathrm{W/cm^2}}\left(\frac{\lambda_0}{1\ \mu\mathrm{m}}\right)^4\left(\frac{n_{at}}{10^{18}\ \mathrm{cm^{-3}}}\right)^{2/3} > 1,$$

where $I(\xi_z)$ is the laser intensity at the moment of maximal ionization rate. Thus, even for sufficiently small intensities of laser radiation and moderate densities of a gas, the potential of an ionized atom cannot be considered as the potential of an isolated atom. For instance, this means that part of the energy levels of excited states of an isolated atom that condense toward the boundary of the continuum of states of a free electron turn into a continuous spectrum. Thus, the potential well of the atomic residual contains only a finite number of energy levels of bound states, as is exactly the case for a short-range potential.

In what follows, we apply potential (16) to investigate the dependence of the spectrum of radiation propagating in an ionized gas on the number of bound states in the atomic potential well. The number of bound states may be varied by changing u and σ_y so that, in addition to the bound state with the ground-state energy, the potential well will contain two, three, etc., energy levels. In this case, the constants σ_y and u are chosen so that the energy of the ground state is constant and equal to $U_1 = 24.6$ eV, which corresponds to the potential of single ionization of helium from the ground state.

For $\sigma_y = 1$ and $u = 0.807$, the ground state with $U_1 = 24.6$ eV proves to be the only state in the potential well. For this case, Fig. 1a represents the calculated values of the dimensionless ionization rate

$$\frac{S}{n_{at}\omega_0} = \omega_0^{-1}\frac{\partial Z}{\partial t},$$

the Keldysh parameter γ, and the average ion charge Z, which is determined by formula (11) in hydrodynamic calculations and by the formula

$$Z = 1 - \sum_n |\langle \Psi | n\rangle|^2 \qquad (18)$$

in quantum-mechanical calculations; here, $|n\rangle$ is the wave function of the nth bound state. Figure 1b shows the calculated spectrum of a transverse electromagnetic field normalized by its maximum, $|\hat{a}(k)|^2$, and

$$\hat{a}(k) = \int_{-\infty}^{\infty} a(\xi)\exp(-ik\xi)d\xi$$

is the Fourier image of the high-frequency field. Figure 1a shows that, in the case of tunneling ionization ($\gamma \approx 0.2$ in the region of maximal ionization rate), Eqs. (12), (14), (11) well describe the time variation of the average charge of ions. At the same time, the diagram of the ionization rate plotted versus the coordinate $(x - ct)/\lambda_0$ related to the momentum, which is obtained from the Schrödinger equation (15), shows additional peaks as compared with the calculations by formulas (12) and (14). These peaks are attributed to the processes of trapping of electrons generated due to the gas ionization by the parent ions (which manifest themselves at negative values of the ionization rate S (see Fig. 1a)) and to the nonadiabatic response of an atomic dipole to the rapidly changing electric field [33].

A comparison of the harmonic spectra (see Fig. 1b) obtained by Eqs. (4), (7), and (15) with the results of hydrodynamic calculations by formulas (4) and (11)–(14) both with and without regard for the ionization current (in the latter case, the second term in (13) is set to zero) shows that the source (13) with the ionization current well describes the amplitudes of the third and fifth harmonics and, satisfactorily, the amplitude of the seventh harmonic. Moreover, this source well describes not only the amplitude but also the shape of the third-harmonic spectrum. Higher order harmonics calculated by the hydrodynamic model have substantially understated amplitude values compared with the quantum-mechanical results. This is associated with the fact that the basic generation mechanism of these harmonics is the electron trapping by the parent ions [10, 16], which is not taken into consideration within the hydrodynamic model and manifests itself in the aforementioned additional peaks on the $S(\xi)$ curve.

Further, we change σ_y and u so that additional energy levels appear in the potential well while the ground-state energy remains unchanged, $U_1 = 24.6$ eV. For $\sigma_y = 1.5$ and $u = 0.685$, an additional level with energy $U_2 = 4.85$ eV appears; for $\sigma_y = 2.4$ and $u = 0.602$, two additional levels appear with energies $U_2 = 10.4$ eV and $U_3 = 1.5$ eV; and, for $\sigma_y = 3.4$ and $u = 0.553$, three additional levels appear with energies $U_2 = 14.54$ eV, $U_3 = 6.63$ eV, and $U_4 = 1.35$ eV. The spectra corresponding to these values of σ_y and u are shown in Fig. 2. The constant C_1 in formula (14), calculated by (17), is approximately equal to 4, 8, 24, and 130 for the cases a, b, c, and d, respectively. Figure 2 shows that an increase in the number of bound states in the potential well leads to a substantial increase in the amplitudes

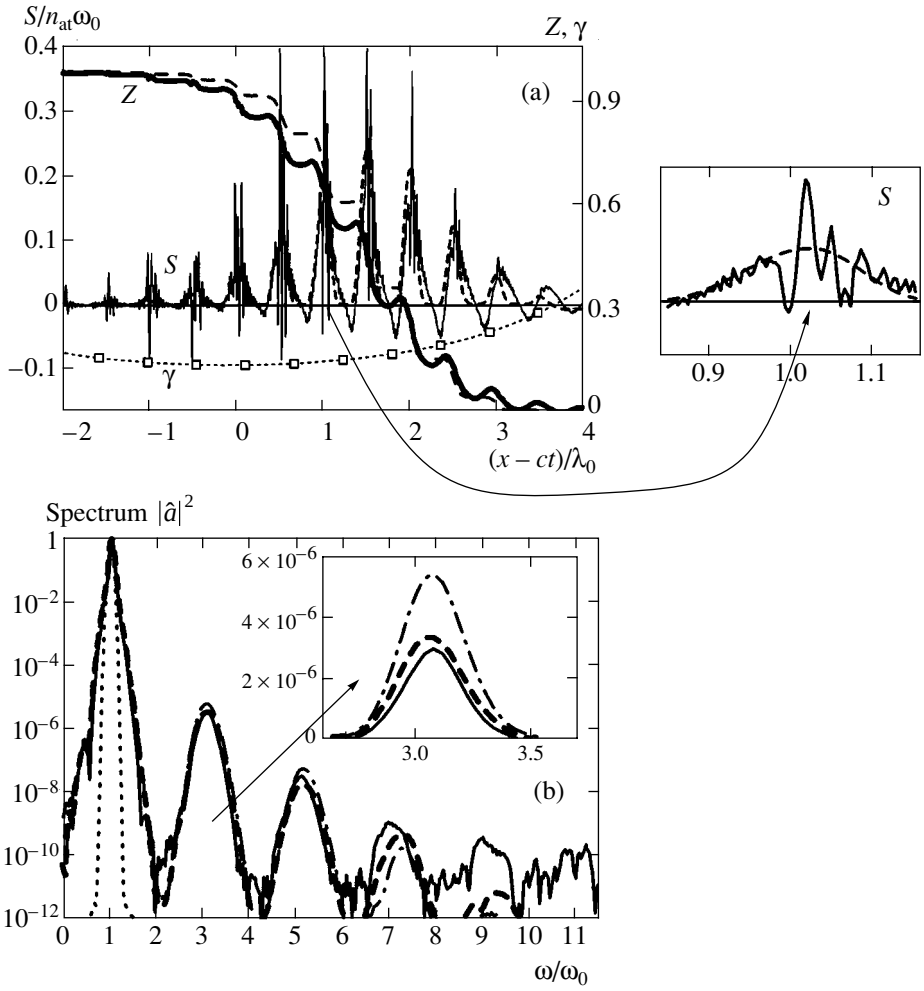

Fig. 1. (a) Dimensionless ionization rate $S(t)/n_{at}\omega_0$ and the average ion charge Z; solid lines represent the results obtained by solving the Schrödinger equation (15); the dashed lines represent the results of hydrodynamic calculations by formulas (11), (12), and (14); and the dotted line represents the Keldysh parameter γ. (b) The spectrum of transverse electromagnetic field radiated due to the propagation of a laser pulse through a gas layer of thickness $x = 6L_{FWHM}$. The solid line represents the result of calculation with the "quantum-mechanical" source (7), the dashed line represents the calculation with the "hydrodynamic" source (13), the dot-and-dash curve illustrates the hydrodynamic calculation without ionization current (the second term in (13) equals zero), and the dotted line represents the initial spectrum of the laser pulse before entering the gas. The parameters of the laser pulse are as follows: intensity $I_{max} = 10^{16}$ W/cm^2, a Gaussian envelope with a full width at half maximum $\tau_{FWHM} = L_{FWHM}/c = 13$ fs, and the wavelength $\lambda_0 = 0.8$ μm; the gas concentration is $n_{at}/n_c = 10^{-2}$, and the ionization potential of a single energy level is $U_1 = 24.6$ eV.

of the fifth, seventh, and higher order harmonics. The amplitude of the third harmonic depends weakly on the number of bound states in the potential well and is well described by the hydrodynamic model (with the ionization current). The amplitude of the fifth harmonic is also well described by the hydrodynamic model when there are no excited energy levels and transitions between them (Figs. 1 and 2a). However, in the presence of exited energy levels and transitions between them, the amplitudes of the fifth, seventh, and ninth harmonics calculated by the quantum-mechanical methods prove to be much greater than the results obtained by the hydrodynamic model. Note that the conclusion that the transitions between excited energy levels are important for the generation of such harmonics is in agreement with the results of [18].

4. THE RELATION BETWEEN THE THIRD-HARMONIC AMPLITUDE AND THE RESIDUAL ENERGY OF ELECTRONS

It follows from Section 3 (see Figs. 1 and 2) that the amplitude of the third harmonic of fundamental frequency ω_0 is well described by the hydrodynamic model. Hence, the amplitude of the third harmonic is fully determined by the laser-field intensity at the moment of the maximal ionization rate. However, the

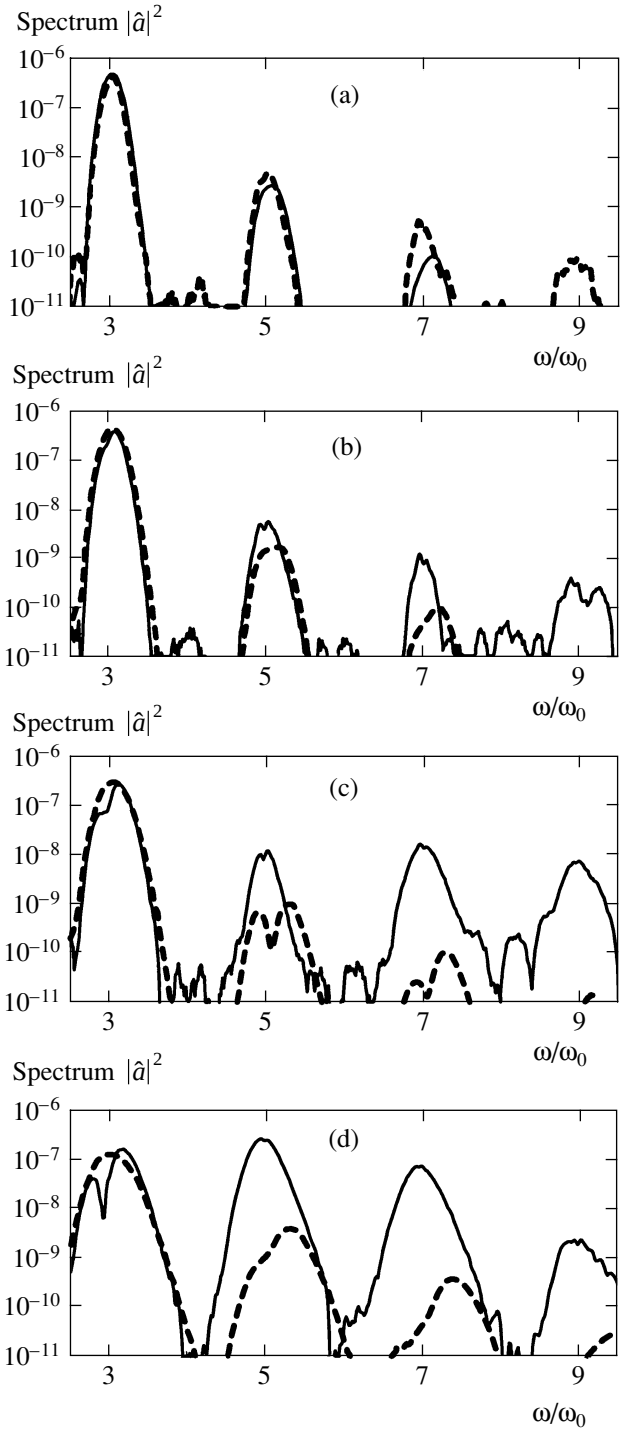

Fig. 2. Spectra of the third, fifth, seventh, and ninth harmonics for atomic potentials with different numbers of energy levels (U_1 = 24.6 eV): (a) with a single energy level U_1, (b) with additional energy level U_2 = 4.85 eV, (c) with two additional energy levels U_2 = 10.4 eV and U_3 = 1.5 eV; (d) with three additional energy levels U_2 = 14.54 eV, U_3 = 6.63 eV, and U_4 = 1.35 eV. The solid lines represent quantum-mechanical calculations, and the dashed lines represent hydrodynamic calculations. The spectra obtained after a laser pulse has passed through a gas layer of thickness $x = 2.2L_{\mathrm{FWHM}}$; τ_{FWHM} = 25 fs; other parameters are the same as in Fig. 1.

residual energy of electrons is also determined by the latter quantity [6]. This fact allows us to express the residual energy of electrons in terms of the amplitude of the third harmonic, which can be used for diagnosis of the residual energy by the third-harmonic spectrum.

For this purpose, we use the Fourier transformation of Eqs. (12) and (4) with source (13) with respect to ξ. Thus, we obtain the following equation for the Fourier image $\hat{a}(k)$ of the high-frequency field, expressed neglecting the second derivative with respect to η in Eq. (4):

$$2\frac{\partial}{\partial\eta}\hat{a}(k) = N\frac{1}{2\pi}\sum_{z=1}^{z_n}\int_{-\infty}^{\infty}\hat{a}(k-k') \times\left[\frac{\hat{S}^{(z)}(k')}{kk'} - \frac{\alpha^2}{2}\frac{U_z}{U_{\mathrm{H}}}\hat{\mathfrak{S}}^{(z)}(k')\right]dk', \quad k\neq 0, \tag{19}$$

where $\hat{S}^{(z)}(k)$ is the Fourier image of the dimensionless ionization rate

$$S^{(z)} \equiv \frac{S^{\{z\}}}{n_{\mathrm{at}}\omega_0}$$

and $\hat{\mathfrak{S}}^{(z)}$ is the Fourier image of $S^{(z)}/|a|^2$.

For further transformations of Eq. (19), we assume that the dimensionless distance $\eta = k_0 x$ traveled by a laser pulse in a gas satisfies the inequality $\eta \ll 6(n_{\mathrm{at}}/n_{\mathrm{c}})^{-1}$; this allows us to neglect the saturation phenomena of the third harmonic that are associated with the difference between the group velocities of the first and third harmonics and with the attenuation of the laser pulse during ionization. In this case, taking into account that the intensity of higher order harmonics is small compared with that of the fundamental harmonic, we obtain from (19) the following linear dependence of the third-harmonic intensity $\hat{a}(3)$ on η:

$$\hat{a}(3) = \frac{N\eta}{6}\frac{1}{2\pi}\sum_{z=1}^{z_n}\int_{2-\Delta_1}^{2+\Delta_1}\hat{a}(3-k') \times\left[\frac{\hat{S}^{(z)}(k')}{k'} - \frac{3\alpha^2}{2}\frac{U_z}{U_{\mathrm{H}}}\hat{\mathfrak{S}}^{(z)}(k')\right]dk', \tag{20}$$

where Δ_1 is the characteristic width of the first-harmonic spectrum, which is determined by the laser-pulse duration. Note that the term proportional to α^2 is attributed to the ionization current J_{ion}.

Formula (20) enables one to express the amplitude of the third harmonic of the laser field in terms of the spectrum of the ionization rate. As will be shown below, the residual energy Q_∞ of electrons can also be expressed in terms of this characteristic. Within the

hydrodynamic model, the residual energy can be represented as $Q_\infty = Q_{\text{fin}}(\xi \longrightarrow -\infty)$, where $Q_{\text{fin}}(\xi)$ is a part of the total energy that determines the fraction of energy irreversibly transferred from the laser field to the electrons generated during the gas ionization [6]. When the ionization occurs in the tunneling regime ($\gamma \ll 1$) and the intensities are nonrelativistic, according to [6, 7], $Q_{\text{fin}}(\xi)$ can be represented as

$$Q_{\text{fin}}(\xi) = \frac{\sum_{z=1}^{z_n} \kappa^{(z)}(\xi) Q_{\text{fin}}^{(z)}(\xi)}{\sum_{z=1}^{z_n} \kappa^{(z)}(\xi)}, \qquad (21)$$

$$Q_{\text{fin}}^{(z)}(\xi) = \frac{1}{\kappa^{(z)}(\xi)} \int_{-\infty}^{\xi} S^{(z)}(\xi^*) Q_e(\xi, \xi^*) d\xi^*.$$

Here, $Q_{\text{fin}}^{(z)}$ is the energy component (reduced to the dimensionless form by the coefficient mc^2 [7]) of a group of electrons that are generated by z-fold ionization, $\kappa^{(z)} = n_z/n_{\text{at}}$ is the degree of the z-fold ionization, and

$$Q_{\text{e}}(\xi, \xi^*) = \frac{1}{2}\left[\int_{\xi^*}^{\xi} a(\xi')d\xi'\right]^2$$

is the dimensionless energy obtained by the moment ξ by an electron generated at the moment ξ^*. Note that, just as in the case of the third harmonic, the hydrodynamic calculation of the energy fraction $Q_{\text{fin}}(\xi)$ is in good agreement with the results of the quantum-mechanical calculation of the total energy of electrons by the formula

$$Q(\xi) = \left\langle \Psi(\xi, y_{\text{a}}) \left| -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial y_{\text{a}}^2} \right| \Psi(\xi, y_{\text{a}}) \right\rangle$$

(see Fig. 3). Figure 3 shows that, everywhere except for the initial stage where the ionization occurs, the curve $Q_{\text{fin}}(\xi)$ represents the lower envelope of the curve $Q(\xi)$.

To establish a relation between Q_∞ and $\hat{a}(3)$, we rewrite expression (21) for

$$Q_\infty^{(z)} = Q_{\text{fin}}^{(z)}(\xi \longrightarrow -\infty)$$

as

$$Q_\infty^{(z)} = (2\pi\kappa_\infty^{(z)})^{-1} \int_{-\infty}^{\infty} \hat{Q}_{\text{e}}(-k')\hat{S}^{(z)}(k')dk', \qquad (22)$$

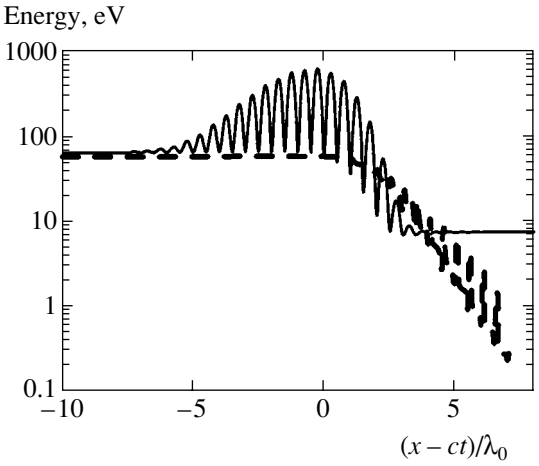

Fig. 3. Quantum-mechanical calculation of electron energy (solid line) and hydrodynamic calculation of the part of energy transferred from the field to electron (dashed line) for the parameters given in Fig. 1.

where $\hat{Q}_{\text{e}}(k)$ is the Fourier image of the energy $Q_{\text{e}}(-\infty, \xi^*)$ and $\kappa_\infty^{(z)} = \kappa^{(z)}(\xi \longrightarrow -\infty)$. To analytically determine the relation between Q_∞ and $\hat{a}(3)$, one has to simplify expressions (22) and (20). To this end, we consider the approximation of a given laser pulse:

$$a(\xi) = a_L(\xi)\cos\xi,$$

$$Q_{\text{e}}(\xi, \xi^*) \approx \frac{1}{2}a_L^2(\xi^*)\sin^2\xi^* + \frac{1}{2}a_L^2(\xi)\sin^2\xi \qquad (23)$$

$$- a_L(\xi^*)a_L(\xi)\sin\xi^*\sin\xi,$$

where $a_L(\xi)$ is the envelope. For $a(\xi)$ and $Q_{\text{e}}(\xi)$ in form (23) under the assumption that

$$k_0 L_{\text{FWHM}} \gg k_0 L_{S^{(z)}} \gg 1$$

(where $L_{S^{(z)}}$ is the characteristic width of the ionization front when z-fold ionized ions are being formed and L_{FWHM} is the full width at half maximum of the laser pulse), one can substitute the following approximations for the Fourier images $\hat{a}(k)$ and $\hat{Q}_{\text{e}}(k)$ into each summand in the sum over z in Eqs. (20) and (22):

$$\hat{a}(k) = a_L(\xi_z)2\pi\left[\frac{\delta(k-1)}{2} + \frac{\delta(k+1)}{2}\right],$$

$$\hat{Q}_e(k) = (1/4)a_L^2(\xi_z)2\pi\left[\delta(k) - \frac{\delta(k-2)}{2} - \frac{\delta(k+2)}{2}\right],$$

where ξ_z is the moment of maximum of the zth ionization front. As a result, we obtain the following algebraic

equations for $\hat{a}(3)$ and Q_∞ from (20) and (22):

$$Q_\infty = \frac{\sum_{z=1}^{z_n} \kappa_\infty^{(z)} Q_\infty^{(z)}}{\sum_{z=1}^{z_n} \kappa_\infty^{(z)}}, \qquad Q_\infty^{(z)} \approx \frac{1}{4} a_L^2(\xi_z)[1-\mu_2^{(z)}], \tag{24}$$

$$\hat{a}(3) \approx \frac{N\eta}{24} \sum_{z=1}^{z_n} a_L(\xi_z)\kappa_\infty^{(z)} \times \left[\mu_2^{(z)} + \frac{\mu_4^{(z)}}{2} - \frac{6\alpha^2}{a_L^2(\xi_z)} \frac{U_z}{U_\mathrm{H}} (2\mu_2^{(z)} - 1)\right], \tag{25}$$

where

$$\mu_{2l}^{(z)} \equiv \frac{\hat{S}^{(z)}(2l)}{\hat{S}^{(z)}(0)}, \quad l = 1, 2,$$

and $a_L(\xi_z)$ is the value of the envelope at $\xi = \xi_z$. When deriving (25), we used the equality

$$a_L^2(\xi_z) = [\hat{\mathfrak{S}}^{(z)}(2) + \hat{\mathfrak{S}}^{(z)}(4)] = 2[2\hat{S}^{(z)}(2) - \hat{S}^{(z)}(0)],$$

which is valid in the approximation of slowly varying amplitudes. Moreover, we took into account that

$$\hat{S}^{(z)}(0) = \int_{-\infty}^{\infty} S^{(z)}(\xi) d\xi = \kappa_\infty^{(z)}.$$

To calculate $\mu_{2l}^{(z)}$ for real multielectron atoms, we use a formula for the ionization frequencies W_z from [29], according to which the constants in (14) are equal to

$$\beta_z = 2n_{*,z} - 1, \quad C_z = \left(\frac{4e^1 z^3}{n_{*,z}^4}\right)^{\beta_z} \frac{e^1 z^2}{2\pi n_{*,z}^4},$$

where $e^1 \equiv \exp(1)$ and $n_{*,z} = z\sqrt{U_\mathrm{H}/U_z}$ is the principal quantum number. Next, consider the approximation of slowly varying amplitudes and take into account that the maximal contribution to the harmonics of the ionization rate are made by the neighborhood of its maximum point ξ_z. Then, we obtain the following expression for $\mu_{2l}^{(z)}$:

$$\mu_{2l}^{(z)} = \frac{\mathfrak{I}_{2l}^{(z)}}{\mathfrak{I}_0^{(z)}}, \quad \mathfrak{I}_{2l}^{(z)} = \frac{2}{\pi} \int_0^{\pi/2} \frac{\cos(2l\xi)}{(\cos^2\xi)^{n_{*,z}-1/2}} \times \exp\left\{-\frac{2}{\varepsilon_z(\cos^2\xi)^{1/2}}\right\} d\xi, \tag{26}$$

where

$$\varepsilon_z = \varepsilon_z(a_L(\xi_z)) = 3\frac{a_L(\xi_z)}{a_\mathrm{a.u.}}\left(\frac{n_{*,z}}{z}\right)^3$$

is a quantity that is small in the ionization region ($\varepsilon_z \lesssim 1/3$).

Using the fact that ε_z is small, we can obtain from (26) the following asymptotic expansion for $\mu_2^{(z)}$, which is written below up to terms on the order of ε_z^4:

$$\mu_2^{(z)} = 1 - \varepsilon_z + \varepsilon_z^2\left(-n_{*,z} + \frac{9}{4}\right) + \varepsilon_z^3\left(-n_{*,z}^2 + \frac{21}{4}n_{*,z} - \frac{207}{32}\right) + \varepsilon_z^4\left(-n_{*,z}^3 + 9n_{*,z}^2 - \frac{201}{8}n_{*,z} + \frac{1419}{64}\right). \tag{27}$$

A similar asymptotic series can also be written for $\mu_4^{(z)}$; however, this series converges more poorly than series (27). At the same time, the general expression (26) is inconvenient for computations. Therefore, it is expedient to apply the following approximate expression for $\mu_4^{(z)}$, which is based on the results of numerical calculations:

$$\mu_4^{(z)} \approx 0.4\mu_2^{(z)}. \tag{28}$$

According to (25) and (28), in the tunneling regime, $\mu_4^{(z)}$ makes a small contribution as compared with $\mu_2^{(z)}$ to the calculated value of $\hat{a}(3)$; therefore, the possible error in determining $\mu_4^{(z)}$ by formula (28) is insignificant.

Generally speaking, Eqs. (24)–(28) allow one to express Q_∞ in terms of $\hat{a}(3)$ only in the case of single ionization of a gas. In the case of multiple ionization, information about the amplitude of the third harmonic is insufficient for determining the residual energy; one also needs information on the spectrum of this harmonic; to this end, one has to consider the original sys-

tem of integral equations (21), (20). However, in this case, system (24)–(28) allows one, in principle, to determine the residual energy $Q_\infty^{(z_{max})}$ of a group of electrons that are generated during the formation of ions of maximal multiplicity of ionization. To this end, it is necessary that the ionization peaks corresponding to the ionization of various electron levels should be separated sufficiently clearly (as in the case of helium). To put it more precisely, it is necessary that the inequality

$$a_L(\xi_{z_{max}}) \gg a_L(\xi_{z_{max}-1})$$

should hold. Let us substitute the threshold field for the z_{max}-fold ionization instead of $a_L(\xi_{z_{max}})$ [34],

$$\frac{a_{z_{max}}}{a_{a.u.}} = \frac{(U_{z_{max}}/U_H)^2}{16 z_{max}}.$$

It is convenient to rewrite this inequality as

$$\left(\frac{U_{z_{max}}}{U_{z_{max}-1}}\right)^2 \frac{z_{max}-1}{z_{max}} \gg 1. \tag{29}$$

When inequality (29) holds, one can omit all the terms in (25) except for those proportional to $a_L(\xi_{z_{max}})$, which make the maximal contribution. Taking into account the numerical result (28) as well, we obtain

$$\hat{a}(3) \approx \frac{N\eta}{24} a_L(\xi_{z_{max}}) \kappa_\infty^{(z_{max})} \times \left[1.2\mu_2^{(z_{max})} - \frac{6\alpha^2}{a_L^2(\xi_{z_{max}})}\frac{U_z}{U_H}(2\mu_2^{(z_{max})} - 1)\right]. \tag{30}$$

Substituting expression (27) for $\mu_2^{(z_{max})}$ into Eq. (30), we obtain that the quantity $a_L(\xi_{z_{max}})$ can be expressed in terms of the ratio $\hat{a}(3)/N\eta\kappa_\infty^{(z_{max})}$.[1] This is achieved by calculating the quantity

$$\varepsilon = 3\frac{a_L(\xi_{z_{max}})}{a_{a.u.}}\left(\frac{n_{*,z_{max}}}{z_{max}}\right)^3,$$

which represents the root of the following equation derived from (30) and (27):

$$B + \varepsilon[A - 2B] - \varepsilon^2[C_\mu - 2BC_2] + \varepsilon^3[C_\mu - 2BC_3] - \varepsilon^4[C_\mu C_2 - 2BC_4] + \varepsilon^5[C_\mu C_3 - 2BC_5] + \ldots + (-1)^{n+1}\varepsilon^n[C_\mu C_{n-2} - 2BC_n] = 0 \tag{31}$$

with the coefficients

$$A = \frac{72}{N\eta\kappa_\infty^{(z_{max})}}\frac{\hat{a}(3)}{a_{a.u.}}\left(\frac{U_{z_{max}}}{U_H}\right)^{-3/2},$$

$$B = 54\left(\frac{\alpha}{\alpha_{a.u.}}\right)^2\left(\frac{U_{z_{max}}}{U_H}\right)^{-2}, \quad C_\mu \approx 1.2,$$

$$C_2 \approx \frac{9}{4} - n_{*,z}, \quad C_3 \approx \frac{207}{32} - \frac{21}{4}n_{*,z} + n_{*,z}^2, \tag{32}$$

$$C_4 \approx \frac{1419}{64} - \frac{201}{8}n_{*,z} + 9n_{*,z}^2 - n_{*,z}^3, \ldots.$$

Here, C_n, $n = 2, 3, \ldots$, are the coefficients of the asymptotic expansion

$$\mu_2^{(z_{max})} = 1 - \varepsilon_{z_{max}} + C_2\varepsilon_{z_{max}}^2 - C_3\varepsilon_{z_{max}}^3 + \ldots$$

(see (27)). Taking into account that ε is small, we can easily represent the solution to (31) as the power series

$$\varepsilon = \varepsilon_{(0)} + \varepsilon_{(1)} + \varepsilon_{(2)} + \ldots,$$

$$\varepsilon_{(0)} = \frac{D_1 + \sqrt{D_1^2 + 4BD_2}}{2D_2},$$

$$\varepsilon_{(1)} = \frac{\varepsilon_{(0)}^3 D_3}{2\varepsilon_{(0)}D_2 - D_1} \sim \varepsilon_{(0)}^2, \tag{33}$$

$$\varepsilon_{(2)} = \frac{3\varepsilon_{(0)}^2\varepsilon_{(1)}D_3 - \varepsilon_{(1)}^2 D_2 - \varepsilon_{(0)}^4 D_4}{2\varepsilon_{(0)}D_2 - D_1} \sim \varepsilon_{(0)}^2\ldots,$$

where D_n is the coefficient multiplying ε^n, $n = 1, 2, \ldots$, in Eq. (31).

Having determined the root ε and the quantity

$$a_L(\xi_{z_{max}}) = \frac{\varepsilon a_{a.u.}}{3}\left(\frac{z_{max}}{n_{*,z_{max}}}\right)^3 \tag{34}$$

from formulas (24) and (27), one can easily find the value $Q_\infty^{(z_{max})}$. The dependence of this value on the amplitude of the third harmonic will be determined by the dependence of the obtained value of ε on coefficient A (32). Equations (31)–(34), (24), and (27) show that the residual energy can be experimentally deter-

[1] The quantity η must not be too large in order that the saturation phenomena of the third harmonic should not manifest themselves.

Comparison of the residual energy (expressed in terms of $\hat{a}$ (3)) calculated by formulas (31)–(34), (24), and (27) with that obtained by formula (21)

Parameters of the problem						Data of hydrodynamic calculations of $\hat{a}$ (3)		$Q_\infty^{(z_{max})}$ calculated by formulas (31)–(34), (24), and (27)	$Q_\infty^{(z_{max})}$ calculated by formula (21)
z_{max}	$U_{z_{max}}$, eV	I_{max}, W/cm^2	λ_0, μm	τ_{FWHM}, fs	$\frac{n_{at}}{n_c}$	$\hat{a}$ (3)	$\frac{\eta/c}{\tau_{FWHM}}$		
1	13.6	10^{15}	0.8	13	10^{-2}	1.86×10^{-4}	3	8.2	6.6
1	13.6	10^{15}	1.08	50	10^{-2}	2.22×10^{-4}	2	6.7	7.0
1	24.6	5×10^{15}	0.8	13	10^{-2}	8.8×10^{-4}	3.0	40	42
1	24.6	5×10^{15}	0.8	50	10^{-2}	10.7×10^{-4}	1.3	25	27
2	54.4	4×10^{16}	0.8	13	5×10^{-3}	13.1×10^{-4}	2.1	215	205
2	54.4	4×10^{16}	0.8	50	5×10^{-3}	14.75×10^{-4}	0.66	172	127
5	97.9	4×10^{17}	0.8	20	2×10^{-3}	13.66×10^{-4}	1.5	484	305

mined in principle by the optical measurements of the third harmonic of laser radiation that generates the plasma.[2]

The comparison of the values of $Q_\infty^{(z_{max})}$ calculated by formulas (31)–(34), (24), and (27) with those obtained by formula (21) is shown in the table. These results are obtained in the case when terms up to $\varepsilon_{z_{max}}^3$ are taken into account in asymptotic series (27), while, when finding the root of polynomial (31), terms up to $\varepsilon_{(2)}$ are taken into account. These calculations show that, with respect to the amplitude of the third harmonic, the diagnosis of the residual energy of electrons ionized from internal energy levels satisfactorily (to an accuracy of no worse than 50%) agrees with the results of calculating the residual energy of this group of electrons by formula (21). Note that the best agreement between the above results is observed in the case of single ionization of helium because, on the one hand, formulas (24) and (25) contain only one term with $z = 1$ in this case and, on the other hand, the tunneling ionization condition $\gamma \ll 1$ is satisfied by a sufficiently large margin ($\gamma \lesssim 0.3$ in calculations). The accuracy of calculations for hydrogen is lower because, in this case, the parameters chosen are on the verge of the tunneling ionization regime ($\gamma \sim 0.5$). For doubly ionized helium, the approach proposed also allows one to describe satisfactorily the quantity $Q_\infty^{(z_{max})}$ since the ionization peaks corresponding to the successive ionization of the first and second electrons are sufficiently well separated,[3] which is also manifested in the fact that the left-hand side of inequality (29) is approximately equal to 2.5 in this case. For nitrogen, the accuracy of the approach proposed is somewhat lower because the appropriate ionization peaks in nitrogen are less separated (in this case, the quantity on the left-hand side of inequality (29) is approximately equal to 1.3).

5. CONCLUSIONS

Thus, we have shown that the spectrum of radiation due to the ionization of a gas by an intense femtosecond laser pulse strongly depends on the structure of atomic levels in the region of the harmonics with frequencies $\omega \geq 5\omega_0$. The intensity of these harmonics proves to be the higher, the greater the number of energy levels of bound states in the field of the atomic potential. The maximum of the spectral density and the shape of the line corresponding to the third harmonic of laser radiation are virtually independent of the intra-atomic structure of the energy levels and are determined by the ionization potential and the parameters of the laser pulse. This fact allows one to construct a simple hydrodynamic model using the quantum probabilities of ionization in the laser-pulse field to describe the third-harmonic generation due to ionization. Within this model, one can determine the residual energy of electrons generated due to ionization by the amplitude of the third-harmonic spectrum. This means that the third-harmonic spectrum can be used for the optical diagnosis of a substance during its ionization.

[2] The degree of ionization $\kappa_\infty^{(z_{max})}$ must be estimated from independent considerations. Below, we assume that $\kappa_\infty^{(z_{max})} = 1$.

[3] The separation of these peaks in the calculations presented in the fifth row of the table is more pronounced than those represented in the sixth row. This fact accounts for the higher accuracy of calculations represented in the fifth row.

ACKNOWLEDGEMENTS

This work was supported in part by the Russian Foundation for Basic Research (project no. 01-02-16723) and by NATO (PST CLG 979372).

REFERENCES

1. F. Brunel, J. Opt. Soc. Am. B **7**, 521 (1990).
2. N. E. Andreev, M. E. Veĭsman, M. G. Keĭdzhyan, and M. V. Chegotov, Fiz. Plazmy **26**, 1010 (2000) [Plasma Phys. Rep. **26**, 947 (2000)].
3. V. P. Kandidov, O. G. Kosareva, and S. A. Shlenov, Kvantovaya Élektron. (Moscow) **28**, 971 (1994).
4. M. Geissler, G. Tempea, A. Scrinzi, *et al.*, Phys. Rev. Lett. **83**, 2930 (1999).
5. V. B. Gildenburg, A. V. Kim, V. A. Krupnov, *et al.*, IEEE Trans. Plasma Sci. **21**, 34 (1993).
6. N. E. Andreev, M. E. Veĭsman, C. P. Goreslavskiĭ, and M. V. Chegotov, Fiz. Plazmy **27**, 296 (2001) [Plasma Phys. Rep. **27**, 278 (2001)]; N. E. Andreev, M. V. Chegotov, M. E. Veisman, *et al.*, Proc. SPIE **3735**, 234 (1998).
7. N. E. Andreev, M. V. Chegotov, M. E. Veĭsman, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **68**, 566 (1998) [JETP Lett. **68**, 592 (1998)].
8. P. Pulsifer, J. P. Apruzese, J. Davis, and P. Kepple, Phys. Rev. A **49**, 3958 (1994).
9. A. L'Huillier, L.-A. Lompre, G. Mainfray, and C. Manus, in *Atoms in Intense Laser Fields*, Ed. by M. Gavrila (Academic, New York, 1992), p. 139.
10. V. T. Platonenko and V. V. Strelkov, Kvantovaya Élektron. (Moscow) **25**, 582 (1998).
11. A. M. Sergeev, A. V. Kim, E. V. Vanin, *et al.*, Proc. SPIE **2770**, 36 (1995).
12. P. Antoine, A. L'Huiller, and M. Lewenstein, Phys. Rev. Lett. **77**, 1234 (1996).
13. V. P. Silin, Kvantovaya Élektron. (Moscow) **26**, 11 (1998).
14. K. N. Ovchinnikov and V. P. Silin, Kratk. Soobshch. Fiz., No. 10, 19 (1998).
15. V. P. Silin, Zh. Éksp. Teor. Fiz. **117**, 920 (2000) [JETP **90**, 805 (2000)].
16. M. Lewenstein, Ph. Balcou, M. Yu. Ivanov, *et al.*, Phys. Rev. A **49**, 2117 (1994).
17. P. B. Corcum, Phys. Rev. Lett. **71**, 1994 (1993).
18. C. F. de M. Faria, M. Dorr, and W. Sander, Phys. Rev. A **58**, 2990 (1998).
19. N. Aközbek, A. Iwasaki, A. Becker, *et al.*, Phys. Rev. Lett. **89**, 143901 (2002).
20. J. L. Krause, K. J. Schafer, and K. C. Kulander, Phys. Rev. A **45**, 4998 (1992).
21. L. Roso-Franko, A. Sanpera, M. L. Pons, and L. Plaja, Phys. Rev. A **44**, 4652 (1991).
22. A. M. Ermolaev and A. V. Selin, Phys. Rev. A **62**, 015401 (2000).
23. S. Blanes and P. C. Moan, Phys. Lett. A **265**, 35 (2000).
24. J. H. Eberly, Q. Su, and J. Javanainen, Izv. Akad. Nauk SSSR, Ser. Fiz. **53**, 1101 (1989).
25. E. V. Vanin, M. S. Dauner, A. V. Kim, and A. M. Sergeev, Pis'ma Zh. Éksp. Teor. Fiz. **58**, 964 (1993) [JETP Lett. **58**, 900 (1993)].
26. W. Becker, S. Long, and J. K. McIver, Phys. Rev. A **50**, 1540 (1994).
27. L. V. Keldysh, Zh. Éksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1964)].
28. M. V. Chegotov, Fiz. Plazmy (Moscow) **26**, 940 (2000) [Plasma Phys. Rep. **26**, 881 (2000)].
29. M. V. Ammosov, N. V. Delone, and V. P. Kraĭnov, Zh. Éksp. Teor. Fiz. **91**, 2008 (1986) [Sov. Phys. JETP **64**, 1191 (1986)].
30. P. Mulser, F. Cornolti, and D. Bauer, Phys. Plasmas **5**, 4466 (1998).
31. A. M. Perelomov, V. S. Popov, and M. V. Terent'ev, Zh. Éksp. Teor. Fiz. **50**, 1393 (1966) [Sov. Phys. JETP **23**, 924 (1966)].
32. N. E. Andreev and M. V. Chegotov, in *Abstracts of XXVII Zvenigorod Conference on Plasma Physics and UTS* (Zvenigorod, 2000), p. 119.
33. I. P. Christov, J. Zhou, J. Peatross, *et al.*, Phys. Rev. Lett. **77**, 1743 (1996).
34. N. B. Delone and V. P. Kraĭnov, Tr. Inst. Obshch. Fiz. Ross. Akad. Nauk **50**, 34 (1995).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 566–572.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 625–632.

FLUIDS

Modified Model of Self-Consistent Field for an Electron Solvated in a Polar Liquid

G. N. Chuev* and M. V. Fedorov**

Institute for Theoretical and Experimental Biophysics, Russian Academy of Sciences, pr. Nauki, Pushchino, Moscow oblast, 142290 Russia

**e-mail: genchuev@rambler.ru*

***e-mail: maximfedorov@rambler.ru*

Received January 21, 2003

Abstract—The self-consistent field model is modified to take into account the effect of orientation correlations on the behavior of an electron solvated in a polar liquid. This model is used as the basis for calculating the maximum of the absorption spectrum as well as the temperature dependence of this maximum for an electron solvated in water and ammonium. The results are in accord with experimental data and with the results of calculation by the quantum molecular dynamics method. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Interest in the problem of excess electrons solvated in a polar liquid has persisted for several decades [1–4]. This attention to the problem is due to the fact that, on the one hand, a large body of experimental data have been obtained [5–7] on the effect of a polar liquid on the behavior and parameters of solvated electrons; on the other hand, excess electrons in a liquid are the simplest examples of a quantum-mechanical–classical system. A solvated electron may serve as an excellent test for verification of various theoretical approaches [8–10]. However, the simulation of the behavior of excess electrons in a liquid taking into account the detailed microscopic structure of the medium involves considerable computational expenditure. The application of various approaches of statistical physics (group-theoretical expansion, the apparatus of correlation functions, etc.) makes it possible to reduce the problem to computation of average equilibrium characteristics and, in some cases, to use analytic estimates instead of computations [9, 10]. In fact, the problem can be reformulated as the problem of self-consistent calculation of the effective field, which determines the behavior of a classical ensemble of particles in a liquid as well as the wave function of the excess electron [10].

Recently, we used this approach for analyzing the behavior of excess electrons in Coulomb liquids [11, 12] and calculated the absorption spectrum for an electron solvated in alkali-halide melts as well as the variations of these spectrum, associated with changes in temperature, density, and composition of the solution. These calculations were made in the standard format of the mean-field model. For strong electrolytes, this approximation is sufficient for obtaining correct estimates of energy and structural parameters of a solvated electron [13]. However, in the case of dipole liquids, this approximation leads to a wrong estimate on the dielectric properties of the liquid due to the long-range nature of orientational correlations and can hardly be used for calculating the parameters of solvated electrons.

Here, we generalize our approach to polar liquids. The main difficulty in the analysis of such system is that orientational correlations between particles must be taken into account for such liquids. Consequently, this study is aimed at the development of a modified mean-field model for calculating the characteristics of an electron solvated in a polar liquid. The paper is constructed as follows. In Section 2, the theory is described briefly, while the results are given in Section 3, where the calculations are performed for a solvated electron in a polar liquid with parameters typical of water and ammonium.

2. THEORY

We consider an excess electron in a polar liquid with density n_0, which consists of classical particles with a dipole moment **m**. The behavior of such a system is determined by the large partition function Ξ,

$$\Xi = \left\langle \left\langle \exp\left[-\beta\left(T + \sum_i^N u_{ej}(\mathbf{r}, \mathbf{R}_i, \mathbf{w}_i)\right)\right] + \frac{1}{2}\sum_{i\neq j}^N u_{ss}(\mathbf{R}_i - \mathbf{R}_j) - \mu N\right\rangle_e \right\rangle_S, \quad (1)$$

where symbols $\langle\ldots\rangle_e$ and $\langle\ldots\rangle_S$ indicate averaging over the electron coordinate and over classical degrees of freedom, the latter including both the coordinates

$\mathbf{R}^{\{N\}} = \{\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_N\}$ of the centers of mass of particles in the liquid and the orientations $\mathbf{w}^{\{N\}} = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_N\}$ of the dipole moments of the particles. In relation (1), β is the inverse temperature, T is the electron kinetic energy, μ is the chemical potential of particles in the liquid, u_{ss} is the potential of interaction between these particles, and u_{ej} is the electron–solvent potential; we assume that these are paired potentials.

When the electron ground state predominates, the averaging over the electron coordinate can be reduced to the averaging over the wave function $\phi(\mathbf{r})$ of the electron ground state [25]. In fact, this leads to the replacement of potential u_{ej} in formula (1) by its mean value

$$\langle u_{ej}(\mathbf{R}, \mathbf{w})\rangle_e = \int u_{ej}(\mathbf{r} - \mathbf{R}, \mathbf{w})\phi^2(\mathbf{r})d\mathbf{r}. \quad (2)$$

This potential $\langle u_{ej}(\mathbf{R}, \mathbf{w})\rangle_e$ can be regarded as an external field acting on an ensemble of classical particles. The source of this external field is an excess electron. Thus, the problem can be reduced to the estimation of the large partition function for a classical liquid in an external potential. Having obtained such an estimate, we must self-consistently calculate the electron density $n_e(\mathbf{r}) = \phi^2(\mathbf{r})$ for the excess election inducing this field.

In the general case, potential u_{ej} includes both the electrostatic interaction and the short-range component. These contributions exert different effects on the solvent. The short-range component leads to the formation of a cavity (solvatophobic effect), while the long-range component induces the polarization of the medium and orientation ordering of particles in the liquid. Since the size $r_e = \langle r^2\rangle_e^{1/2}$ of the electron density distribution is much larger than the characteristic size of variations of the short-range component, we can characterize the short-range contribution u_{ec} by only one parameter $a > 0$ and write $u_{ej}(\mathbf{R}, \mathbf{w})$ in the form

$$\begin{aligned} u_{ej}(\mathbf{R}, \mathbf{w}) &= u_{ec} + \mathbf{m}(\mathbf{w}_j)\nabla\frac{1}{|\mathbf{R}_j|} \\ &= a\delta(\mathbf{R}) + \mathbf{m}\nabla\frac{1}{|\mathbf{R}_j|}. \end{aligned} \quad (3)$$

Similarly, we can separate in the interparticle potential u_{ss} the short-range component u_s and the long-range contribution associated with the dipole–dipole interaction:

$$\begin{aligned} u_{ss}(\mathbf{R}_i - \mathbf{R}_{j,i} - \mathbf{w}_j) &= u_s(\mathbf{R}_i - \mathbf{R}_j) \\ &+ \mathbf{m}(\mathbf{w}_i)\mathbf{M}(\mathbf{R}_i - \mathbf{R}_j)\mathbf{m}(\mathbf{w}_j), \end{aligned} \quad (4)$$

$$\mathbf{M}(\mathbf{R}) \equiv \frac{3\mathbf{R}\mathbf{R}}{R^5} - \frac{\mathbf{I}}{R^3}, \quad (5)$$

where $\mathbf{M}(\mathbf{R}_i - \mathbf{R}_j)$ is the dipole interaction tensor and $\mathbf{I}$ is a second-order unit tensor.

We introduce the collective variable $n_{es}(\mathbf{R}, \mathbf{w})$, i.e., the generalized number density of particles in the liquid,

$$n_{es}(\mathbf{R}, \mathbf{w}) = \sum_j^N \delta(\mathbf{R} - \mathbf{R}_j)\delta(\mathbf{w} - \mathbf{w}_j). \quad (6)$$

Then we can write the large partition function in functional form:

$$\begin{aligned} \Xi &= \langle \exp[-\beta E_{\text{solv}}[n_{es}]]\rangle_S, \\ E_{\text{solv}}[n_{es}] &= T + n_{es}\cdot\frac{(u_{ss} + \mathbf{m}\mathbf{T}\mathbf{m})}{2}\cdot n_{es} \\ &+ n_e\cdot\left(u_{ec} + \mathbf{m}\nabla\frac{1}{|\mathbf{R}|}\right)\cdot n_{es} - \mu\cdot n_{es}. \end{aligned} \quad (7)$$

In this relation, the dot indicates convolution,

$$a\cdot b \equiv \int a(\mathbf{R}_1\mathbf{w}_1)b(\mathbf{R}\mathbf{w} - \mathbf{R}_1\mathbf{w}_1)d\mathbf{R}_1 d\mathbf{w}_1.$$

Our task is to find the functional dependence $E_{\text{solv}}[n_{es}]$ and then calculate the distribution $n_{es}(\mathbf{R}, \mathbf{w})$ of particles in the liquid. For this purpose, we must carry out averaging of the solvation energy E_{solv}; this, in turn, requires knowledge of the Jacobean $J(n_{es}, \mathbf{R}\mathbf{w})$ of the transition from the coordinate representation $\{\mathbf{R}\mathbf{w}\}$ to collective variables n_{es}. Naturally, relation (7) formally does not simplify the averaging procedure, but the bilinear functional dependence $E_{\text{solv}}[n_{es}]$ makes it possible to obtain the approximate estimate (7) from physical considerations.

First of all, we note that the polarization induced by an electron makes the main contribution to the solvation energy. Thus, the estimation of the polarization contributions in Eq. (7) makes it possible to estimate E_{solv} on the whole. For this purpose, we transform Eq. (7) into a continual integral with respect to the electric field strength $\mathbf{E}$ (see the Appendix) and then estimate this integral by the steepest descent method:

$$\begin{aligned} \exp[-\beta\Omega_{\text{solv}}] &= \langle \exp[-\beta E_{\text{solv}}]\rangle_s \\ &\propto \int D[\mathbf{E}]\exp[-\beta S(\mathbf{E})] \approx \exp[-\beta S(\mathbf{E}_m)], \end{aligned} \quad (8)$$

where $\mathbf{E}_m$ is the mean field strength, which is determined from the extremum condition

$$\frac{\partial S(\mathbf{E} = \mathbf{E}_m)}{\partial \mathbf{E}} = 0.$$

After a series of manipulations, we obtain an explicit

expression for the thermodynamic potential,

$$\Omega_{\text{solv}}(n_{es}\mu) = -\frac{1}{2}\mathbf{P}\cdot\mathbf{M}\cdot\mathbf{P} - \beta^{-1}n_0 f(n_{es})\cdot\left(1+\frac{1}{2!}n_0 h_{ss}\cdot f(n_{es})\right), \tag{9}$$

where $h_{ss}(R)$ is the correlation function for a pure solvent, which is determined by the short-range repulsion:

$$n_0^2[1+h_{ss}(|\mathbf{R}-\mathbf{R}'|)] = \left\langle \delta(|\mathbf{R}-\mathbf{R}_1|)\delta(|\mathbf{R}'-\mathbf{R}_2|)\exp\left[-\frac{\beta}{2}\sum_{mk}^{N}u_{ss}(|\mathbf{R}_m-\mathbf{R}_k|)\right]\right\rangle_S . \tag{10}$$

In relation (9), $\mathbf{P}(\mathbf{R})$ is the polarization of the liquid and f is the Mayer function; these quantities can be expressed in terms of the generalized density and the mean field strength:

$$\mathbf{P}(\mathbf{R}) = \int \mathbf{m}(\mathbf{w})n_{es}(\mathbf{R}\mathbf{w})d\mathbf{w}, \tag{11}$$

$$f(n_{es}) = \exp[-\beta(u_{ec}\cdot n_e + \mathbf{m}\mathbf{M}\cdot\mathbf{P}-\mu)]-1. \tag{12}$$

Variation of functional Ω_{solv} with respect to chemical potential μ leads to the following expression for the density $n_{es}(\mathbf{R}\mathbf{w})$ of the solvent:

$$\frac{\delta\Omega_{\text{solv}}}{\delta\mu(\mathbf{R}\mathbf{w})} = n_{es}(\mathbf{R}\mathbf{w}) = n_0[1+h_{es}(\mathbf{R}\mathbf{w})] = n_0[1+n_0 f(n_{es})\cdot h_{es}][f(n_{es})+1]. \tag{13}$$

Thus, combining Eqs. (12) and (13), we can find $\Omega_{\text{solv}}[n_{es}]$. In order to determine the functional $\Omega_{\text{solv}}[n_{es}]$, we must also know the explicit dependence $\mu[n_{es}]$, which can be obtained by inverting relation (13). However, this inversion can be carried out only with the help of additional approximations in view of the nonlinearity and nonlocality of relation (13). For example, in the standard format of the mean field model, we assume that the chemical potential μ = const [15]. In view of the long-range nature of orientational correlations, this approximation for dipole liquids is not quite correct because of a wrong estimate of the dielectric properties of the liquid [16].

In principle, the $\mu[n_{es}]$ dependence can be obtained on the theory of density functional [15]; in the general case, this dependence leads to integral equations of the Ornstein–Zernike type [16]. However, this problem can apparently be solved in the simplest way by using the local approximation

$$\beta\mu(\mathbf{R}) \approx \ln\left(\int \exp[\beta\mathbf{m}\mathbf{E}_m]d\mathbf{w}\right). \tag{14}$$

This approximation takes into account two physical effects: the asymptotic behavior of chemical potential $\mu(R \longrightarrow \infty)$ and the saturation of dipole orientations for $R \longrightarrow 0$. Using this approximation, we obtain an explicit expression for the Mayer function in terms of the electric field strength:

$$f_{es}(\mathbf{R}\mathbf{w}) = \frac{\exp[-\beta(u_{ec}\cdot n_e + \mathbf{m}\mathbf{E}_m)]}{\int \exp[-\beta\mathbf{m}\mathbf{E}_m]d\mathbf{w}} - 1. \tag{15}$$

Here, the mean field strength $\mathbf{E}_m$ is determined through the Maxwell equation (26) and can be expressed in terms of polarization $\mathbf{P}$, whose longitudinal component P_r is, in turn, connected with $\mathbf{E}_m$ via the relation

$$P_r(R) = -n_0 m[\coth(\beta m E_m) - (\beta m E_m)^{-1}] \times [1-n_0 f_{es}\cdot h_{ss}]\exp[-\beta u_{ec}\cdot n_e], \tag{16}$$

where

$$f_{es} = \exp[-\beta u_{ec}\cdot n_e]-1.$$

If we disregard the short-range interaction, expression (15) corresponds to the point dipole model [17]; a similar model was used for estimating the absorption spectrum of a solvated electron in amorphous ice [18].

Using standard thermodynamic relations between the thermodynamic functional and the free energy of the system,

$$F-\mu N = \Omega_{\text{solv}},$$

we obtain the final relations for functional $F[\phi, \mathbf{E}_m]$ of the free energy of a solvated electron:

$$F = \frac{1}{2}\int(d\mathbf{r}[\nabla\phi(r)]^2)d\mathbf{r} - \int\left[\phi^2(r) - \frac{\nabla E_m}{2\pi}\right] \times \int\frac{\phi^2(R)}{2|\mathbf{r}-\mathbf{R}|}d\mathbf{R} - \beta^{-1}n_0 f_{es}\cdot\left(1+\frac{1}{2}n_0 h_{ss}\cdot f_{es}\right) - \frac{1}{2}\mathbf{E}_m\cdot\mathbf{M}^{-1}\cdot\mathbf{E}_m + \mathbf{P}(E_m)\cdot\mathbf{E}_m. \tag{17}$$

Minimizing this functional with respect to the electron wave function, we arrive at the Schrödinger equation with the self-consistent potential $V(\mathbf{r})$:

$$\left[-\frac{1}{2}\nabla^2 + \int\left[\frac{\nabla E_m}{4\pi} - \phi^2(r)\right]\frac{d\mathbf{r}}{|\mathbf{r}-\mathbf{R}|} + n_0 a(f_{es}+1)(1+n_0 h_{ss}\cdot f_{es})\right]\phi(r) = E_0\phi(r). \tag{18}$$

The mean field strength $\mathbf{E}_m$ is determined from Eqs. (26) and (16). Thus, in order to calculate the characteristics of an electron solvated in a polar liquid, we

must find self-consistent solutions to Eqs. (16), (18), and (26). Then we can calculate the first excited level for the fixed self-consistent potential $V(\mathbf{r})$ and estimate the maximum in the absorption spectrum of a solvated electron as

$$\Delta E = |E_0 - E_1|.$$

3. RESULTS

We have calculated the characteristics of a solvated electron, namely, its mean radius r_e, free energy F, total E, potential $\langle V\rangle$ and kinetic $\langle T\rangle$ energies, the absorption spectrum peak ΔE, and the temperature dependence of this peak $\partial\Delta E/\partial t$ for the parameters typical of water and ammonium. For water under standard conditions, we used the following values of parameters: m = 1.854 D, n_0 = 0.0334 Å^{-3}, thermal expansion coefficient $\partial \ln n_0/\partial t = 2.5 \times 10^{-4}$ grad^{-1} [20], and hard sphere diameter σ = 3.15 Å.

It should be noted that among the parameters of our problem, it is most difficult to estimate parameter a. We estimated this parameter on the basis of the electron–hydrogen $u_{e\mathrm{H}}(r)$ and the electron–oxygen $u_{e\mathrm{O}}(r)$ interaction potentials used for calculating hydrated electron by the methods of quantum molecular dynamics or integral equations [21–24]. In these calculations, we assume that this potential at small distances differs from the Coulomb interaction potential, i.e.,

$$u_{ei}(r < R_i) = u_{ei}(R_i) \quad (i = \mathrm{O}, \mathrm{H}).$$

Usually, R_O = 0 Å, while $0 < R_\mathrm{H} < 1$ Å. In our calculations, we used the value of R_H = 0.5 Å. Taking into account approximation (3), we find that

$$a = 4\pi\int_0^\infty [u_{e\mathrm{O}} + 2u_{e\mathrm{H}} - (z_\mathrm{O} + 2z_\mathrm{H})/r]r^2 dr$$
$$= 4\pi z_\mathrm{H} R_\mathrm{H}^2/3,$$

where z_H are partial charges of atoms. We calculated correlation function $h_{ss}(r)$ in the hard sphere model by numerically solving the integral equation with Percus–Yewick closure [25].

In the case of ammonium, for t = –33°C, we took m = 1.471 D, n_0 = 0.024 Å^{-3}, thermal expansion coefficient $\partial \ln n_0/\partial t = 1 \times 10^{-3}$ grad^{-1}, and hard sphere diameter σ = 3.65 Å. Parameter a was estimated and correlation function $h_{ss}(r)$ was calculated in the same way as for a hydrated electron.

In the general case, a detailed analysis of the characteristics of a solvated electron requires numerical solution of Eqs. (16) and (18); our previous investigations [13] proved, however, that these characteristics can be estimated accurately by the variational method also by choosing a certain type of trial functions for the electric field and the electron wave functions, after which function (17) for these functions is minimized directly. For such trial functions of the ground and first excited electron states, we used the following approximations:

$$\phi(r) = \left(\frac{2\alpha^2}{\pi}\right)^{3/4} \exp[-\alpha^2 r^2],$$
$$\phi_1(\mathbf{r}) = \left(\frac{\alpha_1^5}{\pi}\right)^{3/2} r\cos\theta\exp[-\alpha_1 r]; \tag{19}$$

here, α and α_1 are variational parameters characterizing the electron density distribution in the ground and excited states. We approximated the electric field strength E_m as

$$E_m(r) = \frac{\gamma}{r^2}\int_0^r \phi^2(R)R^2 dR. \tag{20}$$

It should be noted that the variational parameter γ has the meaning of reciprocal permittivity; i.e., $\gamma^{-1} = \epsilon_e f$, which the choice of approximation (20) for the electric field is dictated by the fact that this approximation becomes accurate in the asymptotic limit corresponding to the continual model of the liquid, when the radius of short-range interactions tends to zero, and $\gamma^{-1} = \epsilon$.

The results of calculations compared to experimental data are given in the table. A comparison of these data shows that our model makes it possible to estimate correctly the structural and thermodynamic parameters of electrons solvated in water and ammonium. Our calculations are also in qualitative agreement with the data obtained by the quantum molecular dynamics method. For example, for an electron solvated in ammonium, we obtained E = –1.84 eV and ΔE = 1.32 eV [29],

Characteristics of an electron solvated in water and ammonium (experimental values are given in the parentheses)

	Water (t = 25°C)	Ammonium (t = –33°C)
r_e, Å	2.5	3.0
$\langle T\rangle$, eV	1.32	0.96
$\langle V\rangle$, eV	–4.82	–3.51
E, eV	–3.5	–2.55
F, eV	–1.01 (–1.6 [26])	–0.72
ΔE, eV	1.59 (1.72 [27])	1.16 (0.8 [28])
$\partial\Delta E/\partial t$, 10^{-3} eV K^{-1}	–1.15 (–2.9 [28])	–2.2 (–1.5 ± 0.2 [28])

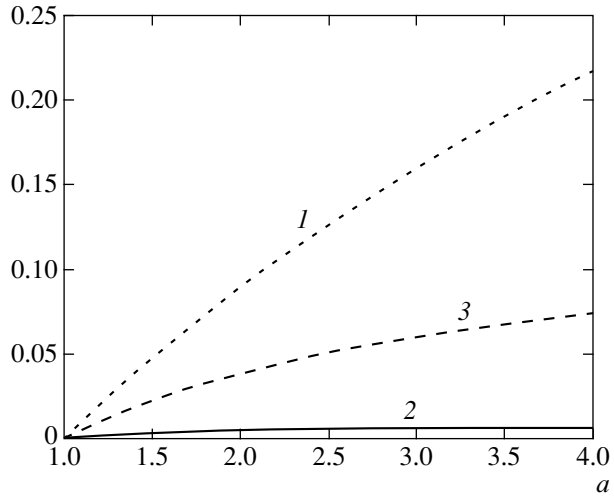

Fig. 1. Relative variations of the total energy δE (*1*), absorption peak $\delta\Delta E$ (*2*), and the mean radius r_e (*3*) for a hydrated electron as functions of parameter a characterizing the short-range repulsion between an electron and the solvent. All quantities are normalized to the corresponding values for a = 1.4 at. units.

while for a hydrated electron, we have E = –2.4 eV and ΔE = 2.2 eV [30]. These values are close to our estimates; however, our model predicts a slightly higher value of the mean radius for a hydrated electron and a slightly lower value for an electron solvated in ammonium (the calculations based on the quantum molecular dynamics method give r_e = 2.3 Å [29] and 3.8 Å [30]). In our opinion, this is associated with the peculiar way the short-range contribution u_{ec} is included in the electron–solvent interaction potential. We analyzed the effect of parameter a, characterizing this contribution, on the results of calculations. Figure 1 shows the relative variation of the total energy E, the absorption maximum $\delta\Delta E$, and the mean radius r_e for a hydrated electron as functions of parameter a. It can be seen that a fourfold increase in this parameter does not lead to an appreciable change in the total energy, while the mean radius increases by 10% and the transition energy increases by 20%. Thus, our results show that a more detailed account of the short-range contribution to the electron–solvent interaction potential is required for estimating the characteristics of a solvated electron mode accurately.

Another important feature of our model is that it enables us to calculate the temperature variations of the transition energy also (see table). Our calculations are in quantitative agreement with experimental data for an electron solvated in ammonium, while only a qualitative agreement is observed for a hydrated electron. This peculiarity is due to anomalous temperature properties of water [20], for which the thermal expansion coefficient is almost an order of magnitude smaller than for other polar liquids. In our opinion, this feature is in turn determined by the peculiarity of formation of hydrogen bonds in water, which is not taken into account in our simple model in any way. By way of example (see Figs. 2–4), let us compare our results with the results of calculations for the temperature dependences of the free, kinetic, and total energies of a hydrated electron, which were obtained by the method of integral equations taking into account the molecular structure of water [24]. It can be seen that our calculations for the free and total energies lead to shifts opposite to those obtained in [24]. This is apparently associated with the specific way in which the effects of cavity formation in

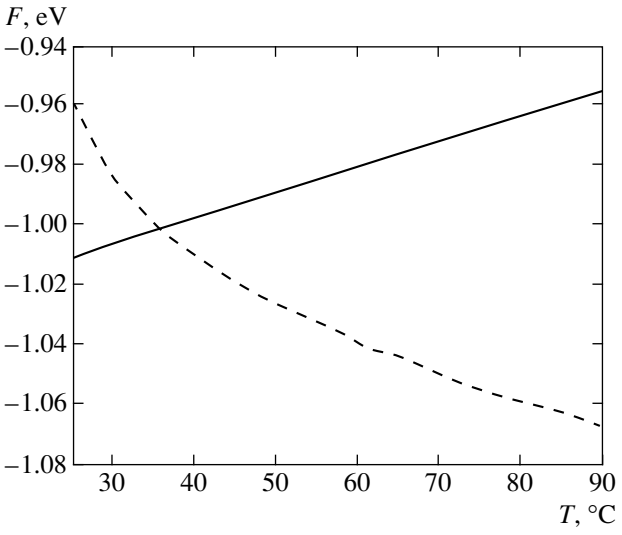

Fig. 2. Temperature dependence of the free energy of a hydrated electron. The solid curve corresponds to our calculations, while the dashed curve corresponds to the results obtained in [24].

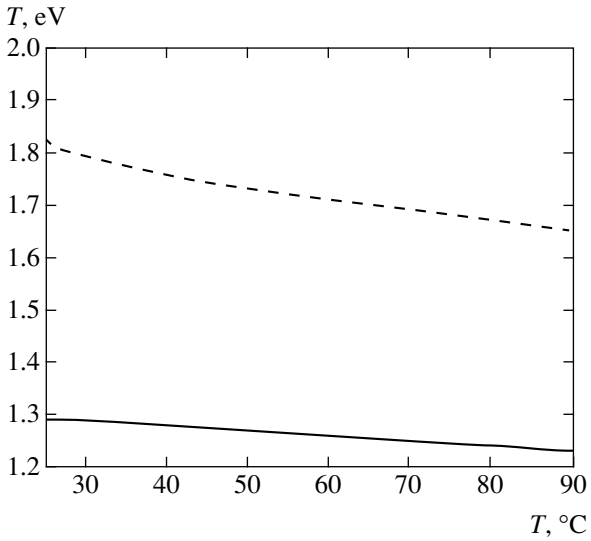

Fig. 3. Temperature dependence of the kinetic energy of a hydrated electron. Notation is the same as in Fig. 2.

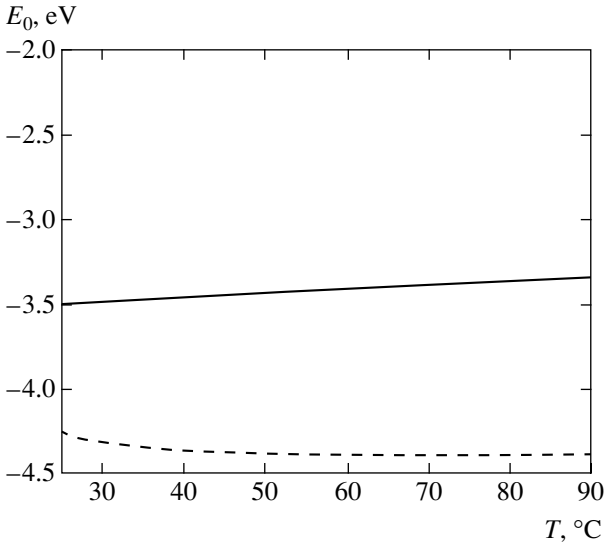

Fig. 4. Temperature dependence of the total energy of a hydrated electron. Notation is the same as in Fig. 2.

the molecular model of water are taken into account (short-range interactions and the temperature factor are significant for these effects). It should be noted that an analysis of such changes requires a very accurate calculation of the parameters of a hydrated electron. For example, a change in the free and total energies of a hydrated electron in the range of 0–50°C amounts to only 5%, which is obviously at the margin of accuracy in our simple model. In our opinion, more adequate molecular models are required for estimating such temperature variations.

It should be noted in conclusion that our model makes it possible to calculate the behavior of an electron solvated in a polar liquid taking into account microscopic parameters of the medium and its thermodynamic state. We believe that the methods developed in this study enable one to quite easily obtain self-consistent estimates for an excess electron in various polar disordered media and then to adapt the results to real systems. The proposed approach reduces the problem to the calculation of average equilibrium parameters and permits the replacement of the calculation by variational estimates. This makes it possible to considerably reduce the computer time as compared to a direct simulation by the Monte Carlo or quantum molecular dynamics methods without deteriorating the accuracy of computations.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project no. 01-03-32529) and partly financed by a grant from NATO CNR.

APPENDIX

Transformation of Partition Function

Transforming a functional to an integral continual in the field, we must take into account the degeneracy of dipole interaction with respect to orientations of particles in a liquid. The physical meaning of this effect is that an additional charge may induce in the general case not only the longitudinal, but also a transverse electric field. Taking this into account, we transform the large partition function (7) into a continual integral with respect to electric field strength **E**, using the relation

$$\exp\left[-\frac{\beta}{2}n_{es}\cdot\mathbf{mMm}\cdot n_{es}\right] = \int D[\mathbf{E}]\exp\left[\frac{\beta}{2}\mathbf{E}\cdot\mathbf{M}^{-1}\cdot\mathbf{E}-\beta n_{es}\cdot\mathbf{mE}\right], \tag{21}$$

where $\mathbf{T}^{-1}$ is a matrix inverse to the dipole interaction tensor. Applying the Hubbard–Stratonovich transformation [19], we ultimately obtain

$$\exp[-\beta\Omega_{\text{solv}}] = \int D[\mathbf{E}]\exp[-\beta S(\mathbf{E})] = \int D[\mathbf{E}] \times \exp\left[\frac{\beta}{2}\left(\mathbf{E}-n_e\nabla\frac{1}{|\mathbf{R}|}\right)\cdot\mathbf{M}^{-1}\cdot\left(\mathbf{E}-n_e\nabla\frac{1}{|\mathbf{R}|}\right)\right] \times\langle\exp[-\beta n_{es}\cdot(\mathbf{mE}+u_{es}\cdot n_e-\mu+u_{ss}\cdot n_{es})]\rangle_s. \tag{22}$$

As a result, action $S(\mathbf{E})$ is factorized, and we can obtain an explicit equation for this quantity in terms of one- and two-particle irreducible correlation functions,

$$S(\mathbf{E}) = -\frac{\beta}{2}\left(\mathbf{E}-n_e\nabla\frac{1}{|\mathbf{R}|}\right)\cdot\mathbf{M}^{-1}\cdot\left(\mathbf{E}-n_e\nabla\frac{1}{|\mathbf{R}|}\right) - \beta^{-1}n_0 f(\mathbf{E})\cdot\left[1+\frac{1}{2!}n_0 h_{ss}\cdot f(\mathbf{E})\right], \tag{23}$$

where $f(\mathbf{E})$ is the Mayer function,

$$f(\mathbf{E}) = \exp[-\beta(u_{ec}\cdot n_e+\mathbf{mE}-\mu)]-1. \tag{24}$$

The extremum of action $S(\mathbf{E})$ with respect to the electric field strength leads to an integral relation between the field strength and polarization:

$$\mathbf{E}_m(\mathbf{R}) = n_e\nabla\frac{1}{|\mathbf{R}|}+\int\mathbf{M}(\mathbf{R}'-\mathbf{R})\mathbf{P}(\mathbf{R}')d\mathbf{R}', \tag{25}$$

which can be written in the form of the Maxwell equation

$$\nabla\mathbf{E}_m = 4\pi n_e - 4\pi\nabla\mathbf{P}. \tag{26}$$

Using Eqs. (24) and (25), we obtain Eq. (9).

Disregard of the short-range interactions in (23)–(26) leads to the Debye–Hückel limit, while the linearization of dependence (24) (for $\mathbf{P} \propto \mathbf{E}$) leads to the linearized Debye–Hückel model. We can also write an auxiliary relation between **P** and **E**,

$$\mathbf{P}(\mathbf{R}) = \frac{\epsilon - 1}{4\pi}\mathbf{E}, \tag{27}$$

which serves as the definition of permittivity ϵ. Combining this expression with Eq. (26) in the case when an excess electron charge is considered as a point charge, we arrive at the Clausius–Mossotti relation for permittivity ϵ. Thus, various approximations can be obtained with the help of Eqs. (24)–(27). The choice of an approximation is determined by the accuracy required for calculating the microscopic structure of the liquid.

REFERENCES

1. J. Jortner and N. R. Kestner, *Electrons in Fluids* (Springer, Berlin, 1973).
2. J. C. Thompson, *Electrons in Liquid Ammonium* (Oxford Univ. Press, London, 1976).
3. R. R. Dogonadze, E. M. Itskovitch, E. Kalman, *et al.*, *The Chemical Physics of Solvation* (Elsevier, Amsterdam, 1988).
4. J. Acrivos, *Physics and Chemistry of Electrons and Ions in Condensed Matter* (Reidel, New York, 1984).
5. E. D. Hart and M. Anbar, *The Hydrated Electron* (Wiley, New York, 1970; Atomizdat, Moscow, 1973).
6. A. K. Pikaev, *Modern Radiation Chemistry. Radiolysis of Gases and Liquids* (Nauka, Moscow, 1986).
7. J. H. Baxendale and F. Busi, *The Study of Fast and Transient Species by Electron Pulse Radiolysis* (Reidel, Dordrecht, 1982).
8. C. Ferradini and J. P. Jay-Gerin, *Excess Electrons in Dielectric Media* (CRC Press, Boca Raton, 1991).
9. D. Chandler and K. Leung, Annu. Rev. Phys. Chem. **45**, 1500 (1994).
10. G. N. Chuev, Usp. Fiz. Nauk **169**, 155 (1999) [Phys. Usp. **42**, 149 (1999)].
11. G. N. Chuev, Zh. Éksp. Teor. Fiz. **118**, 1134 (2000) [JETP **91**, 983 (2000)].
12. G. N. Chuev and V. V. Sychyov, J. Chem. Phys. **112**, 4707 (2000).
13. G. N. Chuev, V. V. Sychyov, and O. Yu. Sokolova, Phys. Rev. E **63**, 061204 (2001).
14. I. R. Yukhnovskiĭ and M. F. Golovko, *Statistical Theory of Classical Equilibrium Systems* (Naukova Dumka, Kiev, 1980).
15. J. K. Percus, in *Recent Developments and Applications of Modern Density Functional Theory*, Ed. by J. M. Seminario (Elsevier, Amsterdam, 1996).
16. G. N. Chuev and N. Russo, Chem. Phys. Lett. **368**, 53 (2003).
17. K. Iguchi, J. Chem. Phys. **48**, 1735 (1968).
18. W. M. Bartczak, M. Hilczer, and J. Kroh, J. Phys. Chem. **91**, 3834 (1987).
19. J. W. Negele and H. Orland, *Quantum Many Particle Systems* (Addison–Wesley, Redwood City, 1988).
20. D. Eisenberg and W. Kauzmann, *The Structure and Properties of Water* (Oxford, London, 1969).
21. M. Sprik, R. W. Impey, and M. L. Klein, J. Stat. Phys. **43**, 949 (1986).
22. D. Laria, D. Wu, and D. Chandler, J. Chem. Phys. **95**, 4444 (1991).
23. J. Zhu and R. I. Cukier, J. Chem. Phys. **99**, 53841 (1993).
24. S. Miura and F. Hirata, J. Phys. Chem. **98**, 9649 (1994).
25. J.-P. Hansen and I. R. McDonald, *Theory of Simple Liquids* (Academic, London, 1986).
26. J. Jortner and R. M. Noyes, J. Phys. Chem. **70**, 770 (1966).
27. E. J. Hart and W. C. Gottschall, J. Am. Chem. Soc. **71**, 2102 (1969).
28. R. K. Quinn and J. J. Lagowski, J. Phys. Chem. **73**, 2326 (1969).
29. M. Sprik and M. L. Klein, J. Chem. Phys. **87**, 5987 (1987).
30. J. Schnitker and P. J. Rossky, J. Chem. Phys. **86**, 3471 (1987).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 573–586.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 633–649.

SOLIDS
Structure

Inhomogeneity of the Ferromagnetic State of $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ Perovskites according to ^{139}La and ^{55}Mn NMR Data

M. M. Savosta[†a], V. D. Doroshev[a,*], V. I. Kamenev[a,*], V. A. Borodin[a,*], T. N. Tarasenko[a,*], A. S. Mazur[a,*], and M. Maryško[b]

[a]*Donetsk Physicotechnical Institute, National Academy of Sciences of Ukraine, Donetsk, 83114 Ukraine*
[b]*Institute of Physics, Academy of Sciences of the Czech Republic, 18221 Praha 8, Czech Republic*
**e-mail: savosta@kinetic.ac.donetsk.ua*

Received December 28, 2002

Abstract—The paper presents results obtained in studying ^{139}La and ^{55}Mn nuclear magnetic resonance spectra and nuclear magnetic relaxation in the temperature range 21–220 K augmented by data on the magnetic and transport properties of ferromagnetic $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ manganites ($x = 0$, 0.075, and 0.125, $\delta \approx 0.03$–0.05). The transition from the ferromagnetic state with semiconductor-type conductivity to the ferromagnetic state with metallic conductivity as the degree of doping x increased was related to the redistribution of the volumes of two phases (two types of regions) with different degrees of freedom of electron holes on manganese sites. The ferromagnet–paramagnet phase transition was a smooth redistribution of the volumes of ferromagnetic and paramagnetic (or antiferromagnetic) regions in a wide temperature range. Ferromagnetic conductivity was caused by the "double exchange" mechanism, whereas, in comparatively nonconducting regions, double exchange was considerably weakened and involved fairly slow thermally activated motion of Jahn–Teller polarons. The dynamics of nuclear spins was evidence of internal inhomogeneity of each of these phases.

1. INTRODUCTION

Ferromagnetic $La_{1-x}A_xMnO_3$ manganites (A = Ca, Sr, Ba, or Pb) with perovskite-like structures have been the objects of extensive studies in recent years because of the so-called giant magnetoresistance effect close to the temperature of magnetic ordering of manganese spins [1–3]. The close relation between the magnetic and electric properties of these compounds is explained in terms of the double exchange model [4], according to which electron holes induced by doping exhibit a tendency to easily jump between neighboring manganese sites if both sites have parallel orientations. As a result, ferromagnetic ordering, for instance, in "optimally" doped samples ($x \approx 0.3$) is accompanied by a sharp decrease in electric resistance directly below the Curie temperature T_C [5, 6]. More recent studies, however, showed that the transition to metallic conductivity often occurred far below T_C [7–9]. What is more, with a low concentration of holes, strong disorder or lattice distortions, and substitutions in manganese sites, the ferromagnetic state can be formed without a decrease in resistance. Such a state (the ferromagnetic insulating state) has a semiconductor-type conductivity even to the lowest temperatures [7, 10, 11]. The nature of such states, which do not fit in with the base concept of double exchange, remains unclear. For this reason, other models are currently being actively discussed, such as the phase separation into ferromagnetic metallic and insulating clusters with charge ordering, ferromagnetic superexchange, charge and/or orbital ordering, cluster spin glass, or a nonuniform ground state (e.g., see [3, 12–15]).

Because of its local nature, nuclear magnetic resonance (NMR) allows valuable information to be obtained about the special features of the magnetic state of such systems. The ^{139}La NMR spectra of various ferromagnetic insulating $LaMnO_{3+\delta}$ and $La_{1-y}Ca_yMnO_{3+\delta}$ manganites were studied in [16]. A substantial decrease in the intensity or the complete disappearance of the NMR signal was observed for all samples in a certain temperature range. The authors of [16] related this observation to slow diffusion of lattice excitations. The source of such excitations was assumed to be small Jahn–Teller polarons. Similar results were obtained in studies of the ^{139}La NMR spectra of $La_{1-x}Ca_xMnO_3$ [17]. A decrease in the intensity of the NMR signal was explained in [17] by quasi-static distortions of manganese octahedra in the ferromagnetic insulating phase. The ^{55}Mn NMR spectra of the corresponding systems were studied in [18–21]. The dynamics of nuclear spins, which contains important information about the

[†] Deceased.

Crystal lattice parameters of $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ samples

x	a, Å	c, Å	V/formula unit, Å^3
0	5.523(1)	13.359(1)	58.8
0.125	5.5265(15)	13.336(3)	58.8

special features of the structure of the ferromagnetic state of such systems, has not, however, been studied.

This work presents the results of a detailed study of the NMR spectra and nuclear magnetic relaxation for the ^{139}La and ^{55}Mn nuclei in the $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ system with comparatively low concentrations of Mn^{4+} ions (x = 0, 0.075, and 0.125). These data are augmented by the results obtained in measuring the magnetic and transport properties of the system. Our purpose was to study the transition between the ferromagnetic insulating (FMI) and ferromagnetic metallic (FMM) states. The objects of study were selected from the following considerations. It is known that, apart from doping with strontium atoms, manganese variable valence and, accordingly, the ferromagnetic state can be formed in manganites by creating cationic vacancies predominantly in lanthanum sites [11, 22–25]. Depending on synthesis conditions, both the ferromagnetic metallic and ferromagnetic insulating states can be obtained. An increase in the degree of doping with strontium under constant synthesis conditions increases T_C and stabilizes the ferromagnetic metallic state. A similar system was studied in [26] by elastic neutron scattering (PDF analysis). The authors advanced the suggestion that carriers in self-doped lanthanum manganites (x_{Sr} = 0) could be captured by cationic vacancies, which explains the low thermally activated electric conductivity, whereas doping with strontium caused the formation of mobile polarons responsible for metallic conductivity. The NMR data obtained in this work show that the compounds under study simultaneously contain two types of ferromagnetic regions with different electron hole mobilities in manganese sites. Depending on the relative volumes of these regions, the ground state may be either conducting or semiconducting. We also found that the phase transition from the ferromagnetic to the paramagnetic state was smeared and occurred as a smooth redistribution of the volumes of ferromagnetic and paramagnetic (or antiferromagnetic) regions in a wide temperature interval.

2. EXPERIMENTAL

Polycrystalline $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ samples (x = 0, 0.075, and 0.125) were synthesized by the standard ceramic technique with sintering in air, because precisely this method of synthesis allows the required concentration of carriers to be obtained via the formation of vacancies in La/Sr sites [27]. The initial materials were La_2O_3, MnO_2, and $SrCO_3$ powders of high purity. These materials were mixed in a stoichiometric ratio and thoroughly ground in an agate mortar under rectified ethanol. After drying, the batch mixture was pressed into bars. The samples were preliminarily heated at 1000°C for 21.5 h. They were then thoroughly ground, pressed into bars, and eventually sintered at 1100°C for 30 h.

The crystal structure of the samples was determined by X-ray diffraction, NiK_α radiation. According to the X-ray data, the samples were single-phase and had a rhombohedral structure ($R\bar{3}c$ symmetry). The cell parameters of the samples of the limiting compositions are listed in the table. The parameter values are in agreement with the literature data on similar compositions [22, 27]. The δ values were estimated by comparing the lattice parameters (table) with the data reported in [22, 27]; this gave $\delta \approx 0.05$ and $\delta \approx 0.03$ for x = 0 and x = 0.125, respectively. Note that the reflections from the sample with x = 0 were broadened compared with the doped x = 0.125 sample; this is likely to be evidence of a greater defect structure of the sample with x = 0 because of the presence of cationic vacancies.

Magnetization and initial magnetic susceptibility measurements were performed on a SQUID magnetometer. The magnetization curves were recorded in fields up to 5 T. Spontaneous magnetic moment values were obtained by extrapolating the curve portions recorded at fields above the saturation field to a zero magnetic field. The resistive measurements were performed by the four-point-probe method. The results of these measurements are shown in Fig. 1, according to which the ferromagnet–paramagnet transition in the sample with x = 0 is smeared over a wide temperature range. The transition becomes sharper as the concentration of strontium grows. Simultaneously, the Curie temperature T_C, determined as the inflection point of the susceptibility curve, systematically increases. In addition, the sample with x = 0 has semiconductor-type resistance in the entire temperature range of measurements, whereas the transition from semiconductor-type to metallic conductivity is observed for the samples with x = 0.075 and 0.125 as temperature decreases. Note that the temperature of the insulator–metal transition is substantially lower than T_C for the doped samples, especially for the sample with x = 0.075.

NMR measurements were performed on an NMR spectrometer with a slow frequency sweep and boxcar detector signal averaging. The NMR spectra were recorded using the two-pulse spin echo sequence τ–τ_{1-2}–τ, where τ_{1-2} is the time interval between exciting pulses of width $\tau \approx 0.5$–1.0 μs. Because of rapid relaxation, the spectra given in this work were recorded using a minimum τ_{1-2}^{min} time interval limited by the "dead" time of the detector and transition processes. The τ_{1-2}^{min} time was 15 μs for measurements on ^{139}La

nuclei and 3–3.5 μs for ^{55}Mn nuclei. Spin–spin relaxation was studied by recording the dependences of the area under the envelope of the spin-echo signal on the τ_{1-2} time interval in two-pulse T_2 experiments. Spin–lattice relaxation was studied by recording signal restoration curves versus time t after the inversion of nuclear spins by a 180° pulse at a fixed τ_{1-2} value, 180°–t–90°–τ_{1-2}–180° (three-pulse T_1 experiments). All recorded signals had a large NMR amplification factor, $\eta \geq 600$, which was evidence that they originated from ferromagnetic regions.

3. THE NMR RESULTS

The ^{139}La and ^{55}Mn NMR spectra of $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ at T = 77 K are shown in Fig. 2.

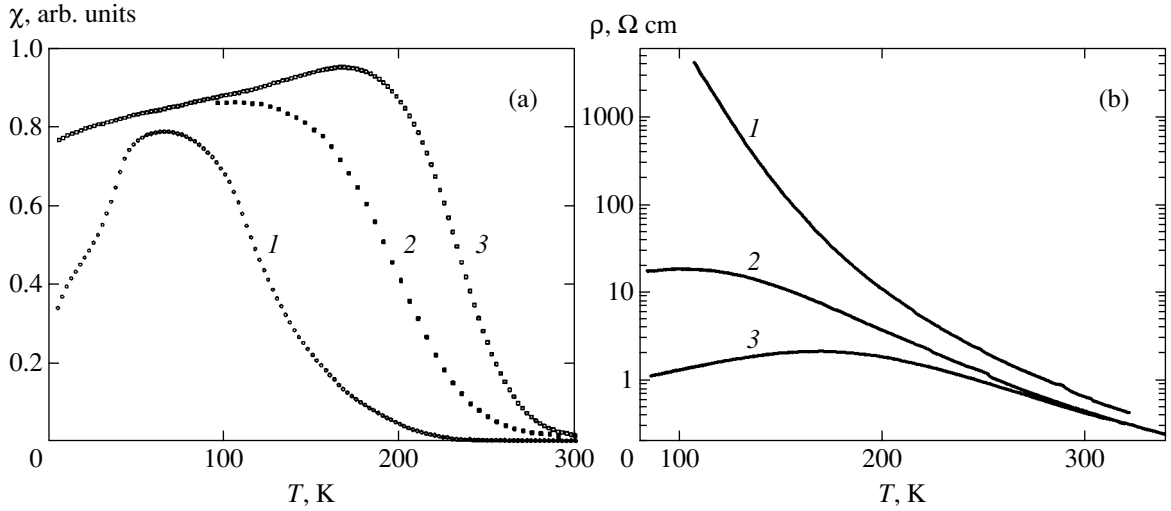

Fig. 1. Temperature dependences of (a) initial magnetic susceptibility and (b) specific electric conductivity of $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ samples with x = 0 (*1*), 0.075 (*2*), and 0.125 (*3*).

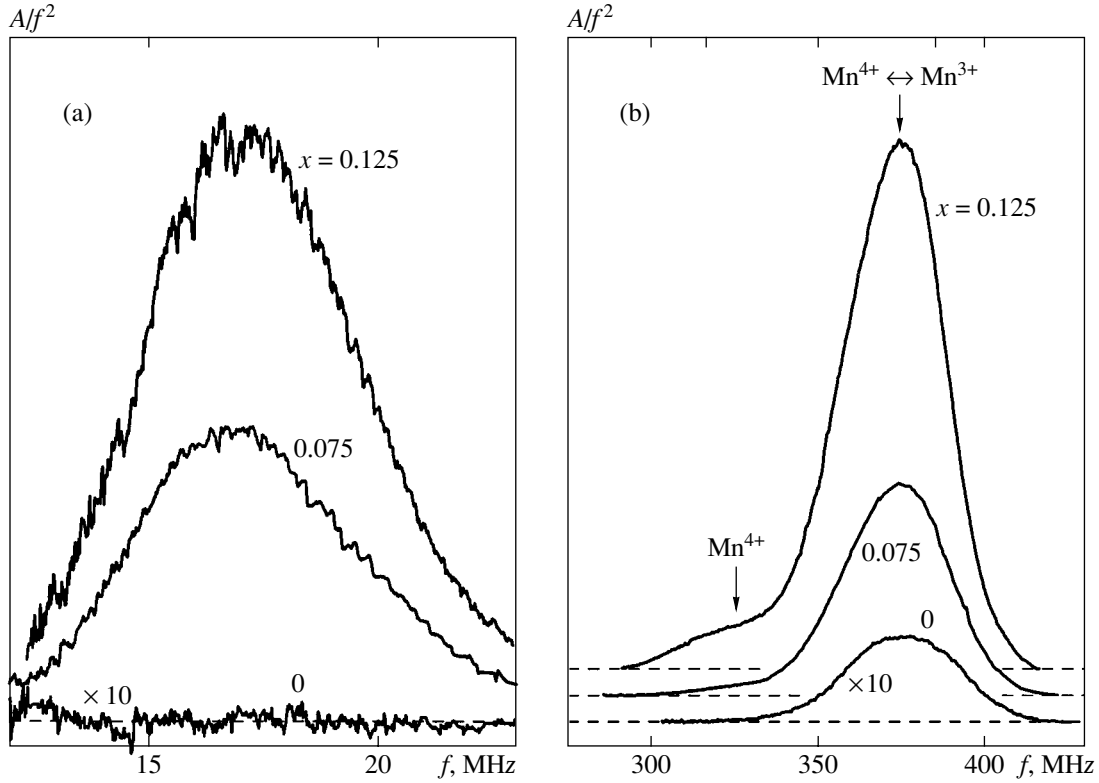

Fig. 2. (a) ^{139}La and (b) ^{55}Mn NMR spectra of $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ at T = 77 K. The intensities of the spectra of the x = 0 sample are increased 10 times.

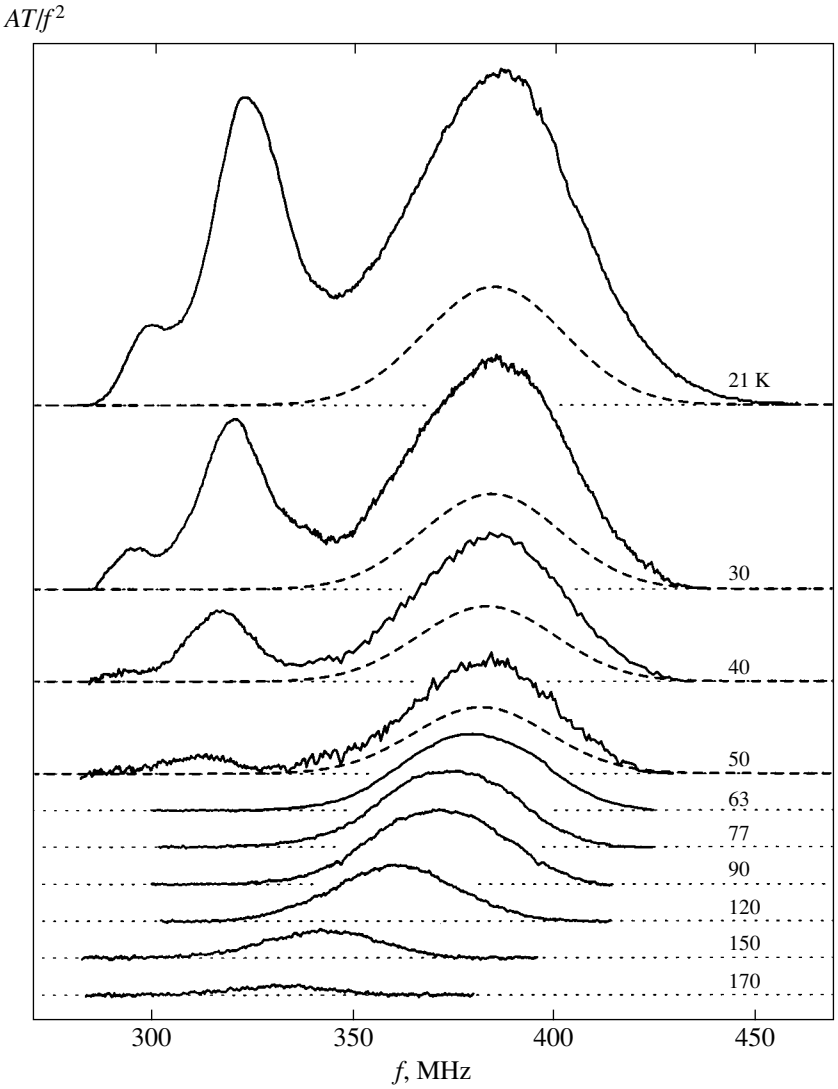

Fig. 3. Temperature evolution of the ^{55}Mn NMR spectra of LMO recorded using $\tau_{1-2} = 3.5$ μs ($T = 21$–50 K) and $\tau_{1-2} = 3$ μs ($T \geq 63$ K). The dashed lines schematically represent the contribution of phase I (see text) to the NMR spectra at low temperatures.

The ^{139}La NMR spectrum consists of a fairly broad single line centered at an $f \approx 17$ MHz frequency for $x = 0.125$ and 0.075. The signal is not observed at $x = 0$. This is in agreement with the data reported in [16, 17], according to which the ^{139}La NMR signal disappears at nitrogen temperatures in ferromagnetic insulating manganites such as the sample with $x = 0$ because of a sharp decrease in the spin–spin relaxation time T_2. Note that the amplitude of the signal for the composition with $x = 0.075$ is two times lower than for $x = 0.125$, which corresponds to the tendency mentioned above toward a decrease in time T_2 in compositions in which the type of conductivity changes from metallic to semiconductor at $T < T_C$. A similar situation is observed in the ^{55}Mn NMR spectra. The amplitude of the NMR signal for the line centered at about 380 MHz in the spectrum of the sample with $x = 0.125$ is 2.5 times larger than for the sample with $x = 0.075$ and 60 times larger than for the sample with $x = 0$. In addition to the line at $f \approx 380$ MHz caused by averaging superfine fields on ^{55}Mn nuclei by fast motions of electron holes over manganese sites $Mn^{4+} \longleftrightarrow Mn^{3+}$ (with a characteristic hopping frequency $f_{hop} > f_{NMR}$), the NMR spectra of the samples with $x = 0.125$ and 0.075 contain a line at $f \approx 325$ MHz, which corresponds to quasi-localized ($f_{hop} < f_{NMR}$) Mn^{4+} states [28].

Further, we concentrate on the samples with $x = 0$ (LMO) and $x = 0.125$ (LSMO), which, at $T < T_C$, as follows from Fig. 1, are manganites with semiconductor-type and metallic conductivities, respectively. The temperature evolution of the ^{55}Mn NMR spectra of LMO is shown in Fig. 3. The data given in this figure were corrected to exclude the $A \propto 1/T$ temperature factor of the NMR signal amplitude caused by the Curie law for nuclear magnetization. The behavior of the intensity of

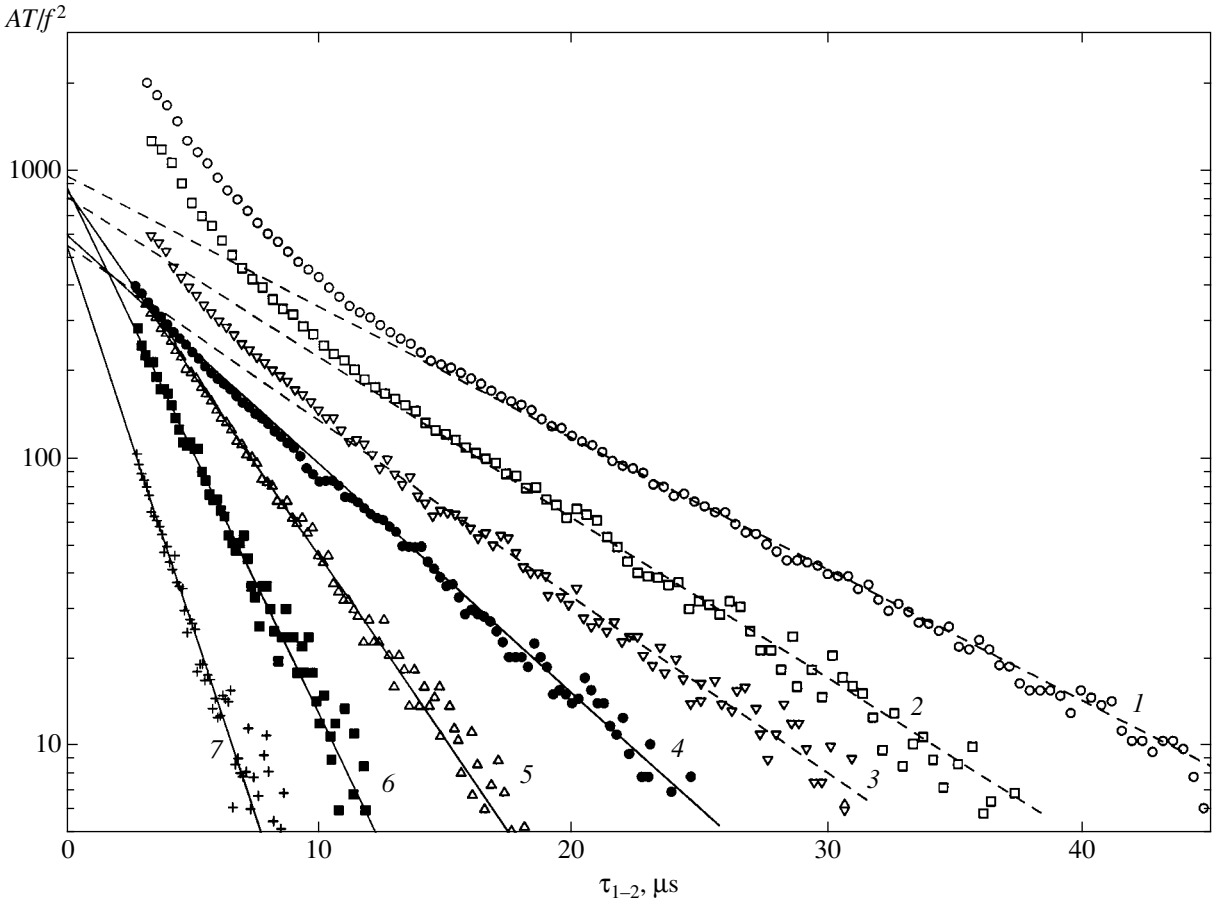

Fig. 4. Dependences of the spin echo amplitude of ^{55}Mn nuclei in LMO on the τ_{1-2} interval at various temperatures: (*1*) 21 K (f = 385 MHz), (*2*) 30 K (f = 384 MHz), (*3*) 50 K (f = 382 MHz), (*4*) 63 K (f = 380 MHz), (*5*) 100 K (f = 369 MHz), (*6*) 130 K (f = 356 MHz), and (*7*) 160 K (f = 337 MHz).

the principal line ($f \approx 380$ MHz at $T = 77$ K) is indicative of the existence of three characteristic temperature intervals. In the temperature range 63 K < T < 120 K, the reduced line intensity is virtually constant. At T < 63 K, the intensity of the signal substantially increases, which is evidence of the appearance of some additional contribution at $f \approx 390$ MHz (T = 21 K). Moreover, a new line at $f \approx 310$–325 MHz corresponding to quasi-localized Mn^{4+} states appears in the spectrum. Its intensity rapidly increases as temperature decreases. Lastly, at T > 120 K, the NMR signal amplitude decreases because of a decrease in the T_2 time for ^{55}Mn nuclei, which becomes comparable with the $\tau_{1-2}^{\min}$ time, and, as shown below, because of a decrease in the number of nuclei that contribute to the resonance, that is, a decrease in the volume of ferromagnetic regions.

Note that, along with the signal of Mn^{4+} ions, the low-temperature NMR spectrum of LMO contains a satellite line at $f \approx 295$–300 MHz. It is likely that this line corresponds to the presence in LMO of manganese sites with a decreased Mn–O distance, that is, Mn–O bonds more covalent in character. The presence in LMO of manganese sites with different Mn–O distances along with a large number of vacancies in lanthanum sites can also be the reason for broadening X-ray reflections in this sample compared with LSMO. A similar shape of ^{55}Mn NMR spectra was observed in [19] for self-doped lanthanum manganite.

The temperature evolution of the ^{55}Mn NMR spectra of LMO shown in Fig. 3 is evidence of the presence of two types of ferromagnetic regions with different dynamics of electron holes in manganese sites. The NMR spectrum of regions of type I consists of a single line ($f \approx 380$ MHz at $T = 77$ K) corresponding to fast motion of electron holes over manganese sites; this line is observed in a broad temperature range. The NMR spectrum of regions of type II is complex. It consists of two lines corresponding to quasi-localized Mn^{4+} ions ($f \approx 325$ MHz) and, as shown below, Mn^{3+} ions ($f \approx 390$ MHz); this spectrum is only observed at low temperatures. The contributions of regions of two types to the line at $f \approx 380$–390 MHz is clearly seen in the dependences of the spin echo amplitude on the τ_{1-2} time interval at various temperatures (Fig. 4). At $T \geq 63$ K,

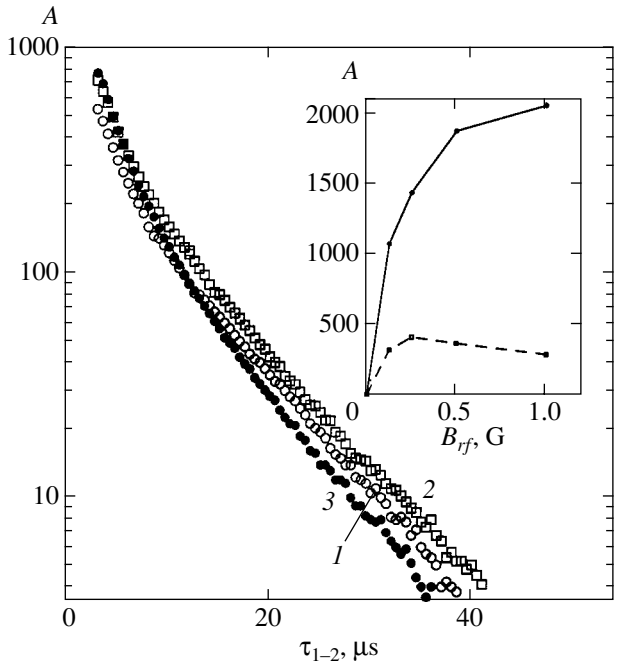

Fig. 5. Dependences of the spin echo signal amplitude for ^{55}Mn nuclei in LMO on the τ_{1-2} interval at f = 385 MHz and various radiofrequency field B_{rf} amplitudes, T = 21 K: (*1*) 0.125, (*2*) 0.25, and (*3*) 1 G. Shown in the inset are the B_{rf} dependences of the amplitudes of the rapidly relaxing (T_2 = 5 μs, the solid curve) and slowly relaxing (T_2 = 17 μs, the dashed curve) echo signal components obtained by decomposing experimental curves into two constituents.

these dependences are well described by a single exponential function, namely,

$$A(\tau_{1-2}) = A_0 \exp(-2\tau_{1-2}/T_2), \quad (1)$$

where T_2 is the spin–spin relaxation time in regions of type I. At T < 63 K, ^{55}Mn nuclei in regions of type II make an additional rapidly relaxing contribution to the spin echo amplitude. Such a contribution in ferromagnets can in principle be related to an admixture of a signal of more rapidly relaxing nuclei in domain boundaries. The NMR amplification coefficient η for nuclei in domain boundaries should then be substantially (as a rule, by one to two orders of magnitude) larger than for nuclei within domains because of the high susceptibility of reversible domain boundary displacements [29]. However, it follows from Fig. 5 that, for our samples, the optimal radiofrequency field amplitude B_{rf} for exciting the rapidly relaxing component of the NMR signal is substantially larger than for the slowly relaxing component. It follows that, as distinguished from the situation with domains and domain boundaries, the η value for the rapidly relaxing component in our spectra is four times lower than η for the slowly relaxing component. We therefore deal with signals of two types of ferromagnetic regions (further called phases I and II) with different effective magnetic fields of crystallographic anisotropy H_A. For phase II, this anisotropy is approximately four times larger ($\eta \sim H_n/H_A$, where H_n is the hyperfine magnetic field).

The presence of the contributions of two phases to the line at $f \approx$ 380–390 MHz also follows from spin–lattice relaxation measurements. The curves of restoration of the longitudinal nuclear magnetization component $M(t)$ at two τ_{1-2} interval values are shown in the insets in Fig. 6. At τ_{1-2} = 14 μs, when the observed signal is determined by the contribution of nuclei in phase I, the curve is well described by the exponential function:

$$A(t) = A_0[1 - \exp(-2t/T_1)], \quad (2)$$

where T_1 = 1.04 ms is the spin–lattice relaxation time in phase I. At τ_{1-2} = 3.5 μs, the curve of restoration of $M(t)$ can be described by the superposition of two exponential functions with relative amplitudes of $A_0^{\mathrm{I}} = 0.36$ and $A_0^{\mathrm{II}} = 0.64$ determined from T_2 measurements and relaxation times T_1 = 1.02 and 0.143 ms for phases I and II, respectively. Note the close agreement between the T_1 times for phase I obtained in two independent experiments. The contribution of phase I to the NMR spectra recorded at low temperatures is schematically shown in Fig. 3 by dashed lines. The amplitude of this contribution was determined by decomposing the dependences of the spin echo amplitude on the τ_{1-2} interval (Fig. 4) into two components.

The results of studying the dynamics of nuclear spins for the line at f = 325 MHz, which corresponds to phase II alone, show that the dependence of the spin echo amplitude on the τ_{1-2} interval is only approximately described by a simple exponential function; that is, in this phase also, we observe a distribution of times T_2. Interestingly, the shape of the curves obtained at various B_{rf} field amplitudes remains almost unchanged, which is evidence of the homogeneity of the NMR amplification factor and, therefore, magnetic anisotropy in phase II.

The temperature evolution of the ^{55}Mn NMR spectra of LSMO is shown in Fig. 7. The shape of the spectrum for this composition weakly changes as temperature varies. The observed decrease in signal intensities as the temperature increases is caused by a decrease in both the T_2 time for nuclei within domains and the volume of the ferromagnetic phase.

At low temperatures (T < 120 K), the shape of the line at $f \approx$ 385 MHz depends on the τ_{1-2} interval (Fig. 8a) because of the characteristic frequency dependence of T_2 with a minimum at the center of the NMR line (Fig. 8c). This feature is related to the Suhl–Nakamura interaction between nuclear spins via virtual spin

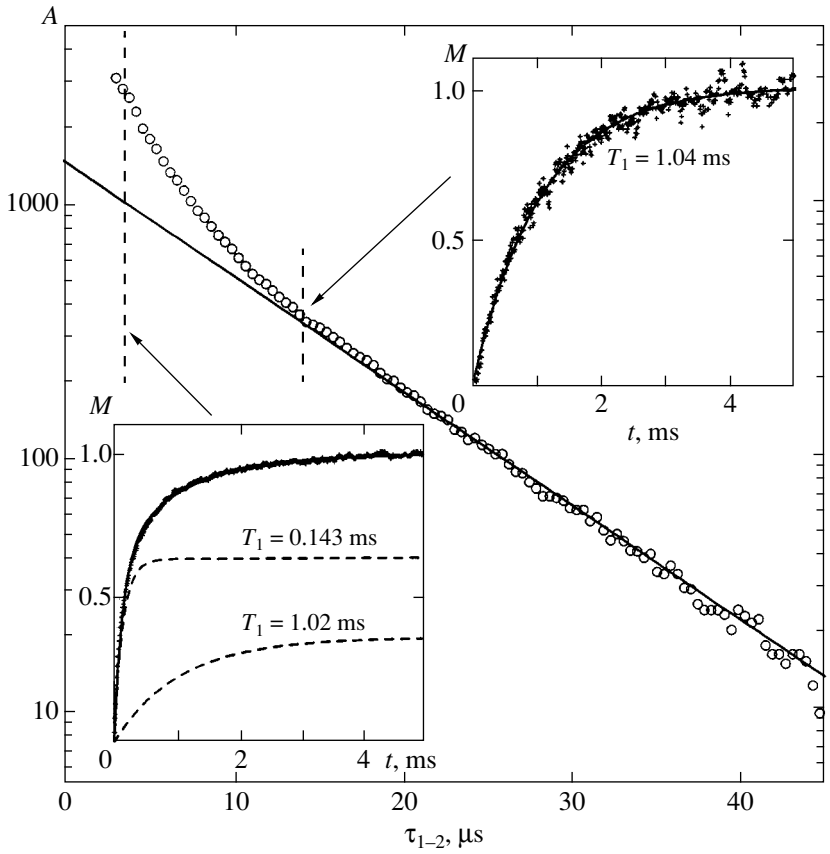

Fig. 6. Dependence of the amplitude of the ^{55}Mn echo signal in LMO on the τ_{1-2} interval at f = 385 MHz, B_{rf} = 0.25 G, and T = 21 K. Shown in the insets are the curves of restoring the longitudinal nuclear magnetization component $M(t)$ recorded with τ_{1-2} = 3.5 and 14 μs.

waves [29], which has the highest intensity at the center of the line because most nuclear spins precess at the corresponding frequency or near it. Interestingly, this mechanism is not effective for the line at $f \approx 330$ MHz, not only at T = 63 K, as is shown in Fig. 8, but also at the lower temperatures.

Another frequency-dependent dynamics of nuclear spins that make contributions to the line at $f \approx 385$ MHz appears at $T > 120$ K. Namely, as the τ_{1-2} time interval increases, the width of the NMR line first decreases and then, after a certain τ_{1-2} value is attained, ceases to change. According to [30], this behavior can be described on the assumption that two unresolved NMR lines (A_1 and A_2) with different times T_2 are present rather than one line. Such a dynamics of nuclear spins is evidence that, as in many manganites with metallic conductivity, ferromagnetic phase I in LSMO is itself spatially inhomogeneous and consists of nanoscopic regions with different mobilities of electron holes and different double exchange intensities [30].

The data obtained in studying spin–spin relaxation of ^{139}La nuclei in LMO and LSMO are shown in Fig. 9. At all temperatures, the dependences of the spin echo amplitude on the τ_{1-2} time interval were recorded at the center of the NMR line. For LMO, these dependences, like those for ^{55}Mn nuclei in phase II, are not described by a simple exponential function. An increase in temperature causes a sharp acceleration in relaxation. Simultaneously, the amplitude of the slowly relaxing contribution decreases (the dashed lines in Fig. 9). As a result, the NMR signal completely disappears at $T >$ 60 K. This is evidence of inhomogeneous relaxation of ^{139}La nuclei with a broad distribution of times T_2, which corresponds to the results obtained in [16, 17]. For LSMO, spin–spin relaxation is slower and weakly depends on temperature, and the NMR signal amplitude is approximately constant up to T = 160 K. A further increase in temperature decreases the intensity of the signal because of a decrease in the volume of the ferromagnetic phase. Note that the data on LSMO given in Fig. 9 were not corrected for the $A \propto 1/T$ temperature dependence of the NMR signal amplitude to prevent the superposition of the curves. Also note that, at low temperatures ($T \le 90$ K), the shape of spin echo decay curves contains a Gaussian contribution of the form $A = A_0 \exp[-(\tau/T_{2G})^2]$ along with the Lorentz contribution of form (1). The Gaussian contribution causes deviations of the experimental dependence at large

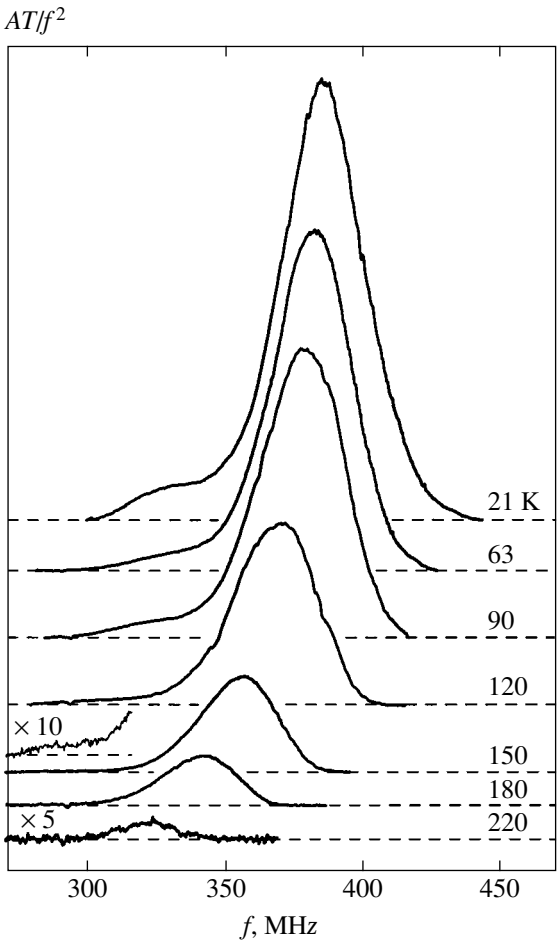

Fig. 7. Temperature evolution of the ^{55}Mn NMR spectra in LSMO. The spectra were recorded with $\tau_{1-2} = 3$ μs.

time lags from the Lorentz component shown by the dashed line.

4. DISCUSSION

The main results that follow from the NMR spectra and relaxation data on LMO and LSMO are summarized in Figs. 10–12. Figure 10 contains the temperature dependences of spin–spin relaxation times T_2 of ^{55}Mn and ^{139}La nuclei. If the relaxation cannot be described by a simple exponential function (Mn^{4+} and lanthanum sites in LMO), the T_2 values for the slowly relaxing component are given, which, as follows, for instance, from Fig. 9, well characterizes the temperature evolution of the dynamics of nuclear spins. For the same reason, the T_2 values for ^{139}La in LSMO at low temperatures correspond to the initial exponential portion of the spin echo decay curves ($\tau_{1-2} < 400$ μs). The relative volume of regions A_1 with a higher mobility of electron holes for phase I in LSMO is shown in the inset in Fig. 10 as a function of temperature. The temperature dependences of the relative ^{55}Mn NMR frequency changes are compared in Fig. 11 with the relative change in the spontaneous magnetic moment. Lastly, the temperature dependences of the volumes of magnetic phases are shown in Fig. 12; these volumes are proportional to the areas of the corresponding ^{55}Mn NMR lines corrected taking into account spin–spin relaxation and the NMR amplification factor. They are shown in comparison with the temperature dependences of the spontaneous magnetic moment. For LMO, only estimates of the corresponding volumes at 21 K are given because of the inhomogeneous distribution of times T_2 in phase II.

Taking into account these data, the phase composition and properties of separate phases in the manganites under consideration can be interpreted as follows. The ferromagnetic state in the LMO and LSMO samples is inhomogeneous and contains two ferromagnetic phases with different mobilities of electron holes in manganese sites. The ^{55}Mn NMR spectrum of phase I contains a single line corresponding to fast motion of electron holes over manganese sites. The temperature dependence of the spin–spin relaxation time (Fig. 10) obeys the exponential law with a slope (on the logarithmic scale) characteristic of single-phase manganites with metallic conductivity below T_C [31]. Lastly, as in other conducting manganites [30], the dynamics of nuclear spins in LSMO is indicative of the appearance of internal nanoscopic inhomogeneity of this phase at $T >$ 100 K; this refers to the mobility of electron holes and, accordingly, double exchange intensity. Taking these observations into account, phase I should be treated as a ferromagnetic metallic phase with the double exchange conductivity mechanism. This phase predominates in LSMO, which results in its metallic conductivity at $T < T_C$. In LMO, this is a residual phase whose volume only amounts to 1.5–3% of the total volume of ferromagnetic regions at low temperatures (Fig. 12).

The results obtained for ferromagnetic phase II in LMO are of the greatest interest. This phase gives a complex ^{55}Mn NMR spectrum, and its spin–spin relaxation time sharply decreases as temperature increases. As a result, the corresponding NMR signals of LMO can only be observed at $T < 60$ K. Important characteristics of this phase are also a sharp decrease in the spin–spin relaxation time of ^{139}La nuclei as temperature increases and inhomogeneous relaxation of both ^{55}Mn and ^{139}La characterized by a broad distribution[1] of times T_2. In [16], the anomalously fast spin–spin relaxation of ^{139}La nuclei in ferromagnetic insulating manganites is related to electric field gradient fluctuations caused by the diffusion of lattice distortions as a result of the slow motion of Jahn–Teller polarons. This sce-

[1] Note that, for LSMO, where the ^{139}La NMR signal largely originates from the ferromagnetic metallic phase, the spin–spin relaxation of ^{139}La nuclei (Fig. 10, LSMO) comparatively weakly depends on temperature, as in other conducting manganites [31, 32].

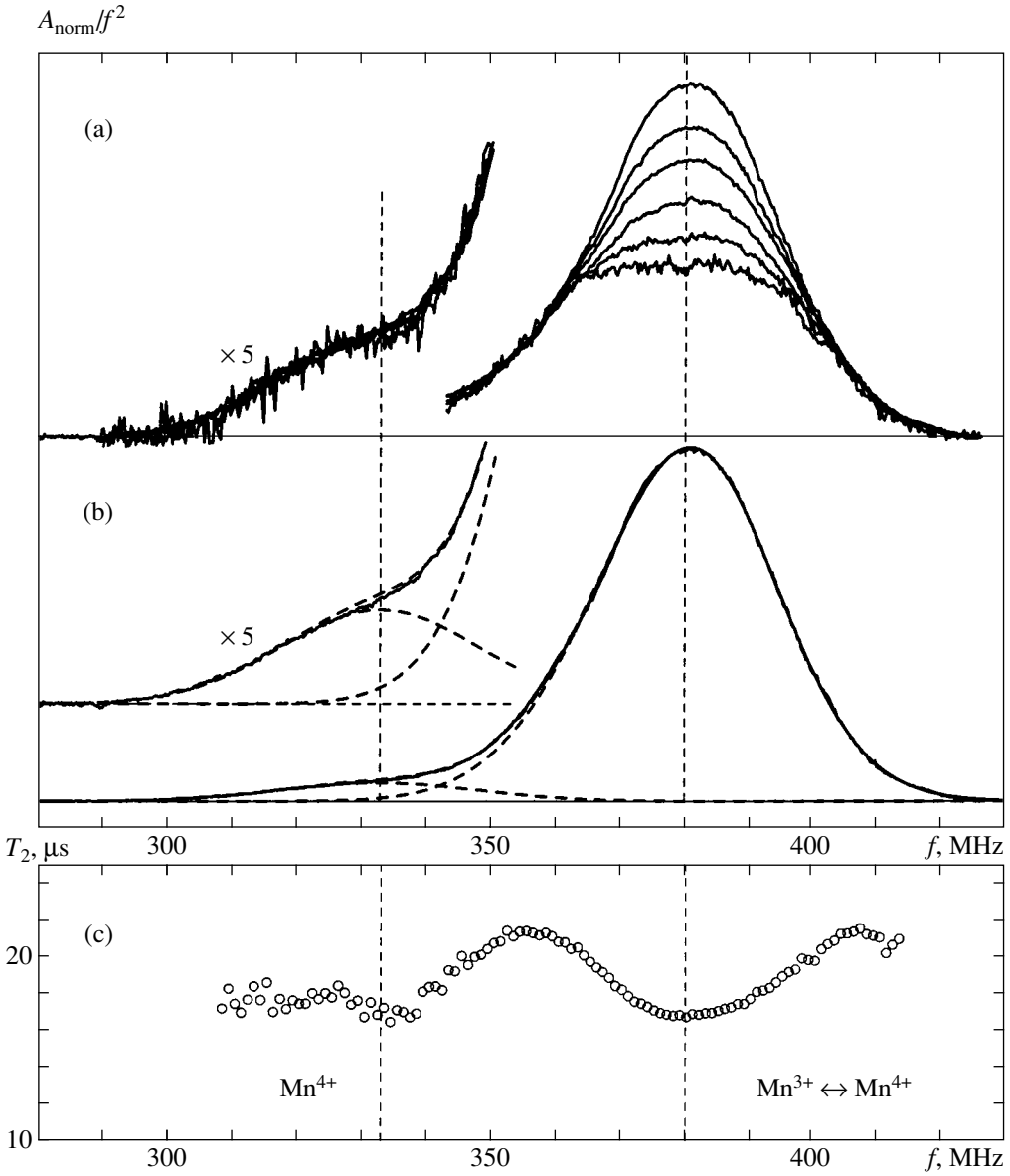

Fig. 8. Evolution of the shape of the ^{55}Mn NMR spectra of LSMO at T = 63 K and τ_{1-2} values of (a) 3, 6.5, 11, 17, 24, and 31 μs (curves from top to bottom) and (b) 3 μs [the spectrum is decomposed into two lines (the dashed curves)]; (c) frequency dependence of time T_2.

nario is in agreement with the results obtained in this work. The motion of Jahn–Teller polarons causes not only electric field gradient fluctuations on ^{139}La nuclei, but also slow manganese valence fluctuations, which explains the sharp temperature dependence of T_2 for ^{55}Mn nuclei, similar to that observed for ^{139}La ones (Fig. 10). This dependence is, however, caused by fluctuations of the longitudinal component of the local magnetic field H_n on ^{55}Mn nuclei. As distinguished from the ferromagnetic metallic phase, where spin–spin relaxation is explained by field H_n fluctuations caused by rapidly moving electron holes ($f \gg f_{NMR}$) [31], we here deal with slow ($f \leq f_{NMR}$) field H_n fluctuations. The mean frequency of these fluctuations passes the f_{NMR} value as temperature decreases to 60 K, and individual contributions, which, in the ionic approximation, correspond to the Mn^{4+} and Mn^{3+} states, appear in the NMR spectrum. The time T_2 sharply increases as the characteristic frequency of fluctuations decreases compared with the NMR frequency. It follows that phase II can be identified with the ferromagnetic insulating phase, in which double exchange is considerably weakened and is largely caused by comparatively slow, thermally activated motions of Jahn–Teller polarons. This phase predominates in the LMO sample (Fig. 12) and, as shown below, is responsible for semiconductor-

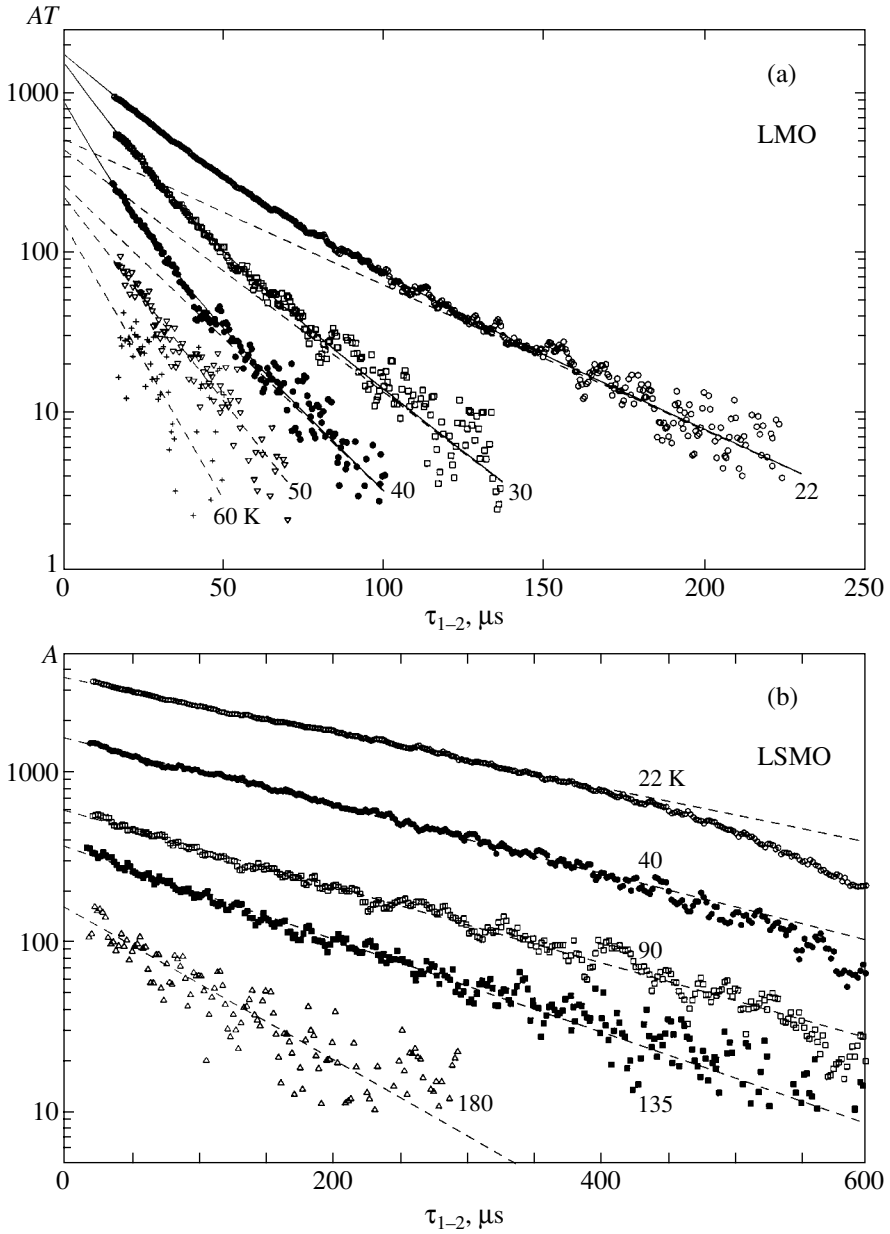

Fig. 9. Dependences of the spin echo amplitude of ^{139}La nuclei in LMO and LSMO on the τ_{1-2} time interval at various temperatures. The solid lines correspond to approximating the data by two exponential functions, and the dashed lines are the slowly relaxing approximation components.

type conductivity together with paramagnetic and/or antiferromagnetic regions. Although its slope is smaller than that observed for LMO, the line corresponding to Mn^{4+} in the spectrum of LSMO is likely to originate from the residual ferromagnetic insulating phase because the temperature dependence of time T_2 for this line is sharper than the corresponding dependence in the ferromagnetic metallic phase (Fig. 10). In principle, the presence of the line corresponding to Mn^{4+} ions in the ^{55}Mn NMR spectrum is not sufficient proof that the system is not single-phase and can also be a consequence of ferromagnetic state nonuniformity, when part of the holes are localized at low temperatures because of local disorder. However, the spin–spin relaxation of ^{55}Mn nuclei in Mn^{4+} sites should then be slower than in Mn^{4+}/Mn^{3+} sites, and an increase in temperature and hole delocalization can level times T_2 [33, 34]. As follows from Fig. 10, no such behavior is observed in our experiments.

It should be stressed that, according to the NMR data, it is precisely slow valence fluctuations rather than the localization of the Mn^{3+} and Mn^{4+} states that occur in the ferromagnetic insulating phase. If localized Mn^{3+} ions with nonzero orbital momenta, which are relax-

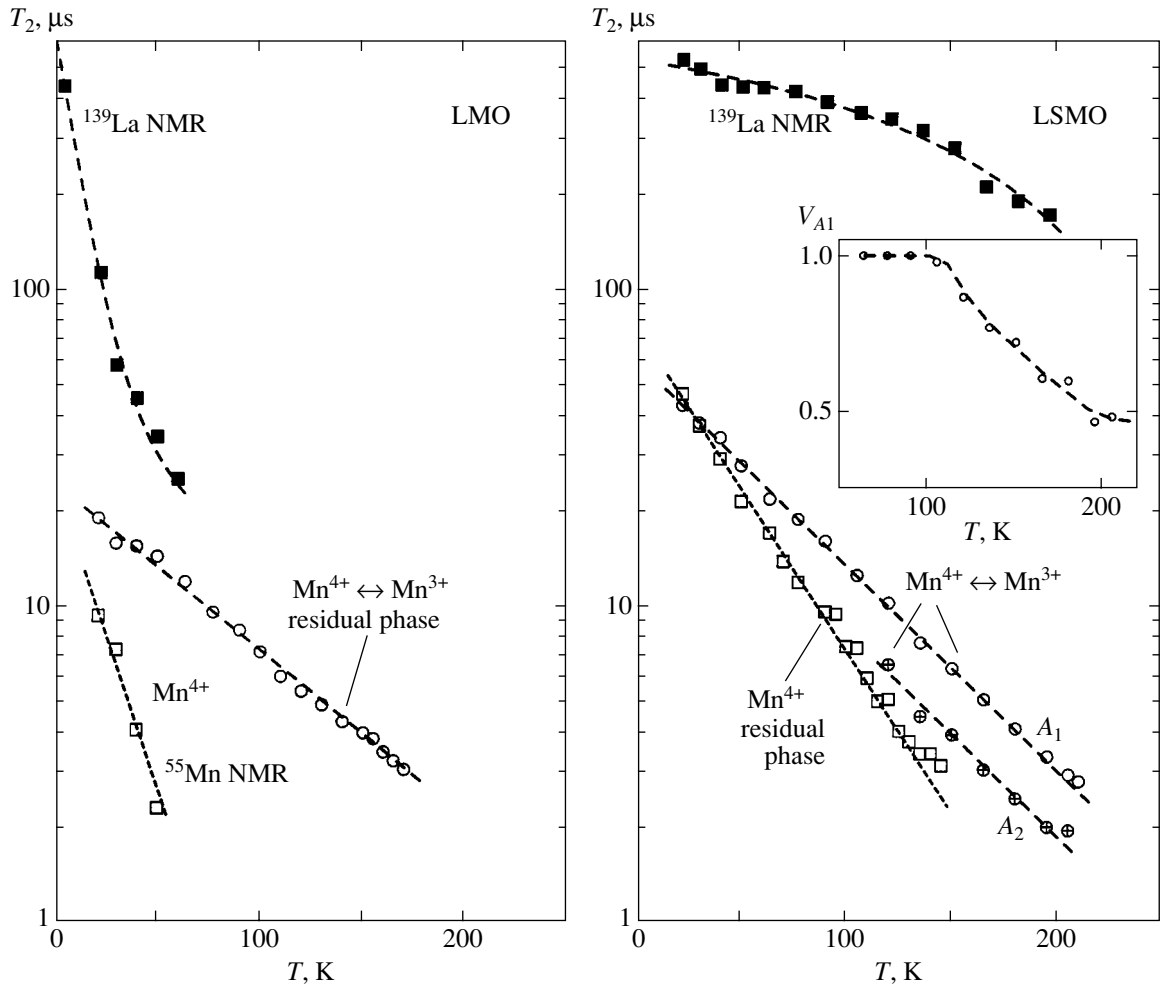

Fig. 10. Temperature dependences of spin–spin relaxation time T_2 for ^{139}La and ^{55}Mn nuclei in LMO and LSMO. Shown in the inset is the temperature dependence of the relative volume of phase A_1 in LSMO (see text).

ation centers with a strongly anisotropic spectrum of fluctuations, were present, it might be expected that nuclear magnetic relaxation would be determined by the spin–lattice contribution and, in the limiting case, we would have the equality $T_2 = 2T_1$. The time T_1 for the ferromagnetic insulating phase in LMO is substantially shorter than for the ferromagnetic metallic phase. Nevertheless, $T_2 \ll T_1$ in both cases, at least at $T \geq 21$ K (see Fig. 6), which corresponds to an isotropic spectrum of local field fluctuations on ^{55}Mn nuclei. The distribution of the times T_2 for the ferromagnetic insulating phase specified above is evidence of the inhomogeneity of Jahn–Teller polaron activation energies resulting from lattice disorder largely caused by vacancies in lanthanum sites in the system under consideration. Such an inhomogeneity can also be the reason for the strong magnetic crystallographic anisotropy of the ferromagnetic insulating phase compared with the ferromagnetic metallic phase. Note that antiferromagnetic interactions between localized states can frustrate the ferromagnetic order when holes slowly move over manganese sites; that is, the specified characteristics of the ferromagnetic insulating phase can also correspond to a state of the type of a cluster spin glass, whose features were observed in self-doped lanthanum manganites [14].

Next, consider the temperature dependences of the volumes of the ferromagnetic metallic and insulating phases. The temperature dependence of the ^{55}Mn NMR frequencies corresponds to a good approximation to the temperature dependence of local magnetic moments on the corresponding manganese sites [35, 36]. Figure 11 shows that the magnetic moment on Mn^{4+} sites (the ferromagnetic insulating phase) decreases as temperature increases more rapidly than the magnetic moment on Mn^{4+}/Mn^{3+} sites with a rapidly changing manganese valence (the ferromagnetic metallic phase), especially in LMO. A similar behavior was reported in [19]; it corresponds with the interpretation given above, according to which double exchange is weakened in the ferromagnetic insulating phase. It is noteworthy that, for both compositions, relative temperature-induced changes in

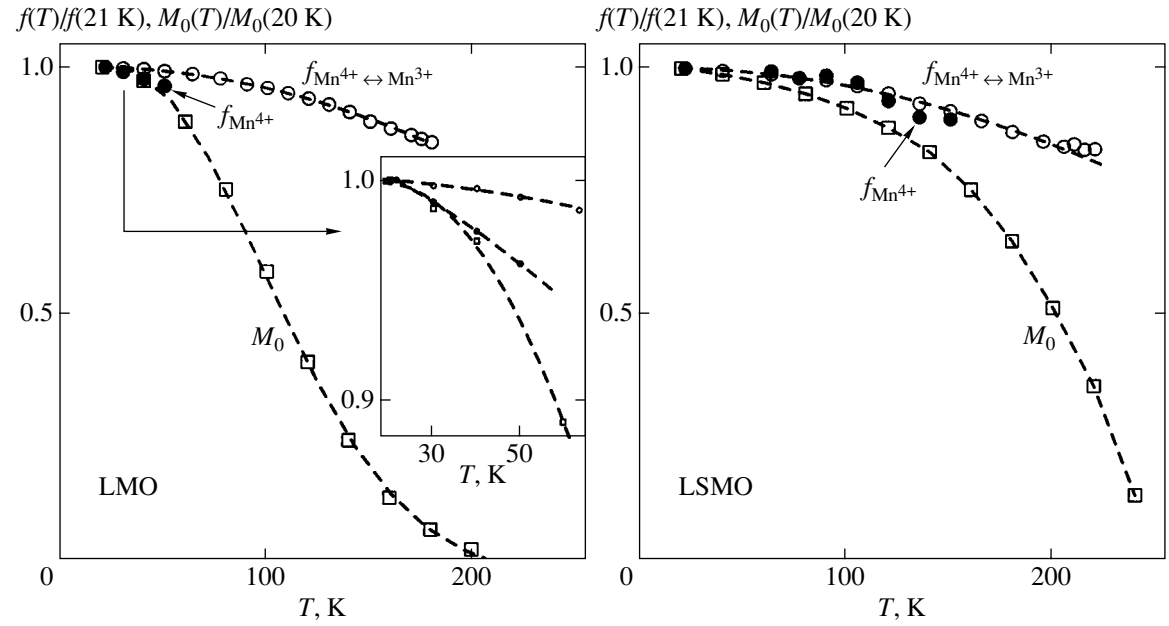

Fig. 11. Temperature dependences of changes in the relative ^{55}Mn NMR frequencies and relative spontaneous magnetic moment M_0 changes for LMO and LSMO.

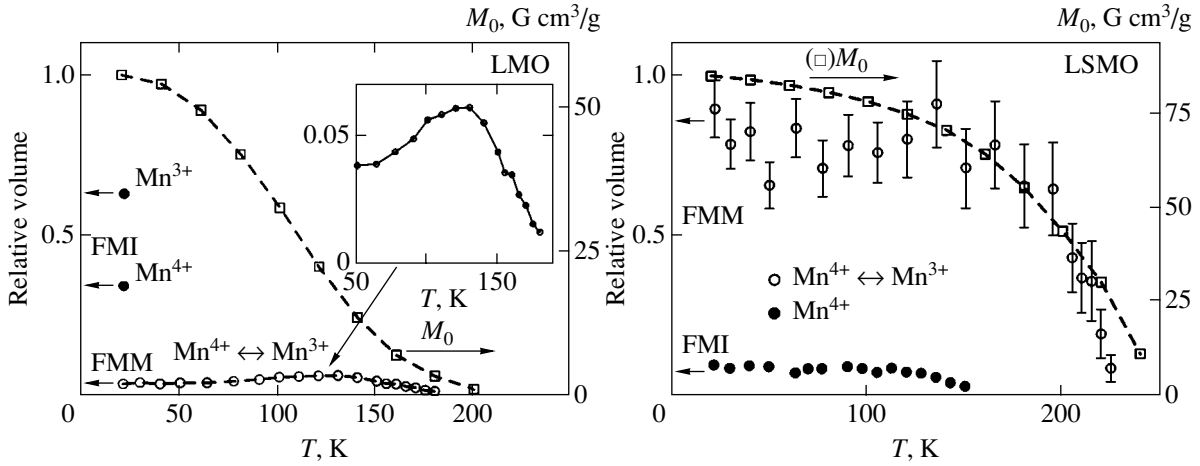

Fig. 12. Temperature dependences of the areas under the ^{55}Mn NMR lines and relative spontaneous magnetic moment M_0 changes for LMO and LSMO.

the NMR frequencies of both the ferromagnetic metallic and insulating phases are substantially smaller than changes in the spontaneous magnetic moment. This is evidence that a decrease in magnetization as temperature increases is largely caused by a decrease in the volume of ferromagnetic regions rather than the local moment on manganese sites. In other words, the transition from the ferromagnetic to the paramagnetic state is a smeared first-order phase transition, which occurs as a smooth redistribution of the volumes of ferromagnetic and paramagnetic (or antiferromagnetic) regions in a wide temperature range. This is clear from Fig. 12, where the temperature dependence of the spontaneous magnetic moment of LSMO is compared with the temperature dependences of the relative volumes of the ferromagnetic metallic and insulating phases determined from the NMR data.[2] The smeared character of the ferromagnet–paramagnet transition is the obvious reason why the dielectric–metal transition temperature for LSMO (T = 170 K) is lower than the Curie temperature T_C = 232 K formally determined as the inflection point of the susceptibility curve. This discrepancy does not contradict the double exchange concept. Actually, the dielectric–metal transition occurs when ferromagnetic

[2] Note that, because of a decrease in time T_2, the NMR spectra at $T >$ 210 K contain only the contribution of line A_1, which amounts to 50% of the total volume of the ferromagnetic metallic phase according to the data given in the inset in Fig. 10.

metallic regions with metallic conductivity occupy the larger part of the crystal volume.

For LMO, the NMR data can only be used to reliably determine the temperature dependence of the relative volume of the more slowly relaxing residual ferromagnetic metallic phase. The volume of the major ferromagnetic insulating phase is difficult to estimate because of inhomogeneity of spin–spin relaxation and its rapid acceleration at higher temperatures. Estimates based on the most reliable data obtained at $T = 21$ K show that the volume of the ferromagnetic insulating phase is about 97% of the total volume of ferromagnetic regions. Accordingly, the ferromagnetic metallic phase volume amounts to about 3% (see Fig. 12). The relative volume of the ferromagnetic metallic phase in this compound can also be estimated by directly comparing the intensities of the NMR signals of LMO and LSMO at $T = 77$ K, which gives a value of about 1.5% of the total volume of ferromagnetic regions. Although we were able to determine the temperature dependence of local magnetic moments in the ferromagnetic insulating phase only to $T \approx 60$ K, the NMR frequency already at these comparatively low temperatures decreased slower than the spontaneous moment (see the inset in Fig. 11). We can therefore suggest that the temperature dependence of the spontaneous magnetic moment in LMO is largely determined by a decrease in the volume of the ferromagnetic insulating phase. The interpretation given above, according to which magnetic interactions are considerably weaker in the ferromagnetic insulating phase, leads us to expect a more rapid decrease in the volume of the ferromagnetic insulating phase in LMO compared with the volume of the ferromagnetic metallic phase in LSMO. Also note that the spontaneous magnetic moment of LSMO per manganese ion at $T = 20$ K, $M_0 \approx 3.53\mu_B$, is close to the spin value $M_0 \approx 3.8\mu_B$ expected for the complete ferromagnetic ordering of magnetic moments. At the same time, the observed magnetic moment of LMO is substantially smaller, $M_0 \approx 2.31\mu_B$, which can either be related to phase separation into the ferro- and paramagnetic (or antiferromagnetic) phases even at this temperature or be evidence of a spin glass–type state, as mentioned above.[3] The complex shape of the temperature dependence of resistance in LMO with a temperature-dependent activation energy (see Fig. 1) can therefore be related to a redistribution of the volumes of paramagnetic (or antiferromagnetic at low temperatures) and ferromagnetic insulating regions in a wide temperature range.

Interestingly, the relative volume of the ferromagnetic metallic phase in LMO shows some tendency to increase in the temperature range 60–130 K in spite of the temperature-induced sharp decrease in the volume of ferromagnetic regions (see the inset in Fig. 12). This behavior can be explained by the delocalization of part of holes in the ferromagnetic insulating phase as temperature increases [33, 34]. Accordingly, the volume of the residual ferromagnetic insulating phase in LSMO also shows a tendency to decrease, however, at higher temperatures ($T \geq 130$ K), which can be explained by a lower mean activation energy of Jahn–Teller polarons in LSMO compared with LMO. According to the above interpretation of the relaxation of ^{55}Mn nuclei in the ferromagnetic insulating phase, the decrease in activation energy should manifest itself by slower relaxation than in LMO, which is actually observed.

The substantial magnetic inhomogeneity of ferromagnetic manganites with cationic vacancies observed in this work by the NMR method is not likely to be related to the defect structure of the polycrystalline samples. Rather, it is the intrinsic property of these samples. Because the NMR method is local, polycrystalline sample defects, whose volume is substantially smaller than that of crystallites, have a weak influence on the reliability of the results obtained in both recording NMR spectra and studying the dynamics of nuclear spins. This, for instance, follows from the results obtained in [21], in which magnetic inhomogeneity of $La_{1-x}Ca_xMnO_3$ single crystals ($0 \leq x \leq 0.15$) grown by zone melting and polycrystalline samples ($x = 0.2$ and 0.3) was studied.

5. CONCLUSIONS

We used the local NMR method to show that the ferromagnetic state in $(La_{1-x}Sr_x)_{1-\delta}MnO_3$ manganites was spatially inhomogeneous because of the simultaneous presence of two types of qualitatively different ferromagnetic regions with different mobilities of electron holes in manganese sites. The ferromagnetic conducting phase is determined by the double exchange mechanism, whereas double exchange is considerably weakened in the ferromagnetic insulating phase and mainly occurs as a result of slow, thermally activated motions of Jahn–Teller polarons at a characteristic frequency comparable with the NMR frequency. The dynamics of nuclear spins is evidence of the internal inhomogeneity of both phases. Additional inhomogeneity is caused by the smearing of the ferromagnet–paramagnet phase transition. A comparison of the NMR and magnetization data shows that this transition occurs as a smooth redistribution of the volumes of ferromagnetic and paramagnetic (or antiferromagnetic)

[3] In principle, the decreased magnetic moment value can also be related to magnetic structure noncollinearity. According to numerous experimental and theoretical results (e.g., see [1, 37), phase separation or spin glass state models should, however, be given preference for the compounds under consideration. In addition, a homogeneous noncollinear model of the magnetic structure of LMO is incapable of explaining the observed temperature dependences of the magnetic moment and NMR frequency. Such an explanation would require us to assume that the magnetic structure continuously changes toward antiferromagnetic ordering as temperature increases. At the same time, magnetic ordering in LMO exists up to $T \approx 200$ K, which is substantially higher than the antiferromagnetic ordering temperature $T_N \approx 140$ K in the stoichiometric $LaMnO_3$ compound.

regions in a wide temperature range. The ferromagnetic conducting phase is a residual phase in self-doped $La_{1-\delta}MnO_3$ but becomes predominant as the concentration of strontium increases to $x = 0.125$. Combined with a narrowing of the temperature interval of the ferromagnet–paramagnet transition, this causes a change in the character of electric conductivity from semiconductor-type conductivity at $x = 0$ to metallic conductivity at low temperatures for the $x = 0.075$ and 0.125 compositions.

ACKNOWLEDGMENTS

The authors thank P. Novák, Z. Jirák, and J. Hejtmanek of the Institute of Physics (Prague) for valuable discussions of the experimental results. This work was financially supported by the Ukrainian Foundation for Basic Research (project no. F7/471-2001).

REFERENCES

1. É. L. Nagaev, Usp. Fiz. Nauk **166**, 833 (1996) [Phys. Usp. **39**, 781 (1996)].
2. A. P. Ramirez, J. Phys.: Condens. Matter **9**, 8171 (1997).
3. E. Dagotto, T. Hotta, and A. Moreo, Phys. Rep. **344**, 1 (2001).
4. C. Zener, Phys. Rev. **82**, 403 (1951).
5. J. H. van Santen and G. H. Jonker, Physica (Amsterdam) **16**, 599 (1950).
6. G. J. Snyder, R. Hiskes, S. DiCarolis, *et al.*, Phys. Rev. B **53**, 14434 (1996).
7. R. M. Thomas, L. Ranno, and J. M. D. Coey, J. Appl. Phys. **81**, 5763 (1997).
8. D. Cao, F. Bridges, M. Anderson, *et al.*, Phys. Rev. B **64**, 184409 (2001).
9. A. K. Heilman, Y. Y. Xue, B. Lorenz, *et al.*, Phys. Rev. B **65**, 214423 (2002).
10. Y. Moritomo, A. Asamitsu, and Y. Tokura, Phys. Rev. B **56**, 12190 (1997).
11. S. De Brion, F. Ciorcas, G. Chouteau, *et al.*, Phys. Rev. B **59**, 1304 (1999).
12. Y. Endoh, O. Hino, S. Nohdo, *et al.*, Phys. Rev. Lett. **82**, 4328 (1999).
13. J.-S. Zhou, J. B. Goodenough, A. Asamitsu, and Y. Tokura, Phys. Rev. Lett. **79**, 3234 (1997).
14. L. Ghivelder, I. A. Castillo, M. A. Gusmão, *et al.*, Phys. Rev. B **60**, 12184 (1999).
15. B. Martinez, R. Senis, Ll. Balcells, *et al.*, Phys. Rev. B **61**, 8643 (2000).
16. G. Allodi, M. C. Guidi, R. De Renzi, *et al.*, Phys. Rev. Lett. **87**, 127206 (2001).
17. G. Papavassiliou, M. Belesi, M. Fardis, and C. Dimitropoulos, Phys. Rev. Lett. **87**, 177204 (2001).
18. Cz. Kapusta and P. C. Riedi, J. Magn. Magn. Mater. **196–197**, 446 (1999).
19. Cz. Kapusta, P. C. Riedi, W. Kocemba, *et al.*, J. Appl. Phys. **87**, 7121 (2000).
20. M. Belesi, G. Papavassiliou, M. Fardis, *et al.*, Phys. Rev. B **63**, 180406(R) (2001).
21. G. Allodi, M. C. Guidi, R. De Renzi, and M. W. Pieper, J. Magn. Magn. Mater. **242–245**, 635 (2002).
22. J. F. Mitchell, D. N. Argyriou, C. D. Potter, *et al.*, Phys. Rev. B **54**, 6172 (1996).
23. V. Golovanov, L. Mihaly, C. O. Homes, *et al.*, Phys. Rev. B **59**, 153 (1999).
24. R. Suryanarayanan, V. Gasumyants, and N. Ageev, Phys. Rev. B **59**, R9019 (1999).
25. V. Markovich, E. Rozenberg, G. Gorodetsky, *et al.*, Phys. Rev. B **63**, 054423 (2001).
26. D. Louca, E. L. Brosha, and T. Egami, Phys. Rev. B **61**, 1351 (2000).
27. F. Prado, R. D. Sanchez, A. Caneiro, *et al.*, J. Solid State Chem. **146**, 418 (1999).
28. G. Matsumoto, J. Phys. Soc. Jpn. **29**, 615 (1970).
29. E. A. Turov and M. P. Petrov, *Nuclear Magnetic Resonance in Ferro- and Antiferromagnets* (Nauka, Moscow, 1969; Wiley, New York, 1972).
30. M. M. Savosta and P. Novák, Phys. Rev. Lett. **87**, 137204 (2001).
31. M. M. Savosta, V. A. Borodin, and P. Novák, Phys. Rev. B **59**, 8778 (1999).
32. G. Allodi, R. De Renzi, and G. Guidi, Phys. Rev. B **57**, 1024 (1998).
33. M. M. Savosta, P. Novák, Z. Jirák, *et al.*, Phys. Rev. Lett. **79**, 4278 (1997).
34. M. M. Savosta and E. E. Solov'ev, Pis'ma Zh. Éksp. Teor. Fiz. **70**, 522 (1999) [JETP Lett. **70**, 533 (1999)].
35. M. M. Savosta, V. A. Borodin, P. Novák, *et al.*, Phys. Rev. B **57**, 13379 (1998).
36. P. Novák, M. Maryško, M. M. Savosta, and A. N. Ulyanov, Phys. Rev. B **60**, 6655 (1999).
37. A. Moreo, S. Yunoki, and E. Dagotto, Science **283**, 2034 (1999).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 587–592.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 650–655.

SOLIDS
Structure

Magnetic Ordering in a Random System of Point Ising Dipoles

E. Z. Meilikhov

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia
e-mail: meilikhov@imp.kiae.ru

Received December 30, 2002

Abstract—An exact expression for the random magnetic field distribution function is obtained for a simple model of a random system of Ising magnetic dipoles. The magnetic phase diagram for such a system is determined within the framework of the random mean field theory. The magnetic characteristics of individual phases of this system are described. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The magnetic state of a random system of interacting magnetic dipoles cannot be described within the framework of the traditional mean field theory since the latter is applicable only for homogeneous systems in which local fields are identical for all interacting magnetic fields. In contrast to thermal fluctuations, spatial fluctuations of a local field, which exist in a random system, prevent the establishment of a magnetic order even at zero temperature. For this reason, a correct description of such systems requires a more general approach.

Zhang and Widom [1] used one of the possible ways for generalizing the mean field theory to analyze a random system of Ising spherical dipoles of a finite diameter.[1] They proved that the ground state of such a system becomes ferromagnetic only when the mean distance between spherical dipoles is comparable to their size; otherwise (in the case of large distances between dipoles), the system is paramagnetic even at $T = 0$. In other words, the ferromagnetic state is possible only in a system in which the dipole concentration exceeds a certain critical value; this is generally in accordance with tendencies in the behavior of dipole system with configuration disorder [3].

However, the above result was obtained under two assumptions, both of which are dubious. First, the spatial dipole distribution function was factorized as is usually done in an analysis of random systems. This function was represented by the product of two identical one-particle distribution functions $g(r)$ each of which depends on only one spatial coordinate, viz., the distance r between a given particle (dipole) and the other particle placed at the origin): $g(r) \propto r^2$ for $r > a$ (a is the particle diameter) and $g(r) = 0$ for $r < a$. This is justified for a system of point particles ($a = 0$) but is not observed in the case considered here since the arrangement of a finite-size particle is determined by the position of not one, but many neighboring particles even in the absence of magnetic correlations. One could hope that this approximation would not lead to considerable errors in the case of strongly rarefied systems, for which the probability of a close neighborhood of more than two particles is low ($na^3 \ll 1$, n being the particle concentration), but it was used in [1] for $na^3 \sim 1$. Second, the approximation used is equivalent to the assumption that the correlation length of spatial fluctuations of the magnetic moment directions of dipoles is larger than the dipole diameter, which can also only be justified for strongly rarefied systems.

These assumptions set a limit on the applicability of the obtained conclusions. In this study, another approach is proposed, which makes it possible to obtain an exact result for point Ising dipoles and to trace the modification of this result for finite-size dipoles. Using this approach, a magnetic phase diagram is obtained for a random system of Ising dipoles and the magnetic parameters of individual phases of this system are determined.

2. GENERALIZED MEAN FIELD THEORY FOR POINT (SPHERICAL) DIPOLES

The magnetic moment $\mathbf{m}_i$ of Ising dipoles can assume only two values: $\mathbf{m}_i = \pm m\mathbf{e}_0$, where $\mathbf{e}_0$ is the unit vector parallel to the direction of the dipoles. In the traditional theory, it is assumed that each dipole is acted upon by the same local field $\mathbf{H}_0$, which determines the mean value $\langle \mathbf{m}_i \rangle_T$ of the magnetic moment of any magnetic dipole (angle brackets correspond to averaging over the ensemble and subscript T corresponds to ther-

[1] The energy of interaction of uniformly magnetized spherical dipoles (one-domain spherical grains) coincides with the energy of interaction of two equivalent point dipoles [2].

modynamic averaging). For Ising dipoles, $\mathbf{H}_0 \parallel \mathbf{e}_0$ and we obtain

$$\langle \mathbf{m}_i \rangle_T = \frac{\sum_{\mathbf{m} = \pm m\mathbf{e}_0} \mathbf{m} \exp(\mathbf{m}\mathbf{H}_0/kT)}{\sum_{\mathbf{m} = \pm m\mathbf{e}_0} \exp(\mathbf{m}\mathbf{H}_0/kT)} = m\mathbf{e}_0 \tanh\left(\frac{mH_0}{kT}\right), \quad \mathbf{I} = n\langle \mathbf{m}_i \rangle_T, \tag{1}$$

where $\mathbf{I}$ is the system magnetization and n is the dipole concentration. For a sample in the form of a long cylinder whose axis is parallel to the direction of magnetic dipoles, the local field $\mathbf{H}_0$ is also parallel to the cylinder axis and is proportional to the magnetization:

$$H_0 = \frac{4\pi}{3} I + H_3. \tag{2}$$

Here,

$$H_3 = m \sum_{ikl} \frac{3\cos^2 \alpha_{ikl} - 1}{\rho_{ikl}^3} \tag{3}$$

is the component of the local field produced by all dipoles of the Lorentz sphere along the direction of $\mathbf{e}_0$; ρ_{ikl} is the distance between the chosen dipole placed at the center of the sphere (origin of the coordinate system) and a dipole in position (ikl); and α_{ikl} is the angle formed by the line connecting these dipoles with the direction of $\mathbf{e}_0$.

Let us suppose that a system with a random arrangement of dipoles is in the ferromagnetic state characterized by the average magnetization $\mathbf{I} \parallel \mathbf{e}_0$. Local magnetic fields H_3 are different for different dipoles and can be characterized by the distribution function $F(j; H_3)$, which generally depends on the relative magnetization $j \equiv I/mn$ of the system determined by the fraction $\eta = (1/2)(1 + j)$ of dipoles whose average magnetic moments are directed along the magnetization $\mathbf{I}$ of the system [$F(1; H_3)$ is the distribution function in the case when all the dipoles in the Lorentz sphere are parallel to one another].

The generalized mean-field theory proposed in [1] for a random system of spherical dipoles is based on an approximate and not quite correct (see the Introduction) method for calculating the distribution function $F(1; H_3)$ for local magnetic fields. At the same time, this function can be determined exactly using a simple but quite reasonable model. For this purpose, we apply the Markov method for determining the probability of the sum of a large number of random quantities [4], according to which

$$F(1; H_3) = \frac{1}{2\pi} \int_{-\infty}^{\infty} A(s) \exp(-isH_3) ds,$$
$$A(s) = \lim_{N \to \infty} \left[\int_{\alpha\rho = 0}^{\rho_{\max}} \int \exp[ish(\rho, \alpha)] \tau(\rho, \alpha) d\alpha s\rho \right]^N, \tag{4}$$

where $h(\rho, \alpha) = m(3\cos^2\alpha - 1)/\rho^3$ is the effective magnetic field produced by a point dipole with coordinates ρ, α at the origin (in the case considered here, this field is equal to the component of the magnetic field of the dipole along its direction), $\tau(\rho, \alpha)$ is the distribution function for these coordinates, and N is the number of dipoles in a sphere of radius $\rho_{\max}$ over which integration is carried out. If we assume further that (i) the random nature of arrangement of dipoles does not change their average concentration n (i.e., $N \longrightarrow 4\pi\rho_{\max}^3 n/3$ for $\rho_{\max} \longrightarrow \infty$); (ii) the distribution of angles α is uniform; and (iii) coordinates ρ and α are not correlated, then

$$\tau(\rho, \alpha) d\rho d\alpha = \frac{\rho^2 d\rho \cdot 2\pi \sin\alpha d\alpha}{4\pi\rho_{\max}^3/3}. \tag{5}$$

Substituting Eq. (5) into (4), we obtain

$$A(s) = \exp[-nC(s)],$$
$$C(s) = 2\pi \int_{\alpha} \sin\alpha d\alpha \int_0^{\infty} \{1 - \exp[ish(\rho, \alpha)]\} \rho^2 d\rho. \tag{6}$$

While integrating with respect to ρ, we must use the above expression for $h(\rho, \alpha)$ and take into account the fact that

$$\int_0^{\infty} \frac{1 - e^{iu}}{u^2} du = \frac{\pi}{2}.$$

This gives

$$nC(s) = 2\frac{\pi^2 mns}{3} \int_0^{\pi/2} \sin\alpha \left|3\cos^2\alpha - 1\right| d\alpha = sh_0, \tag{7}$$

where $h_0 = 8\pi^2 mn/9\sqrt{3}$ is the characteristic field approximately equal to the field of a dipole separated from the origin by a distance on the order of the mean distance $n^{-1/3}$ between dipoles; factor 2 takes into account the existence of two regions equivalent to the angular interval $0 < \alpha < \pi/2$.

For the distribution function $F(1; H_3)$, we finally obtain

$$F(1; H_3) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\cos(|H_3|s)\exp(-|s|h_0)ds = \frac{1}{\pi}\left(\frac{h_0}{h_0^2 + H_3^2}\right). \quad (8)$$

This is a Lorentzian distribution[2] of width h_0, which is centered at the mean field $\langle H_3\rangle = 0$. The result obtained, which differs in principle from the Gaussian distribution predicted in [1] for a similar system (in the approximate model), can be easily interpreted if we assume that strong fields H are mainly produced by nearest neighbors. The law $w_\rho(\rho)$ for distribution of distances ρ to the nearest neighbor is defined by the formula [4]

$$w_\rho(\rho) = 4\pi\rho^2 n\exp(-4\pi\rho^3 n/3) \propto \rho^2,$$

while the field H produced by this distribution is proportional to $1/\rho^3$. It follows hence that

$$F(H) = w_\rho[\rho(H)]|d\rho/dH| \propto 1/H^2,$$

which correctly describes the "wings" of the Lorentzian distribution in the range of strong fields.

A remarkable property of the obtained distribution, which is not obvious at first glance, is that it predicts equal probabilities for fields H_3 of opposite polarities.[3] It will be shown below that such a symmetric spread of Lorentzian fields "hampers" the emergence of magnetic ordering in the system under investigation. As a matter of fact, the noted property is preserved for any random distribution of these directions in spite of the fact that distribution (8) was obtained for a random system in which the directions of the magnetic moments of all dipoles coincide. In other words, the emergence of a magnetic order in a random system of dipoles does not in any way affect the form of the distribution function for the Lorentzian field; i.e., $F(j; H_3) \equiv F(1; H_3)$.

Let us now demonstrate that magnetic ordering cannot exist in the random system of point Ising magnetic moments considered here. The magnetization of the system is defined as $\mathbf{I} = n\langle\mathbf{m}_i\rangle_T$, where the average magnetic moment $\langle\mathbf{m}_i\rangle_T$ must be calculated taking into account the spread in Lorentzian fields H_3 using the obvious generalization of relation (3) pertaining to a regular system:

$$I = n\langle m_i\rangle_T = mn\int_{-\infty}^{\infty}\tanh\left[\frac{m(4\pi I/3 + H_3)}{kT}\right]F(j; H_3)dH_3. \quad (9)$$

Using expression (8) for the distribution function $F(1; H_3) \equiv F(j; H_3)$, we derive the generalizing equation for mean field:

$$j = \frac{1}{\pi}\int_{-\infty}^{\infty}\tanh\left[\frac{1}{\Theta}\left(Au + \frac{4\pi}{3}j\right)\right]\frac{du}{1+u^2} \equiv -\frac{A}{\pi\Theta}\int_{-\infty}^{\infty}\frac{\arctan u\,du}{\cosh\left[\frac{1}{\Theta}\left(Au + \frac{4\pi j}{3}\right)\right]}, \quad (10)$$

where $j = I/I_s$ ($I_s = nm$) is the reduced magnetization of the system, $\Theta = kT/m^2n$ is the reduced temperature (equal to the ratio of the thermal energy to the characteristic energy of magnetic interaction between dipoles), and $A = 8\pi^2/9\sqrt{3} \approx 5.065$. The paramagnetic state ($j = 0$) is obviously a solution to this equation. Analysis shows that this equation has no other solutions;[4] this obviously corresponds to the absence of ferromagnetic ordering in a random system of point Ising dipoles. This conclusion coincides with the result obtained in [4], according to which the system could become ferromagnetic only for a high dipole concentration $n\rho_{\min}^3 > 0.300$, where $\rho_{\min}$ is the minimal spacing of spherical dipoles.[5] In our model, we assume that $\rho_{\min} = 0$ (not spherical but pointlike dipoles); for this reason, the dipole concentration is always "low."

[2] The fact that distribution (8) predicts that $\langle H_3^2\rangle = \infty$ is obviously connected with the assumption that the distance between dipoles can be infinitely small. Deviations from this distribution should be expected for fields $H_3 > m/\rho_{\min}^3$, where $\rho_{\min}$ is the minimal possible spacing of dipoles (e.g., their characteristic size).

[3] This is formally connected with the specific form of the angular dependence of the field of a point dipole, $h(\rho, \alpha) \propto (3\cos^2\alpha - 1)$, which, together with the angular dependence of distribution $\tau(\rho, \alpha) \propto \sin\alpha$, ensures the equality of the "angular factors" for the regions $3\cos^2\alpha - 1 < 0$ and $3\cos^2\alpha - 1 > 0$ corresponding to negative and positive values of field H_3.

[4] As $\Theta \longrightarrow 0$, the right-hand side of Eq. (10) asymptotically approaches $(2/\pi)\arctan(4\pi j/3A) < 8j/3A = 0.526j < j$.

[5] In [4], the criterion for the emergence of spontaneous magnetization in a random system of rigid spherical Ising dipoles of diameter a is written in the form $(\pi/6)na^3 > 0.157$. The minimal distance between the centers of dipoles in this system is $\rho_{\min} = a$, which leads precisely to the above condition $n\rho_{\min}^3 > 0.300$.

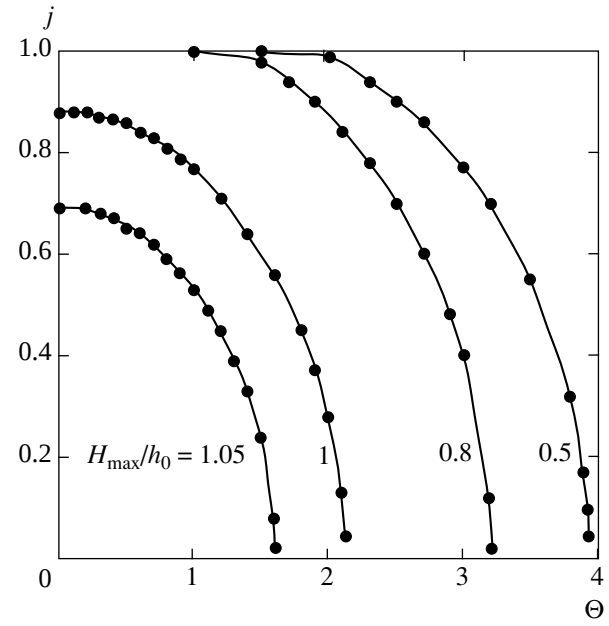

Fig. 1. Temperature dependences of spontaneous magnetization j of a system with a random field distribution truncated from above for different values of the truncation parameter (external field $H_e = 0$).

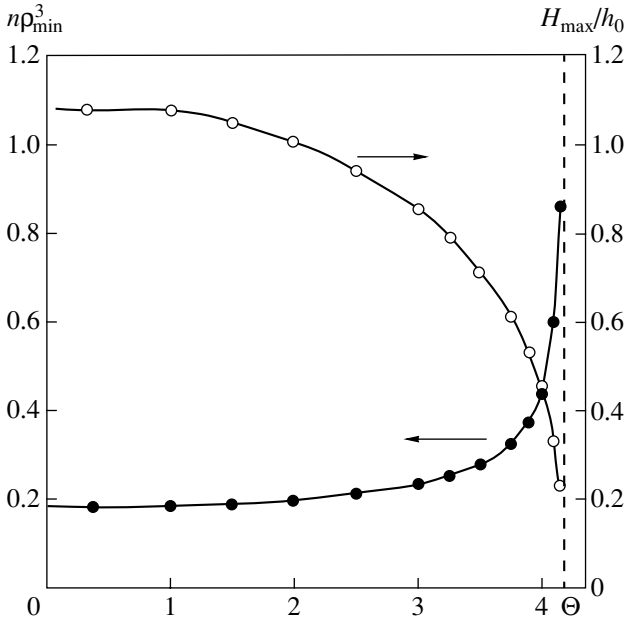

Fig. 2. Temperature dependences of the threshold value of parameter $n\rho_{min}^3$ for which a spontaneous magnetic moment arises in a system of random Ising dipoles and of the maximum value $H_{max} = m/\rho_{min}^3$ of random field H_3 corresponding to this value ($H_e = 0$).

The impossibility of unlimited decrease in the distance between dipoles leads to a "truncation" of the values of magnetic field from above by a value of $H_{max} \sim m/\rho_{min}^3$. For this reason, we can easily pass to spherical dipoles in the framework of our model by truncating the wings of distribution function (8), which correspond to fields $|H_3| > H_{max} = m/\rho_{min}^3$:

$$F(1; H_3) = \begin{cases} \left[2\arctan\left(\frac{H_{max}}{h_0}\right)\right]^{-1}\left(\frac{h_0}{h_0^2 + H_3^2}\right), & |H_3| \le H_{max} \\ 0, & |H_3| > H_{max}. \end{cases} \quad (11)$$

In this case, instead of Eq. (10), we obtain the following equation for the mean field:

$$j = \left[2\arctan\left(\frac{H_{max}}{h_0}\right)\right]^{-1} \times \int_{-H_{max}/h_0}^{H_{max}/h_0} \tanh\left[\frac{1}{\Theta}\left(Au + \frac{4\pi}{3}j\right)\right]\frac{du}{1+u^2}. \quad (12)$$

An analysis shows that this equation acquires the second (nonzero) root corresponding to spontaneous magnetization of the system for $H_{max}/h_0 < 1.08$ (at zero temperature). This condition can be written in the form of a criterion for the concentration of spherical dipoles (which is required for the emergence of spontaneous magnetization): $n\rho_{min}^3 > 0.183$; this criterion differs significantly from that obtained in [4] ($n\rho_{min}^3 > 0.300$). In order to ensure a ferromagnetic ordering at $T > 0$, the truncation of the distribution function $F(1; H_3)$ upon an increase in temperature must be more radical; at a constant concentration of dipoles, this can be ensured only by increasing the minimal dipole spacing. (Parameter ρ_{min} is bounded from above by the natural "geometrical" condition $n\rho_{min} \lesssim 1$.) However, at a high temperature ($\Theta \gtrsim 4.2$), the system remains paramagnetic even for $\rho_{min} \longrightarrow \infty$. The corresponding temperature dependences of spontaneous magnetization and threshold values of $n\rho_{min}^3$ and H_{max} are shown in Figs. 1 and 2, respectively.

The magnetization of the system in question in the paramagnetic state is determined by the external magnetic field H_e. In order to find the magnetization, it is sufficient to carry out the substitution $4\pi I/3 \longrightarrow 4\pi I/3 + H_e$ in Eq. (9), after which the argument of the hyperbolic tangent in Eq. (10) is supplemented with the term $(H_e/h_0)\Theta^{-1}$. The solution of the equation obtained in this way describes the temperature dependence of the system magnetization in the applied magnetic field. Figure 3 shows a series of such dependences for various magnetic fields. As expected, these dependences differ considerably from the Langevin dependence, according to which magnetization $I = I_s$ under saturation (i.e., at low temperatures), which means that $j = 1$. It can be

seen from Fig. 3 that the saturation of the magnetization of a random system always corresponds to $j = j_0 < 1$.[6] The field dependence of the low-temperature magnetization $j_0(H_e)$ is shown in the inset to Fig. 3.

3. NUMERICAL EXPERIMENT: COMPARISON WITH THEORY

The magnetic "behavior" of random systems of point Ising dipoles is obviously determined to a considerable extent by the distribution function $F(1; H_3)$ of local magnetic fields. In this connection, it is expedient to obtain an additional proof of the correctness of the theoretical scheme used for determining this function. For this reason, in addition to the derived analytic expression (8) for function $F(1; H_3)$, we calculated this function by simulating a random system of Ising dipoles.

The system was "created" using a random uniform arrangement of dipoles in a sphere (the total number of dipoles is 5×10^3 or 4×10^4), after which the magnetic field H_3 was calculated at the center of this sphere.

Function $F(1; H_3)$ was found by sorting out a large number (approximately 10^4) of realizations of such a system. The distribution function determined in this way is shown in Fig. 4. It can be seen that the shape of this function is close to the Lorentzian (solid curve), especially on the wings, where the function remains Lorentzian up to very high values of field $H_3 \approx 100mn$ (see the inset to Fig. 4). At the same time, an attempt to approximate function $F(1; H_3)$ by a Gaussian function proved futile (dashed curve in Fig. 4). Since strong fields are mainly created by nearest neighbors, the shape of the distribution function in this region is not sensitive to the choice of the total number of dipoles taken into account in calculations. This, however, does not apply to the range of weak magnetic fields ($H \lesssim h_0$) produced by a more or less symmetric aggregate of a large number of neighboring dipoles. Verification (involving an increase in the number of dipoles taken into consideration) confirms the correctness of the results depicted in Fig. 4.

4. CONCLUSIONS

We have proved that the system studied here can be in the ferromagnetic state in the case of appropriate truncation of the region of strong random fields (which is equivalent to a transition from point dipoles to finite-size dipoles). To find out whether this is the ground state, note that the energy density of the system at $T = 0$

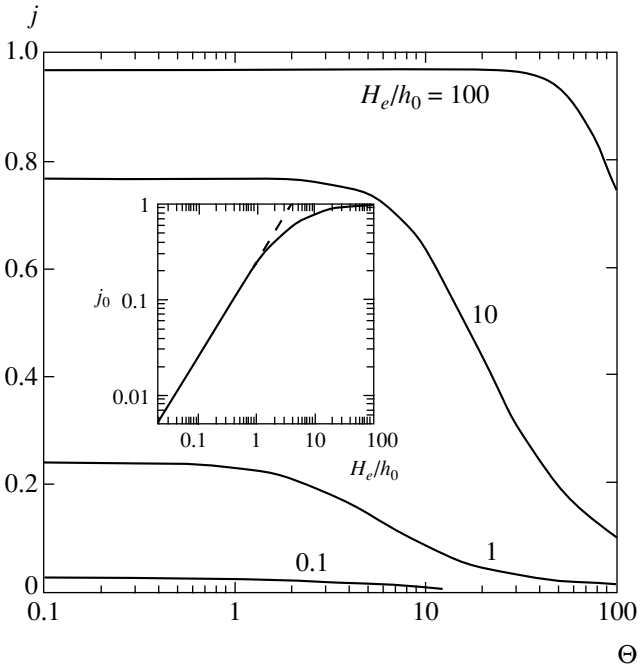

Fig. 3. Temperature dependences of magnetization j of a system in the paramagnetic state for different values of the external magnetic field H_e. The inset shows the field dependence of magnetization at $T = 0$ (the dashed straight line depicts the linear dependence $j_0 = 0.265H_e$).

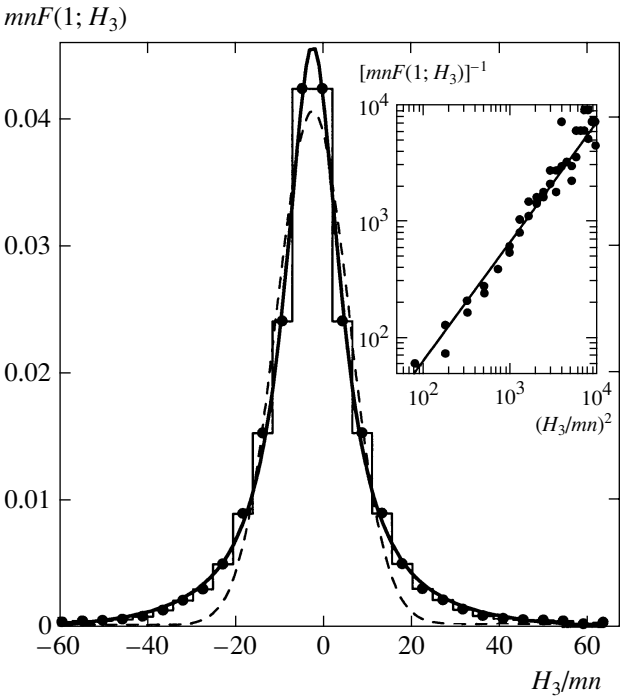

Fig. 4. Distribution function $F(1; H_3)$ for a random system of point Ising dipoles, approximated by the Lorentzian (solid) and the Gaussian (dashed) curves. The inset gives a representation of function $F(1; H_3)$ demonstrating the correctness of the Lorentzian approximation in the range of strong fields.

[6] For $\Theta \longrightarrow 0$, Eq. (10) assumes the form $j = (2/\pi)\arctan(4\pi j/3A + H_e/Ah_0)$; for $j \ll 1$, we have $j_0 = (2/\pi)(A - 8/3)^{-1}H_e = 0.265H_e$.

is $W = n\langle w\rangle$, where the mean energy of a dipole is given by

$$\langle w\rangle = -\langle \mathbf{m}\cdot\mathbf{H}_0\rangle = [-\eta m + (1-\eta)m]\left\langle \frac{4\pi}{3}I + H_3\right\rangle$$

(averaging is carried out over the ensemble of particles), and $I = mn(1-2\eta)$, η being the fraction of dipoles with magnetic moments directed along the magnetization. Since the distribution function $F(1; H_3)$ is even, we obtain

$$\left\langle \frac{4\pi}{3}I + H_3\right\rangle = \int_{-\infty}^{\infty}\left(\frac{4\pi}{3}I + H_3\right)F(1; H_3)dH_3$$
$$= \frac{4\pi}{3}I = \frac{4\pi}{3}mn(1-2\eta)$$

and finally

$$W = -m^2n^2(2\eta-1)^2 \le 0.$$

The paramagnetic state ($\eta = 1/2$) corresponds to $W = 0$; in the ferromagnetic state, we have $W < 0$; consequently, the ferromagnetic phase predominates at low temperatures.

It should be emphasized that all the results obtained in this study pertain exclusively to a "liquid" random system, in which magnetic dipoles can be arranged in any spatial region. For this reason, it would be incorrect to apply these results for describing the properties of "crystalline" random systems, in which dipoles can occupy (with a certain probability) only quite definite positions in a certain crystalline structure. The method applied here is unsuitable for studying such systems since these are not random systems in the Markovian sense. In this case, the initial crystal should be divided into nonequivalent sublattices and the distribution functions for random fields in each sublattice should be calculated numerically.

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 03-02-17029, 02-02-16974, and 01-02-16420).

REFERENCES

1. H. Zhang and M. Widom, Phys. Rev. B **51**, 8951 (1995).
2. J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999; Inostrannaya Literatura, Moscow, 1965).
3. B. E. Vugmeister and M. D. Glinchuk, Rev. Mod. Phys. **62**, 993 (1990).
4. S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 593–600.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 656–663.

SOLIDS
Structure

Magnetic Properties of a Random System of Rodlike Ising Dipoles

E. Z. Meilikhov* and R. M. Farzetdinova

Russian Research Centre Kurchatov Institute, pl. Kurchatova 1, Moscow, 123182 Russia

**e-mail: meilikhov@imp.kiae.ru*

Received December 30, 2002

Abstract—A generalized mean field theory is constructed for a system of rodlike Ising magnetic dipoles of a finite length a with configuration disorder. The theory is based on an analysis of the local magnetic field distribution function. It is shown that the magnetic state of such a system is determined by the dipole concentration n: the system is paramagnetic for $na^3 \lesssim 5 \times 10^2$, while ferromagnetic ordering exists for $na^3 \gtrsim 5 \times 10^2$. The susceptibility of the system in the paramagnetic state is determined. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The magnetic state of a system of interacting magnetic dipoles is determined by a complex "game" of long-range interaction and anisotropy of such an interaction with thermal and spatial magnetic field fluctuations (the latter are associated with a random arrangement and orientation of dipoles). In the traditional mean field theory, thermal fluctuations are included, but the mean local field is assumed to be the same for all dipoles. Consequently, spatial fluctuations of a local field, which are significant in a random system and prevent the establishment of a magnetic order even at zero temperature, are not taken into account.

To obtain a correct description of the magnetic systems in question, the traditional mean field theory must be generalized appropriately. One possible way for doing this was proposed by Klein *et al.* [1] and then used by Zhang and Widom [2] to analyze a random system of spherical dipoles.[1] They proved that the ground state of such a system is paramagnetic as long as the mean distance between spherical dipoles is larger than their size. The ferromagnetic state is possible only in a system in which the dipole concentration exceeds a certain critical value. The same result was obtained in [4] by a different method.

However, many real systems of this type (e.g., ferroliquids or patterned media, which are promising materials for magnetic recording [5]) consist of essentially nonspherical dipoles. For this reason, for a high concentration of such dipoles (when the mean distance between the dipoles is comparable with their size), the type of their interaction differs considerably from the interaction of point dipoles. We will illustrate this statement for a system of rodlike uniformly magnetized rodlike dipoles with magnetic moments m. The interaction energy for two such dipoles with parallel magnetic moments (the center of one dipole is at the origin and the center of the other dipole is determined by radius vector $\mathbf{R}$) is given by the relation [6]

$$w_d = \frac{m^2}{4a^2}\left(\frac{2}{|\mathbf{R}|} - \frac{1}{|\mathbf{R} + 2a\mathbf{e}_0|} - \frac{1}{|\mathbf{R} - 2a\mathbf{e}_0|}\right), \qquad (1)$$

where $\mathbf{e}_0$ is the unit vector in the direction of the magnetic moment of the dipoles and $2a$ is the length of each dipole. If the angle between vectors $\mathbf{e}_0$ and $\mathbf{R}$ is equal to α, Eq. (1) can be written in the form

$$w_d = \frac{m^2}{4a^3} \times \left[\frac{2}{\rho} - \frac{1}{\sqrt{\rho^2 + 4 + 4\rho\cos\alpha}} - \frac{1}{\sqrt{\rho^2 + 4 - 4\rho\cos\alpha}}\right], \qquad (2)$$

where $\rho = |\mathbf{R}|/a$. For $\rho \gg 1$, relation (2) is transformed into the well-known expression for the energy of interaction between point dipoles: $w_d = m^2(1 - 3\cos^2 a)/|\mathbf{R}|^3$.

The sign of energy w_d determines the type of the ground state of a system of two dipoles: if $w_d < 0$, the magnetic moments of the dipoles are parallel (ferromagnetic state); otherwise ($w_d > 0$), these moments are antiparallel (antiferromagnetic state). The plane containing both dipoles under investigation (for definiteness, the xz plane; the magnetic moments are directed along the z axis) splits into two regions with different

[1] The energy of interaction of uniformly magnetized spherical dipoles (one-domain spherical grains) coincides with that of two equivalent point dipoles [3].

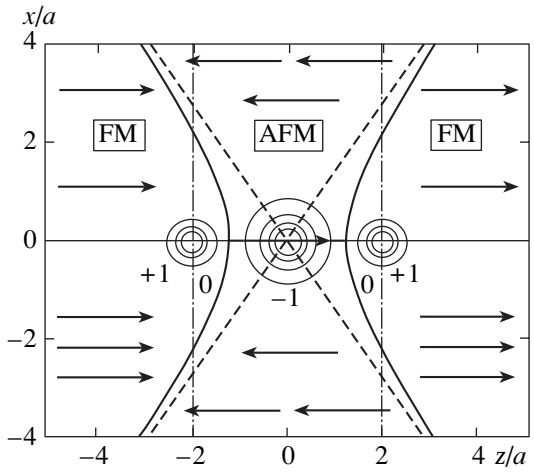

Fig. 1. Boundaries of the regions of ferromagnetic (FM) and antiferromagnetic (AFM) ground states of a system of two Ising dipoles. Dashed lines $x = \pm z\sqrt{2}$ correspond to point dipoles and solid curves are rodlike dipoles of length $2a$. Solid curves correspond to a constant energy of interaction of rodlike dipoles (the digits on the curves indicate energy w_d in units of $m^2/4a^3$).

signs of w_d (and, hence, with different types of the magnetic ground state). Figure 1 shows these regions for point and rodlike dipoles. For point dipoles, the relation between the sizes of these regions is independent of the distance between the dipoles, while the fraction of space corresponding to the antiferromagnetic ground state of a pair of closely spaced rodlike dipoles ($\rho = \sqrt{(x/a)^2 + (z/a)^2} \lesssim 1$) is much larger.

However, this does not lead to a conclusion on a strong tendency of a system of a large number of rodlike dipoles toward antiferromagnetism since the degree of anisotropy of the interaction of rodlike dipoles differs significantly from that for point dipoles.

Thus, nonsphericity of dipoles must considerably affect the magnetic state of a dipole system. This study is devoted to generalization of the mean field theory for a random system of rodlike dipoles.

2. GENERALIZED MEAN FIELD THEORY FOR RODLIKE DIPOLES

Let us suppose that a system consisting of randomly arranged dipoles is in the ferromagnetic state characterized by average magnetization $\mathbf{I} \parallel \mathbf{e}_0$. Local effective fields $-w_d/m$ produced at the origin of coordinates by randomly arranged dipoles are different for different dipoles, and the random value of their sum H_3 can be characterized by the distribution function $F(j; H_3)$, which generally depends on the relative magnetization $j \equiv I/mn$ of the system (n is the dipole concentration) determined by the fraction $\eta = (1/2)(1 + j)$ of dipoles whose average magnetic moments are directed along the magnetization $\mathbf{I}$ of the system ($F(1; H_0)$ is the distribution function in the case when all the dipoles in the Lorentz sphere are parallel to one another).

In order to determine the form of this function, we apply the Markov method [7], according to which

$$F(j; H_3) = \frac{1}{2\pi}\int_{-\infty}^{\infty} A(q)\exp(-iqH_3)dq,$$

$$A(q) = \lim_{N\to\infty}\left[\iiint_{\zeta\,\alpha\,\rho=0}^{\rho_{max}} \exp[iqh_\zeta(\rho, \alpha, \zeta)]\tau(\rho, \alpha, \zeta)d\alpha d\rho d\zeta\right]^N, \quad (3)$$

where (see Eq. (2))

$$h_\zeta(\rho, \alpha, \zeta) = \zeta h(\rho, \alpha),$$

$$h(\rho, \alpha) = \frac{m}{4a^3}\left[-\frac{2}{\rho} + \frac{1}{\sqrt{\rho^2 + 4 + 4\rho\cos\alpha}} + \frac{1}{\sqrt{\rho^2 + 4 - 4\rho\cos\alpha}}\right] \quad (4)$$

is the effective magnetic field produced by a point dipole with random coordinates ρ, α at the origin (in the case considered here, this field is equal to the component of the magnetic field of the dipole along its direction); the random parameter ζ assumes values +1 and –1 (with a probability η and $1 - \eta$, respectively) and determines the direction of the magnetic moment of the dipole, $\tau(\rho, \alpha, \zeta)$ is the distribution function for random values of coordinates and parameter ζ, and N is the number of dipoles in a sphere of radius ρ_{max}, over whose volume the integration is carried out. If we assume further that (i) the random nature of the arrangement of dipoles does not change their average concentration n (i.e., $N \longrightarrow 4\pi(a\rho_{max})^3 n/3$ for $\rho_{max} \longrightarrow \infty$); (ii) the distributions of coordinates and parameter ζ are uniform; and (iii) correlations are absent, then

$$\tau(\rho, \alpha, \zeta)d\rho d\alpha d\zeta = [(1-\eta)\delta(\zeta + 1) + \eta\delta(\zeta - 1)]d\zeta \cdot (4\pi\rho_{max}^3/3)^{-1} \times \rho^2 d\rho \cdot 2\pi\sin\alpha d\alpha. \quad (5)$$

Substituting Eqs. (4) and (5) into (3), we obtain

$$A(q) = \exp[-na^3C(q)],$$

$$C(q) = 2\pi(1-\eta)\int_0^\infty \rho^2 d\rho \times \int_0^\pi \{1-\exp[-iqh(\rho,\alpha)]\}\sin\alpha d\alpha + 2\pi\eta\int_0^\infty \rho^2 d\rho \int_0^\pi \{1-\exp[iqh(\rho,\alpha)]\}\sin\alpha d\alpha \quad (6)$$

$$= 2\pi\int_0^\infty \rho^2 d\rho \times \int_0^\pi \{1-\cos[qh(\rho,\alpha)] - ij\sin[qh(\rho,\alpha)]\}\sin\alpha d\alpha.$$

In particular, it follows from Eqs. (3) and (6) that in the absence of magnetization ($j = 0$), the distribution function $F(0; H_3)$ of local fields is even, i.e., symmetric with respect to $H_3 = 0$.

Relations (6) do not lead to a simple analytic expression for the distribution function $F(j; H_3)$. For this reason, we determined this function using two mutually complementing methods: (i) the "small q approximation" based on the fact that the region of large values of q is insignificant in the inverse Fourier transformation of (3) [2] and (ii) numerical calculations for a model random system of rodlike Ising dipoles.

In the first approach, functions $\cos[qh(\rho, \alpha)]$ and $\sin[qh(\rho, \alpha)]$, which are to be integrated in Eq. (6), are replaced by their approximate power expansions in the small argument qh (up to the first nonvanishing term in qh). After this, it becomes possible to carry out integration, leading to

$$\int_0^\infty \rho^2 d\rho \int [1-\cos[qh(\rho,\alpha)]]\sin\alpha d\alpha = \frac{\pi^2}{4}\left(\frac{mq}{a^3}\right)^2, \quad (7)$$

$$\int_0^\infty \rho^2 d\rho \int \sin[qh(\rho,\alpha)]\sin\alpha d\alpha = -\frac{2}{3}\frac{mq}{a^3}.$$

Substituting Eqs. (7) into (6), we obtain

$$F(j; H_3) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left[-\frac{(H_3 + jH_j)^2}{2\sigma^2}\right],$$

$$\frac{H_j}{h_0} = \frac{4\pi}{3} \approx 4.2 \quad (8)$$

$$\frac{\sigma}{h_0} = \frac{\sqrt{\pi}}{2}\frac{1}{\sqrt{na^3}} \approx \frac{0.89}{\sqrt{na^3}}, \quad h_0 = mn.$$

Thus, in the small q approximation, distribution functions $F(j; H_3)$ have the form of Gaussian functions whose maximum ($H_3 = -jH_j$) is shifted linearly upon an increase of magnetization to the region of negative values of magnetic field H_3.[2] The approximation used here is valid if the probability of the emergence of strong local fields for which $qh \gtrsim 1$ is low. The Gaussian distribution function obtained above (with exponentially decreasing "wings") does not possess such a property. However, as the concentration of rodlike dipoles decreases, their finite size becomes less and less significant, and such dipoles can be treated as point dipoles for $na^3 \ll 1$. In this case, the system is known to be characterized by a Lorentzian distribution function [4] with "long" wings decreasing according to a power law. Consequently, distribution (8) holds only in the concentration region $na^3 \gtrsim 1$. On the other hand, strong fields are mainly created by the nearest neighbors [7], and the probability of their production increases with the dipole concentration n, when the mean distance between dipoles becomes smaller than the dipole size a. Consequently, the approximation used becomes inapplicable for $na^3 \gg 1$. In order to determine the limits of its applicability more precisely, we numerically calculated the local field distribution functions in a model random system of rodlike Ising dipoles.

The system was "created" via a uniform random arrangement of dipoles over a sphere (the total number of dipoles is 10^4), after which the magnetic field H_3 at the center of this sphere was calculated. Functions $F(j; H_3)$ were determined by running through a large number (about 10^4) of realizations of such a system. Figure 2 shows the distribution functions $F(j; H_3)$ determined as a result of such a numerical calculation for a system with $na^3 = 1$. It can be seen that these functions are close to Gaussian functions, their widths and the positions of peaks being close to those predicted by relations (8). Figure 2 also shows that the shape of the distribution functions in the wings differs considerably

[2] Formally, this is connected with the specific form of the angular dependence (4) of the rodlike dipole field. As was mentioned earlier, there is an increase (as compared to point dipoles) in the fraction of space corresponding to the antiferromagnetic interaction of a pair of closely spaced rodlike dipoles ($\rho \lesssim 1$).

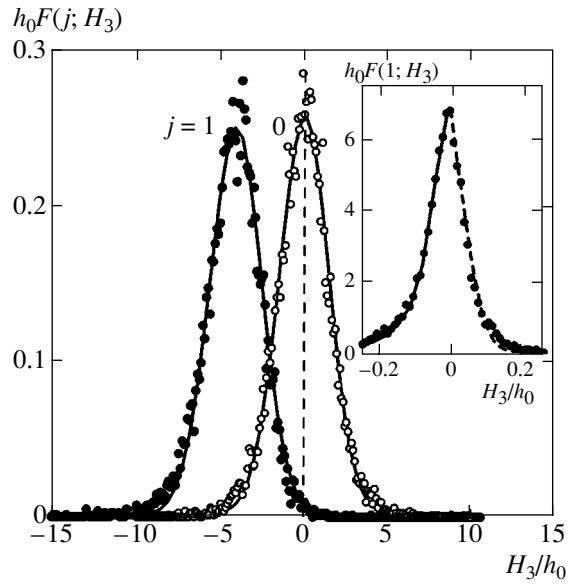

Fig. 2. Distribution functions $F(0; H_3)$ and $F(1; H_3)$ of local fields for a model system of randomly distributed rodlike Ising dipoles with $na^3 = 1$. Solid curves correspond to approximated Gaussian functions. The discrepancies are most noticeable on the left wings of the distributions. The inset shows the distribution function $F(1; H_3)$ for a system with $na^3 = 0.01$ close to a system of point dipoles (the solid curve corresponds to the Lorentzian approximation, and the dashed curve, to the Gaussian function).

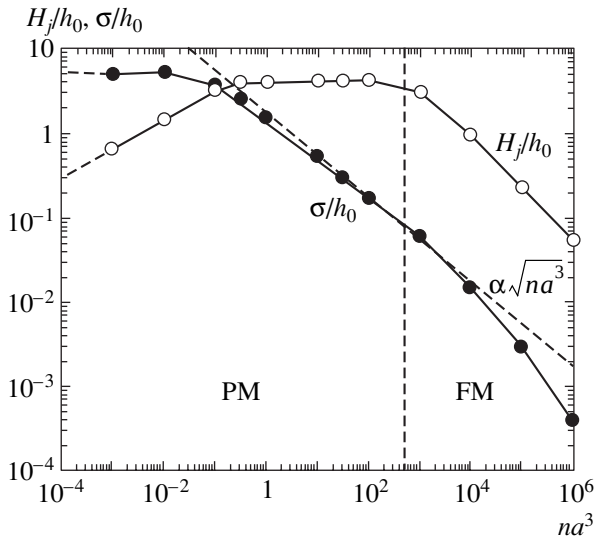

Fig. 3. Concentration dependences of calculation parameters (position of the H_j peak and width σ) of distribution $F(1; H_3)$ for a model system. The inclined dashed line corresponds to the $\sqrt{na^3}$ power dependence, while the vertical dashed line is the left boundary between the ferromagnetic (FM) and paramagnetic (PM) phases.

from the Gaussian shape especially for fields lying to the left of the peak). These region correspond to strong (in absolute magnitude) fields which cannot in principle be described correctly by the approximation of small q. The shape of the distribution function $F(1; H)$ in this region will be determined in the Appendix.

As the dipole concentration decreases to values $na^3 \ll 1$ and the system becomes equivalent to a random system of point dipoles, the shape of the distribution indeed becomes close to Lorentzian shape. This can be seen from the inset to Fig. 2, corresponding to a system with $na^3 = 0.01$: the right wing of the distribution function is close to the Gaussian distribution, while its left wing is already Lorentzian.

Figure 3 shows the dependences of parameters H_j and σ of the distribution function $F(1; H_3)$ on the dipole concentration determined by the value of na^3. It can be seen that the functional dependences $H_j/h_0 = \text{const}$ and $\sigma/h_0 \propto (na^3)^{-1/2}$, which are predicted by the small q approximation, are observed in the concentration range $1 \lesssim na^3 \lesssim 10^3$, i.e., for systems in which the mean distance between dipoles is 1–10 times smaller than their size.

3. MAGNETIC PROPERTIES OF A RANDOM SYSTEM OF RODLIKE DIPOLES

The considerable shift of the distribution function $F(1; H_3)$ to the range of negative magnetic fields for $na^3 \gtrsim 1$ "prevents" the formation of magnetic ordering in the system considered here. The system magnetization is $\mathbf{I} = n\langle \mathbf{m}_i \rangle_T$, where the mean magnetic moment $\langle \mathbf{m}_i \rangle_T$ must be calculated taking into account the spread of local fields H_3 via the obvious generalization of the equation

$$\frac{I}{mn} = \tanh\left[\frac{m(4\pi I/3 + H_3)}{kT}\right],$$

corresponding to a regular system (here, the term $4\pi I/3$ corresponds to the magnetic field produced by dipoles on the surface of the Lorentz sphere):

$$\frac{I}{mn} = \int_{-\infty}^{\infty} \tanh\left[\frac{m(4\pi I/3 + H_3)}{kT}\right] F(j; H_3)\, dH_3. \quad (9)$$

Obviously, Eq. (9) may have the solution $j = 1$ at $T = 0$ only under the necessary condition $F(1; H_3 < 4\pi I/3) = 0$, which can never be realized in a system of dipoles arranged absolutely at random. As in the case of point dipoles [4], the emergence of ferromagnetism can be "facilitated" by limiting the configuration randomness (e.g., by appropriately setting the lower limit on dipole spacing, which is equivalent to "trimming" the wings of the random field distribution function).

Using expression (8) for the distribution function $F(j; H_3)$, we arrive at an equation generalizing the standard equation for mean field:

$$j = \frac{1}{\sqrt{2\pi}(\sigma/h_0)}\int_{-\infty}^{\infty}\tanh\left[\frac{1}{\Theta}\left(\frac{4\pi}{3}j + u\right)\right] \times \exp\left[-\frac{(u + jH_j/h_0)^2}{2(\sigma/h_0)^2}\right]du, \quad (10)$$

where $\Theta = kT/m^2n$ is the reduced temperature (equal to the ratio of the thermal energy to the characteristic energy of magnetic interaction between dipoles).

The paramagnetic state ($j = 0$) is obviously a solution to this equation since its right-hand side is equal to zero for $j = 0$. In order to find out whether this equation has a solution corresponding to the ferromagnetic state ($j \neq 0$), we note that this equation for $\Theta \longrightarrow 0$ assumes the form

$$j = -\mathrm{sgn}(z_0)\Phi(|z_0|j), \quad (11)$$

where

$$\Phi(x) = \frac{2}{\sqrt{\pi}}\int_0^x \exp(-x^2)dx$$

is the probability integral and $z_0 = (H_j/h_0 - 4\pi/3)(h_0/\sqrt{2}\,\sigma)$. Equation (11) has a solution $j \neq 0$ only in the case when $z_0 < 0$, $|z_0| > \sqrt{\pi}/2$, i.e., for

$$\frac{4\pi}{3}mn > H_j + \sigma\sqrt{\frac{\pi}{2}}. \quad (12)$$

The meaning of the last condition is that the collective field $(4\pi/3)mn$ created by dipoles on the surface of the Lorentz sphere must "overcome" local fields of the reverse sign with the mean value $-H_j$ and a spread $\pm\sigma$. It can be seen from Fig. 3 that this condition holds only for $na^3 \gtrsim 5 \times 10^2$, i.e., for a fairly high dipole concentration. In this case, the system is ferromagnetic and Eq. (10) can be used for deriving the temperature dependence of its magnetization, which is shown in Fig. 4. The temperature range in which the magnetization differs from zero expands upon an increase in the dipole concentration na^3. The corresponding concentration dependence of the Curie temperature is given in Fig. 5. It can be seen from Fig. 3 that $H_j, \sigma \longrightarrow 0$ for $na^3 \longrightarrow \infty$. This means that the distribution function $F(j; H_3)$ approaches a delta function centered in the vicinity of $H_3 = 0$. In this case, Eq. (10) can be reduced to the equation $j = \tanh(4\pi j/3\Theta)$; for $\Theta \approx \Theta_C$ (for $j \ll 1$), this equation shows that $\Theta_C = 4\pi/3$ for high dipole concentrations.

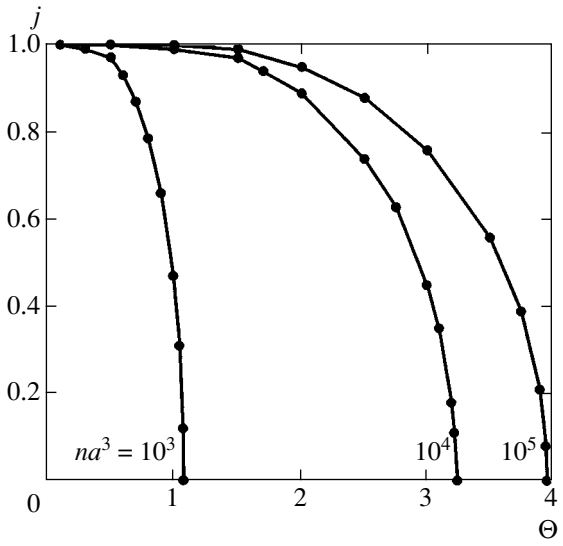

Fig. 4. Temperature dependences of the system magnetization in the ferromagnetic state for different dipole concentrations.

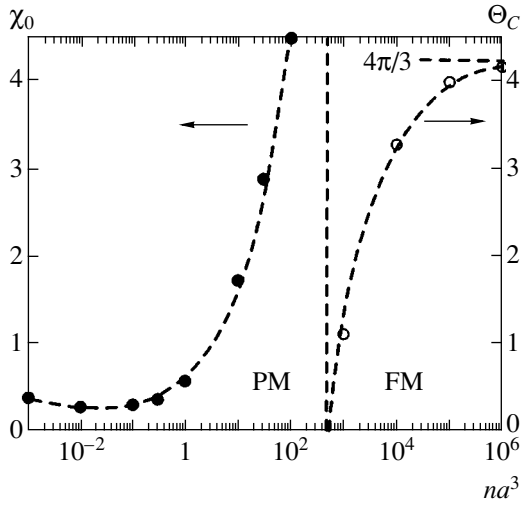

Fig. 5. Concentration dependences of the low-temperature susceptibility of the paramagnetic phase and the Curie temperature of the ferromagnetic phase of a random system of rodlike Ising dipoles. As the dipole concentration increases ($na^3 \longrightarrow \infty$), the Curie temperature attains saturation ($\Theta_C \longrightarrow 4\pi/3$). The vertical dashed line is the left boundary of the ferromagnetism region.

In the limit of low dipole concentrations ($na^3 \ll 1$), the system under study resembles a random system of point dipoles without ferromagnetic ordering [4].[3]

[3] We are speaking of a system in which dipoles can be arranged at infinitely small distances relative to one another and the distribution function is Lorentzian. In such a system, strong random fields corresponding to long wings characteristic of such functions suppress ferromagnetism.

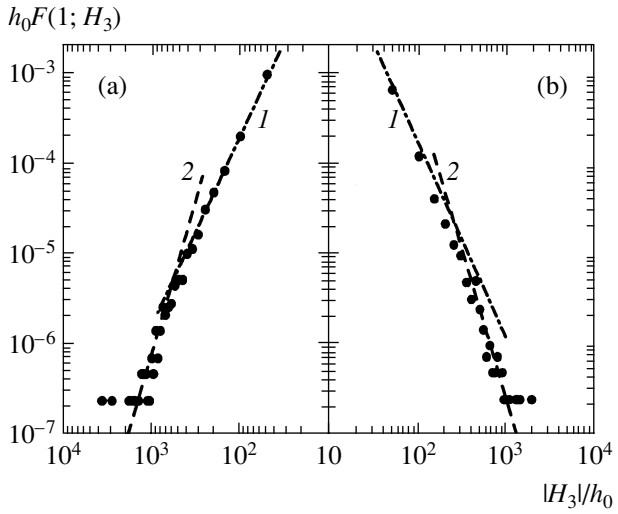

Fig. 6. Left (a) and right (b) wings of the distribution function $F(1; H_3)$ of local fields for rodlike dipoles with a concentration corresponding to $na^3 = 10^{-3}$: Lorentzian function (*1*) and power dependences $1/H^4$ (*2*).

Nevertheless, complete equivalence between the systems of linear and point dipoles is attained only in the limit of their vanishingly low concentration. As a matter of fact (see the Appendix), the distribution function in the region of strong fields for any finite dipole concentration decreases much more rapidly than the Lorentz function, namely, according to the law $1/H^4$ (see Eq. (20)) (naturally, the low-field boundary of this region depends on the dipole concentration and is shifted towards stronger fields upon a decrease in the concentration).

One of the model distribution functions corresponding to the dipole concentration $na^3 = 10^{-3}$ is shown in Fig. 6, demonstrating the wings of this function rapidly decreasing according to the $1/H^4$ law corresponding to formula (20).

In the range of low concentrations ($na^3 \lesssim 5 \times 10^2$), the system is paramagnetic. In order to describe the properties of this system in an external magnetic field H_e, it is sufficient to carry out the substitution $H_3 \longrightarrow H_3 + H_e$ in the argument of the hyperbolic tangent in Eq. (9). This will lead to the substitution $u \longrightarrow u + H_e/h_0$ in Eq. (10). In a weak external field ($H_e/h_0 \ll 1$), the magnetization of the paramagnetic system is low ($j \ll 1$). Expanding the functions appearing in Eq. (10) in H_e and j, we obtain

$$j = I_1\left(\frac{4\pi}{3}j + \frac{H_e}{h_0}\right) - I_2\frac{H_j}{h_0}j, \qquad (13)$$

where

$$I_1 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{e^{-x^2/2}}{\Theta\cosh^2[(\sigma/h_0\Theta)x]}dx,$$
$$I_2 = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}x\tanh\left[\frac{\sigma}{h_0\Theta}x\right]e^{-x^2/2}dx. \qquad (14)$$

Relation (13) makes it possible to find the low-field magnetic susceptibility of the system $\chi = I/H_e = j/(H_e/h_0)$ and its temperature dependence.

At high temperatures ($\Theta \gg 1$), we have $I_1 = 1/\Theta \longrightarrow 0$, $I_2 \longrightarrow 0$ and it follows from Eq. (13) that $\chi(\Theta \gg 1) = 1/\Theta$, or $I = mn(mH_e/kT)$ is the conventional Curie law for noninteracting Ising dipoles.

At low temperatures ($\Theta \ll 1$), we have

$$I_1 = I_2 = \sqrt{\frac{2}{\pi}}\frac{h_0}{\sigma},$$

which gives

$$\chi_0 \equiv \chi(\Theta \ll 1) = \left(\sqrt{\frac{\pi}{2}}\frac{\sigma}{h_0} + \frac{H_j}{h_0} - \frac{4\pi}{3}\right)^{-1},$$

i.e., a temperature-independent susceptibility. In the vicinity of the ferromagnetic transition (whose boundary is determined by condition (12)), it increases indefinitely. The concentration dependence of χ_0 is shown in Fig. 5.

4. CONCLUSIONS

As applied to random systems, the mean field theory for a system with magnetic interaction requires a generalization taking into account nonequivalence of individual magnetic moments forming the system. We have considered a version of such a theory, based on the determination and analysis of the distribution functions for random magnetic fields produced by magnetic moments with an irregular spatial distribution. These distribution functions have been determined by two methods: analytically, with the help of the Markov method [7], and numerically, by processing statistically the results of calculations of random fields in a model system. In the case considered here, the former method makes it possible to derive only approximate equations for distribution functions in a limited range of dipole concentrations; however, these expressions give a reliable qualitative idea on the form of these functions in the magnetic state of the system under study.

The main results were obtained by numerically simulating a random system of rodlike Ising dipoles. It has been shown in the framework of the generalized mean field theory that the magnetic state of such a system is determined by the dipole concentration n: for dipole

concentrations exceeding $na^3 \gtrsim 5 \times 10^2$, the system is ferromagnetic, while for lower concentrations, it is paramagnetic. In the region of low concentrations ($na^3 \ll 1$), the system is close to an ensemble of point dipoles but is not equivalent to such an ensemble since the distribution function in the region of strong random fields is characterized not by long Lorentzian wings, but by a more rapid decay (the $1/H^4$ law).

ACKNOWLEDGMENTS

This study was supported by the Russian Foundation for Basic Research (project nos. 03-02-17029, 02-02-16974, and 01-02-16420).

APPENDIX

We can get an idea of the form of the distribution function $F(1; H)$ in the region of strong fields by assuming that field H is mainly produced by the nearest dipole. The $w_R(R)$ law describing the distribution of distances R to the nearest neighbor for a random (noncorrelated) distribution of dipoles is given by [7]

$$w_R(R) = 4\pi R^2 n \exp(-4\pi R^3 n/3), \quad \text{(A.1)}$$

while the distribution $w_\alpha(\alpha)$ of angles α between vector $\mathbf{e}_0$ and vector $\mathbf{R}$ determining the position of the nearest dipole will be regarded as uniform: $w_\alpha(\alpha) = |\sin\alpha| d\alpha$.

In contrast to the energy of interaction of point dipoles, which increases indefinitely upon a decrease in their spacing, the energy of interaction of rodlike dipoles has three (infinitely large) extrema (see Fig. 1) corresponding to the convergence of their centers ($R \ll a$) or ends ($R - 2a \ll a$, $\sin\alpha \longrightarrow 0$). It follows from Eq. (2) that

$$H \approx -\frac{a}{2R} H_1 \quad \text{(A.2a)}$$

in the vicinity of these extrema for $R \ll a$, while

$$H \approx \frac{a}{4\sqrt{(R-2a)^2 + 4a^2\alpha'^2}} H_1 \quad \text{(A.2b)}$$

for $R - 2a \ll a$, $\alpha' = \alpha$ or $\alpha' = \pi - \alpha$, where $H_1 = m/a^3$. The three indicated regions determine the shape of the distribution function $f(H)$ for large absolute values of H. For the function corresponding to negative values of the magnetic field (A.2a), we obtain

$$F(1; H) \propto w_R[R(H)]\left|\frac{dR(H)}{dH}\right| = \frac{\pi}{2}\frac{na^3}{H_1}\left(\frac{H_1}{H}\right)^4 \times \exp\left[-\frac{\pi}{6} na^3 \left(\frac{H_1}{|H|}\right)^3\right], \quad H < 0. \quad \text{(A.3)}$$

For regions with positive magnetic field (A.2b), we can use the following rule of transformation of random quantities [8]: if the distribution functions $f_r(r)$ and $f_\chi(\chi)$ for independent random quantities r and χ are known, the distribution function for the random quantity $\Lambda = \sqrt{r^2 + \chi^2}$ is given by

$$f_\Lambda(\Lambda) = \Lambda \int_0^{2\pi} f_r(\Lambda\cos\varphi) f_\chi(\Lambda\sin\varphi) d\varphi.$$

It follows hence that the distribution function $f_h(h)$ for random quantity $h = 1/\Lambda = 1/\sqrt{r^2 + \chi^2}$ is given by

$$f_h(h) = f_\Lambda[\Lambda(h)]\left|\frac{d\Lambda}{dh}\right| = \frac{1}{h^3}\int_0^{2\pi} f_r\left(\frac{\cos\varphi}{h}\right) f_\chi\left(\frac{\sin\varphi}{h}\right) d\varphi.$$

In our case, $r = R - 2a$, $\chi = 2a\alpha'$, $H = (m/4a^2)h$ and $f_r(r) = w_R(2a + r)$, $f_\chi(\chi) = (1/2a) f_\alpha(\chi/2a)$. Consequently,

$$F(1; H) \propto 2\frac{4a^2}{m} f_h\left(\frac{H}{m/4a^2}\right) = \frac{\pi n a^3}{H_1}\left(\frac{H_1}{H}\right)^3 \int_0^{2\pi}\left[1 + \frac{1}{8}\frac{H_1}{H}\cos\varphi\right]^2 \times \exp\left\{-\frac{32\pi n a^3}{3}\left[1 + \frac{1}{8}\frac{H_1}{H}\cos\varphi\right]^3\right\}\left|\sin\left[\frac{1}{8}\frac{H_1}{H}\sin\varphi\right]\right| d\varphi \quad \text{(A.4)}$$

(factor 2 on the right-hand side of the first relation in (A.4) takes into account the existence of two extrema of the positive magnetic field). For fields $H/H_1 \gg \sqrt{na^3}$, distribution function (A.4) has a form similar to Eq. (A.3):

$$F(1; H) \propto \frac{\pi}{2}\frac{na^3}{H_1}\left(\frac{H_1}{H}\right)^4 \exp\left(-\frac{32\pi n a^3}{3}\right) \times \left\{\frac{\sinh[4\pi n a^3 (H_1/H)]}{4\pi n a^3 (H_1/H)}\right\}, \quad H > 0. \quad \text{(A.5)}$$

It follows from Eqs. (A.3) and (A.5) that the left wing of the distribution function $F(1; H)$ must be higher than its right-hand wing. This can clearly be seen in the inset to Fig. 2. For a low dipole concentration ($na^3 \ll 1$),

both wings of the distribution function have, in accordance with Eqs. (A.3) and (A.5), identical power form

$$F(1; H) \propto \frac{na^3}{H_1}\left(\frac{H_1}{H}\right)^4, \qquad \text{(A.6)}$$

which considerably differs, however, from the Lorentzian profile.[4]

As regards the region of weak magnetic fields ($H < H_1$), such fields are produced by a more or less symmetric set of all neighboring dipoles and it is impossible to determine the form of the corresponding distribution function $f(H)$ using the nearest neighbors approximation. In this case, an appropriate statistical method (e.g., the Markov method) has to be used.

[4] In contrast to Eq. (A.6), the wings of the Lorentzian distribution decrease at a much slower rate (according to the law $1/H^2$), i.e., so slowly that the root-mean-square value $\sqrt{\langle H^2 \rangle}$ of the random field is infinitely large. It is for this reason that ferromagnetism is impossible in the system of point dipoles; consequently, the wings of the Lorentzian distribution corresponding to $|H| \gtrsim 5h_0$ must be trimmed for the emergence of spontaneous magnetization [4].

REFERENCES

1. M. W. Klein, C. Held, and E. Zuroff, Phys. Rev. B **13**, 3576 (1976).
2. H. Zhang and M. Widom, Phys. Rev. B **51**, 8951 (1995).
3. J. D. Jackson, *Classical Electrodynamics*, 3rd ed. (Wiley, New York, 1999; Inostrannaya Literatura, Moscow, 1965).
4. E. Z. Meilikhov, Zh. Éksp. Teor. Fiz. **124**, 650 (2003) [JETP **97**, 587 (2003)].
5. R. L. White, J. Magn. Magn. Mater. **209**, 1 (2000).
6. M. D. Costa and Yu. G. Pogorelov, Phys. Status Solidi A **189**, 923 (2001).
7. S. Chandrasekhar, Rev. Mod. Phys. **15**, 1 (1943).
8. V. S. Pugachev, *Theory of Probabilities and Mathematical Statistics* (Fizmatlit, Moscow, 2002).

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 601–605.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 664–669.

SOLIDS
Structure

Effect of a Magnetic Field on the Starting Stress and Mobility of Individual Dislocations in Silicon

M. V. Badylevich, Yu. L. Iunin, V. V. Kveder, V. I. Orlov*, and Yu. A. Osip'yan

Institute of Solid-State Physics, Russian Academy of Sciences,
Chernogolovka, Moscow oblast, 142432 Russia

**e-mail: orlov@issp.ac.ru*

Received February 18, 2003

Abstract—The strong effect of a magnetic field on the starting stress and mobility of individual dislocations is discovered in silicon grown by the Czochralski method with a high concentration of dissolved oxygen. It is shown that exposure of dislocations preliminarily introduced into the sample to a magnetic field considerably reduces the starting stresses for the motion of these dislocations. The effect is not observed in samples with a low oxygen concentration. It is assumed that the magnetic field induces singlet–triplet transitions in thermally excited states of silicon–oxygen complexes in the dislocation core, thus stimulating a change in the state (atomic configuration) of oxygen already located at dislocations. As a result, the mean binding energy of oxygen with a dislocation decreases. *© 2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Analysis of the interaction between dislocations in silicon and intrinsic and impurity point defects is essential for the development of modern electronic technologies involving, in 90% of cases, silicon crystals grown by the Czochralski method. In integrated circuit technologies, such studies are important, first, due to the wide use of the dislocations for gettering harmful impurities and, second, in view of the dislocation pinning required for preventing their multiplication and penetration into the active regions of integrated circuits during technological operations involved in production. In the past few years, the interest in studies of the interaction between dislocations and impurities has grown still further in connection with the use of cheap polycrystalline silicon for manufacturing solar cells. Dislocations strongly affect the lifetime of minority carriers in this material and determine the efficiency of solar cells to a considerable extent. It is well known that "pure" and defect-free 60° dislocations in silicon determine the existence of comparatively shallow (about 70–80 meV) one-dimensional bands that weakly affect the lifetime of minority carriers; consequently, dislocations are effective recombination centers only due to the presence of some specific defects and impurities in their cores [1]. For this reason, it is important to study the interaction and reactions between impurities and dislocations, as well as the effect of impurities on the properties of dislocations, and to find methods for controlling these properties.

In some cases, the interaction between impurities and dislocations has the form of solid-state chemical reactions (see, for example, [2, 3]); at the intermediate stages of such a reaction, "broken" valence bonds with an unpaired electron spin may appear. Hence, it cannot be ruled out that an external magnetic field mixing triplet and singlet states at intermediate stages of reactions may strongly affect the kinetics and course of such reactions. Some spin-dependent chemical reactions strongly affected by magnetic field were discovered in liquids and gases and studied in detail [4]. Unfortunately, spin-dependent reactions of defects in crystals have been studied insufficiently.

The measurement of dislocation mobility is a sensitive integrated method for detecting various reactions of defects in crystals. The effect of magnetic field on plastic properties of alkali–halide and some other crystals has been established quite reliably [5–7]. This phenomenon is known as the magnetoplastic effect. It is generally assumed that a change in the dislocation mobility as a result of exposure of samples to magnetic field is associated with changes in the properties of intrinsic and impurity defects; however, the microscopic mechanisms of such processes have not been studied in detail.

In view of the considerable practical importance of silicon as a basic material for electronics, analysis of spin-dependent processes in silicon is of prime importance (especially on account of growing interest in spintronics). Only spin-dependent recombination of electrons and holes at defects has been investigated earlier in detail (see, for example, [8]), while reactions between defects and impurities have been studied insufficiently. It should be noted that the effect of magnetic field on some electrical and mechanical properties of silicon samples, which can apparently be attributed to reactions with defects, was observed in some cases [9, 10]. However, the effects described in these publications have not only been insufficiently studied, but also

poorly reproducible. For example, in spite of several attempts, we failed to reproduce the results presented in [9, 10]. The effects described in [9, 10] are determined by certain poorly controllable factors and require additional experiments for clarifying their origin.

In this study, we report on the results of analysis of the effect of a magnetic field on starting stresses and mobility of individual dislocations in silicon single crystals.

2. EXPERIMENTAL TECHNIQUE

The experiments were performed on initially dislocation-free *n*-type silicon single crystals with a phosphorus concentration of $(3–5) \times 10^{13}$ cm^{-3}. We used two types of crystals differing in the concentration of dissolved oxygen, viz., crystals grown using the Czochralski method (Cz–Si) and those containing 10^{18} cm^{-3} of dissolved oxygen and grown using float zone method (Fz–Si) with an oxygen concentration on the order of 10^{16} cm^{-3}.

Samples having a shape of right prisms with a size of $35 \times 4 \times 1$ mm^3 and with (111), $(11\bar{2})$, and $(1\bar{1}0)$ face orientations (the symbols of faces are given in decreasing order of their areas) were cut by a diamond saw and polished mechanically by diamond powder to obtain an optically smooth surface. Then the surface layer, 20–50 μm in thickness, which was damaged during mechanical treatment, was removed by chemical polishing in a $7HNO_3 : 1HF$ solution.

After polishing, dislocations were introduced into the samples by pricking the wide (111) face with an indenter (diamond pyramid). The load applied to the pyramid during indentation, reduced shear stresses τ, duration Δt, and loading temperature T were chosen so that individual dislocation half-loops 10–20 μm in diameter were generated, as a rule, as a result of each indentation. In our experiments, these parameters were as follows: $T = 600$°C, $\tau = 50$ MPa, and $\Delta t = 10$ min. Then the dislocation sources (in the indentation regions) were removed by chemical polishing and the samples were loaded again at $T = 600$°C to increase the size of the initial dislocation half-loops increased to a diameter on the order of 500–600 μm.

After slow cooling of the samples (together with the furnace), the dislocation outcrops at the sample surface were revealed by selective chemical etching (Sirtl etchant: $1CrO_3 + 2H_2O + 3HF$); the samples prepared in this way were used for measuring the velocity of dislocations and starting stresses for their motion.

The mobility of individual dislocations at a fixed stress τ was measured by loading the samples using four-point bending around the $[11\bar{2}]$ axis over time Δt. The track lengths L of individual dislocations over the loading time Δt were determined by repeated selective etching. Length L of the dislocation tracks was measured using a Neophot-32 optical microscope with a micrometer eyepiece having a scale division value of about 0.2 μm/division. In order to measure the dependence of dislocation velocity $V = L/\Delta t$ on stress τ and to determine starting stresses τ_{st} for moving dislocations, we use three-point bending around the $[11\bar{2}]$ axis leading to a linear law of τ distribution along the long edge of the sample.

In order to improve the precision of measurements of dislocation velocity for a fixed load, we determined in each experiment the velocity of several tens of dislocations and determined the average velocity. The deformation temperature T was monitored by a platinum–platinum-rhodium thermocouple and was maintained at a constant level to within ±0.5 K in the course of the experiment.

In order to study the effect of magnetic field, the samples with introduced dislocations were held for a certain time at room temperature in a field of $B = 20$ kOe prior to velocity measurements. The magnetic vector was perpendicular to the (111) face.

The experimental results considered below were obtained for 60° segments of dislocation half-loops of the $(1\bar{1}1)[011]$ slip system, which had been introduced on the compressed, as well as extended, side of the sample.

3. EXPERIMENTAL RESULTS AND DISCUSSION

It was found that the average velocity of individual 60° dislocations in Cz–Si crystals after exposure of the samples to magnetic field at room temperature increases by 30–40% as compared to the average velocity of dislocations in the same sample, measured before exposure to magnetic field. Figure 1 shows the relative variation $(V_B - V_0)/V_0$ of the dislocation velocity under a load of $\tau = 50$ MPa at a temperature of 550°C as a function of the exposure time in field $B = 20$ kOe at room temperature. Here, V_0 and V_B are the average velocities of dislocations measured before and after exposure of the sample to magnetic field. The average velocities of dislocations depicted in Fig. 1 were obtained as a result of averaging the path lengths of 50–100 individual dislocations in each sample. It can be seen that the strongest effect is attained after 3 h of exposure to magnetic field; a further increase in the exposure time does not lead to a considerable change in the effect. For this reason, the samples were treated in magnetic field for three hours in hour experiments.

Within the measurement error, the effect of an increase in the dislocation velocity due to exposure of the samples to magnetic field remains unchanged after holding for 100–150 h at room temperature and decreases by half only after 300–400 h. This means that magnetic field induces changes in the system of defects

that become active immediately after exposure to the field. The effect persists for a long time.

The measurements of the temperature dependence of the velocity of individual dislocations before and after exposure to magnetic field proved that such exposure in the temperature range 500–650°C studied in our experiments does not lead to an appreciable change in the activation energy $U_d = 2.15 \pm 0.01$ eV of the mobility of dislocations.

Similar measurements were made on Fz–Si single crystals. It was found that within the experimental error, exposure of the samples to magnetic field does not change the mobility of individual dislocations in Fz–Si. Since the main difference between Cz–Si and Fz–Si samples is that the oxygen concentration in the former samples is two orders of magnitude higher than in the latter, it is natural to assume that the effect is associated with spin-dependent reactions involving oxygen.

It is well known that a large number of impurities and defects (including oxygen-containing complexes in silicon) may affect the motion and multiplication of dislocations to a considerable extent [11]. A natural question arises whether the observed effect of magnetic field is associated with a change in the state of oxygen in the bulk of the sample or with a change in the state of oxygen accumulated directly at dislocations in the course of sample preparation.

Our measurements of optical absorption spectra at $T = 77$ K in a region of 1100–1150 cm^{-1}, which are associated with vibrational modes of oxygen in silicon, have not revealed any significant difference between the states of oxygen in the bulk of the samples before or after the action of magnetic field. This suggests that the observed effect is associated with the change in the state of oxygen located directly at dislocations, which is stimulated by magnetic field.

The only sensitive method revealing changes in the system of defects and impurities in the vicinity of a dislocation core is the measurement of starting stresses, i.e., the minimal stresses under which dislocations begin to move [12, 13]. It was shown in [12, 13] that the magnitude of starting stresses τ_{st} is directly connected with the interaction of a dislocation core with surrounding impurities and point defects. In the first (rough) approximation, we have $\tau_{st} \sim NU_B$, where N is the number of impurity atoms per unit dislocation length and U_B is the binding energy between the impurity atoms and the dislocation core.

We measured the starting stresses in Cz–Si samples before and after the action of magnetic field. For this we measured the dependences of dislocation path length L on the shearing stress (T = 600°C, $\Delta t = 10$ min) before and after exposure to magnetic field using the three-support bending method. Figure 2 shows typical dependences of $L/\Delta t$ on τ obtained in this way. Each point corresponds to the path length of an individual dislocation, while continuous curves were obtained by averaging (parabolic smoothing) over 13 points (dislocations). It can be seen that exposure of the sample to magnetic field leads to a noticeable (to almost one-half the initial value) decrease in starting stresses for dislocation motion.

Thus, these results can be used to formulate the following preliminary model of the observed effect.

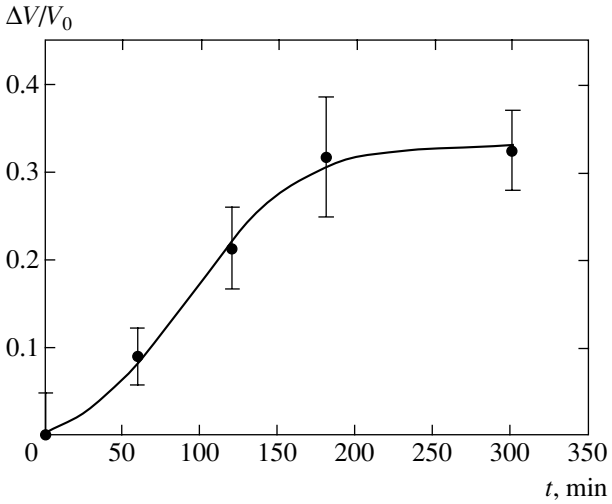

Fig. 1. Variation of the velocity $\Delta V/V_0 = (V_B - V_0)/V_0$ of 60° dislocations in Cz–Si as a function of exposure time for a sample in a magnetic field of $B = 20$ kOe at room temperature. The dislocation velocity was measured at 550°C under a shearing stress of $\tau = 50$ MPa.

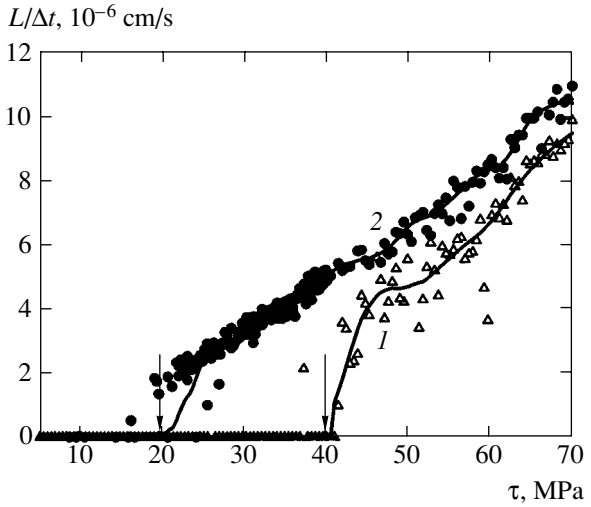

Fig. 2. Dependence of the velocity of individual 60° dislocations on the shearing stress, measured at 600°C on a Cz–Si sample before (curve *1*) and after (curve *2*) exposure of the sample for 3 h to a magnetic field of $B = 20$ kOe at room temperature. Symbols represent data for solitary dislocations, while the curves were obtained by averaging (parabolic smoothing) over 13 dislocations. Arrows indicate starting stresses.

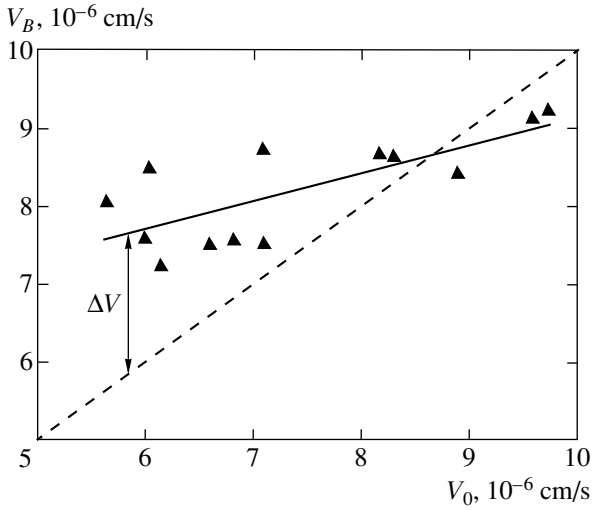

Fig. 3. Velocity V_B of 60° dislocations after exposure to magnetic field as a function of velocity V_0 of these dislocations before exposure to the field for several different Cz–Si samples. The dashed line shows the dependence in zero magnetic field, i.e., for $V_B = V_0$. The velocity of dislocations was measured at 600°C under a sharing stress $\tau = 50$ MPa.

The Cz–Si samples exhibiting a decrease in the starting stresses and an increase in the mobility of individual dislocations as a result of their exposure to magnetic field differ from the Fz–Si samples in which this effect is not observed at a considerably higher concentration of interstitial oxygen.

Theoretical calculations [2, 3] show that the energy of oxygen molecules in dislocation cores is considerably lower than the energy of interstitial molecular oxygen in the ideal silicon lattice; i.e., we can speak of the presence of a considerable binding energy between oxygen molecules and a dislocation. Oxygen molecules in dislocation cores in silicon may be in several different configurations differing in binding energy. Thus, we can assume that the dislocations implanted into the crystal accumulate oxygen in their cores due to reactions between dissolved oxygen and a moving dislocation; this must reduce the dislocation mobility and lead to the emergence of considerable starting stresses. It is known from experiments that the presence of oxygen in silicon indeed increases the starting stresses for dislocation motion [12, 14].

The diffusion coefficient for interstitial oxygen at room temperature is very small (below 10^{-22} cm^2/s) [15]; for this reason, the average diffusion length for oxygen over the time of magnetic field treatment (3 h) hardly exceeds the unit cell size for silicon. Thus, we can disregard the effect of a magnetic field on the growth of oxygen precipitates or the mean oxygen concentration in the bulk of the crystal. We can assume that the magnetic field leading to singlet–triplet transitions in thermally excited states of silicon–oxygen complexes stimulates the change in the state (configuration) of oxygen molecules, which have already occupied dislocation cores, so that the average binding energy U_B between oxygen and dislocations decreases. The decrease in the binding energy reduces starting stresses ($\tau_{st} \sim NU_B$). This also leads to oxygen loss by dislocations after their motion begins [16]. Impurities located in a dislocation core decelerate its motion; consequently, a decrease in the oxygen concentration in a dislocation core increases the velocity of the dislocation.

It should be noted that the amount of oxygen accumulated by a dislocation during the formation of a dislocation half-loop depends on the path traversed by the dislocation, the time the sample is held at a high temperature, the concentration of the defects of dislocation core reconstruction, and other parameters depending on past history. This must lead to a spread in the values of starting stress and velocity of dislocations in different samples (the larger the amount of accumulated oxygen, the smaller the dislocation velocity V_0 at a fixed load must be). Assuming that the magnetic field stimulates the detachment of oxygen from a dislocation, we can expect that the dislocation velocities V_B in different samples exposed to magnetic field must tend to a certain value typical of "pure" dislocations. This is actually observed in experiment. Figure 3 shows the dependence of the average velocity V_B of dislocations after the action of magnetic field on velocity V_0 of these dislocations before the action of magnetic field measured for a large number of samples with different past histories. It can be seen that magnetic field weakly affects the velocity of dislocations in samples with a high initial velocity of dislocations, in which the concentration of oxygen strongly coupled with dislocations is low.

The proposed model is a preliminary one; a number of additional experiments are required for determining the microscopic mechanisms of spin-dependent reactions occurring in the oxygen-dislocation system.

ACKNOWLEDGMENTS

The authors are grateful to V.I. Alshits and Yu.I. Golovin for fruitful discussions.

This study was supported financially by the Russian Foundation for Basic Research (project no. 02-02-17024).

REFERENCES

1. V. Kveder, M. Kittler, and W. Schroeter, Phys. Rev. B **63**, 115208 (2001).
2. A. Umerski and R. Jones, Philos. Mag. A **67**, 905 (1993).
3. M. I. Heggie, S. Jenkins, C. P. Ewels, *et al.*, J. Phys.: Condens. Matter **12**, 10263 (2000).
4. A. L. Buchachenko, R. Z. Sagdeev, and K. M. Salikhov, *Magnetic and Spin Effects in Chemical Reactions* (Nauka, Novosibirsk, 1978), p. 296.

5. V. I. Al'shits, E. V. Darinskaya, T. M. Perekalina, and A. A. Urusovskaya, Fiz. Tverd. Tela (Leningrad) **29**, 467 (1987) [Sov. Phys. Solid State **29**, 265 (1987)].
6. Yu. I. Golovin, R. B. Morgunov, A. A. Baskakov, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 114 (1999) [JETP Lett. **69**, 127 (1999)].
7. Yu. A. Osip'yan, Yu. I. Golovin, D. V. Lopatin, *et al.*, Pis'ma Zh. Éksp. Teor. Fiz. **69**, 110 (1999) [JETP Lett. **69**, 123 (1999)].
8. V. V. Kveder, Yu. A. Osip'yan, and A. I. Shalynin, Zh. Éksp. Teor. Fiz. **83**, 699 (1982) [Sov. Phys. JETP **56**, 389 (1982)].
9. M. N. Levin and B. A. Zon, Zh. Éksp. Teor. Fiz. **111**, 1373 (1997) [JETP **84**, 760 (1997)].
10. A. A. Skvortsov, A. M. Orlov, and L. I. Gonchar, Zh. Éksp. Teor. Fiz. **120**, 134 (2001) [JETP **93**, 117 (2001)].
11. V. N. Erofeev and V. I. Nikitenko, Zh. Éksp. Teor. Fiz. **60**, 1780 (1971) [Sov. Phys. JETP **33**, 963 (1971)].
12. S. Senkader, A. Giannattasio, R. Flaster, and P. R. Wilshaw, J. Phys.: Condens. Matter **14**, 13141 (2002).
13. B. V. Petukhov, Fiz. Tverd. Tela (Leningrad) **24**, 439 (1982) [Sov. Phys. Solid State **24**, 248 (1982)].
14. K. Sumino, Phys. Status Solidi A **171**, 111 (1999).
15. S. Senkader, P. R. Wilshaw, and R. J. Falster, J. Appl. Phys. **89**, 4803 (2001).
16. B. V. Petukhov, Zh. Tekh. Fiz. **60** (10), 64 (1990) [Sov. Phys. Tech. Phys. **35**, 1150 (1990)].

Translated by N. Wadhwa

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 606–614.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 670–679.

SOLIDS
Structure

Stability of Negative Ions near the Surface of a Solid¶

D. I. Zhukhovitskii[a,*], W. F. Schmidt[b], and E. Illenberger[b]

[a] *Joint Institute for High Temperatures, Russian Academy of Sciences, Izhorskaya ul. 13/19, Moscow, 125412 Russia*

[b] *Institut für Physikalische und Theoretische Chemie, Freie Universität Berlin D-14195, Berlin, Germany*

**e-mail: dmrzh@orc.ru*

Received February 4, 2003

Abstract—Stationary states of molecular negative ions (anions) near the surface of a solid are investigated. The lone electron is assumed to interact with a diatomic molecule and the surface of the solid. The energies of electron levels are determined by solving the 2D Schrödinger equation. It is shown that its stable solutions exist at distances from the surface greater than some critical distance, otherwise the electron is detached from the anion. In the case of attraction between the electron and the solid, the interaction potential between the anion and the solid appears to have the Lennard–Jones form and the ion is separated from the surface by some equilibrium distance. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The interaction of electronegative molecules and negative molecular ions (anions) with surfaces of condensed state has extensively studied within the last decades. One of the problems most interesting to us is the formation and evolution of negative ion resonances (NIRs) on surfaces [1–4]. As in the gas phase, NIRs can also be generated at surfaces by the attachment of free electrons to absorbed molecules at a defined energy [1–3]. It has been recognized that the photochemical behavior of adsorbed molecules can effectively be governed by the photoinduced electron transfer from the substrate to the adsorbed molecule [5–7]. In some systems, NIRs thus formed are considered the driving force for the respective photochemical reaction.

The formation and evolution of NIRs are usually appreciably modified when passing from the gas phase to the surface [8, 9]. This concerns the energy of the temporary negative ion, its lifetime (with respect to the electron loss and dissociation), and the branching ratios between the dissociative attachment channels.

Some electron scattering experiments have been performed on molecules deposited on cold noble gas substrates (solids). Here, a noble gas layer of variable thickness can be used to study the influence of the metallic substrate on the particular process, e.g., the (dissociative) attachment cross section or the energy shift of the negative ion resonance [8–10]. In addition, substrate-induced electron transfer reactions via the initial formation of an electron exciton precursor in the noble gas layer could be observed [3, 13].

The stability of molecular negative ions at or near the surface is an essential point in the investigation of low-energy electron-driven reactions on adsorbed and condensed molecules [14]. In addition to its importance from the standpoint of basic science, the stability of negatively charged particles at or near a solid surface is an important issue in many technological processes like photocopying, laser printing, etc.

In [15], it was shown that the interaction of a negative ion with a nonpolar liquid results in a considerable shift of its photodetachment threshold. Apparently, noticeable shifts can be expected in the interaction of an anion and the surface of a solid. In [16–18], the dynamics of the processes of charge transfer and production of molecular anions in the vicinity of a surface were investigated. It was recognized that an accurate calculation of electron energies requires solving the 2D Schrödinger equation, because the spherical symmetry for the lone electron is broken at a short distance from the surface [18]. Consequently, on the basis of one-dimensional perturbation theory, sufficiently accurate values of the electron energy cannot be found at short distances from the surface because higher order terms require nonspherical corrections to the wave function.

The objective of this paper is to calculate the stationary-state energy of the lone electron of a molecular anion near the surface of a solid by solving the stationary 2D Schrödinger equation. The interaction of the lone electron with the diatomic molecule is described by a polarization pseudopotential; the interaction with the solid is characterized by a single parameter, the effective energy of the electron inside the solid, V_0.

In this formulation, the problem has a two-dimensional axial symmetry with the axis perpendicular to the surface, and we must therefore solve the 2D

¶This article was submitted by the authors in English.

Schrödinger equation. Solutions of the multidimensional Schrödinger equation are known in the cases where spatial variables can be separated (as in the theory of the Stark effect). In the case of inseparable variables, due to the complexity of the general mathematical formulation, no regular methods, either analytic or numerical, have been developed to date. As a rule, the multidimensional Schrödinger equation is reduced to a quasi-one-dimensional one, specific for the problem under consideration. There are several approaches to the treatment of multidimensional equations. Among them, the split-step Fourier scheme [19] was used in [20] for the investigation of white noise in the 2D nonlinear Schrödinger equation. The 3D wave packet propagation method was used in [21] to describe the propagation of an electron near the surface of a solid.

We propose a solution of this problem assuming that the electron state is in fact a superposition of states with different values of the angular momentum in the corresponding effective spherically symmetric potentials. The state with zero angular momentum is assumed to predominate. An exact solution of the Schrödinger equations is then sought as a linear combination of the wave functions corresponding to different values l of the angular momentum. Thus, the 2D Schrödinger equation is reduced to an infinite set of one-dimensional equations for the radial wave functions. It can be shown that its solution can be approximated to a very good accuracy by the truncation of the infinite set to two equations for $l = 0$ (zero approximation) and $l = 1$ (first approximation). The latter is easily solved numerically by the iteration method. This makes possible calculation of the lone electron energy as a function of the parameters characterizing its total interaction potential.

Two cases must be distinguished, the repulsive surface (potential barrier) and the attractive surface (potential well). Repulsion of the electron from the surface causes pure repulsion of the anion from the surface; there is a minimum distance at which a stationary state is possible. At shorter distances, nonstationary states emerge, which prove to have decay times too short to be detected experimentally. Thus, detachment of the electron from a diatomic molecule occurs, and the electron is removed into the vacuum.

In the case of an attractive surface, the existence of an electron stationary state depends on the values of potential parameters. If the potential well in a solid is too deep, no stationary state is possible and the electron is detached from the anion at the distance where the anion decay time becomes shorter than its residence time near the surface of the solid. If the well depth is moderate, there is a finite range with some minimum and maximum distances from the surface where a stationary solution exists. If the potential well is shallow, there is only the minimum distance, as in the case of a potential barrier. In the case of surface attraction, detachment of the lone electron implies its tunneling into the potential well of a solid. If stationary states exist, the curve of interaction between the anion and the surface of the solid has a form characteristic of the interatomic interaction (the Lennard-Jones potential). As is known, this potential has an equilibrium distance. Hence, a molecule-like equilibrium state of the anion near the surface emerges. This makes it possible to predict not only the shift of the electron level (of the electron photodetachment threshold), but also the distance from the surface at which the anion can be found.

This paper is organized as follows. In Section 2, the interaction potential for the lone electron and the Schrödinger equation to be solved are written; in Section 3, its asymptotically exact solution is found and the first-order approximation to this solution is considered. Potential curves for anions near the surface are calculated in Section 4; the results obtained are analyzed in Section 5.

2. PROBLEM FORMULATION

We consider a system consisting of a highly polarizable diatomic molecule, the surface of a solid, and a lone electron. The interactions between the molecule and the surface are assumed to be negligibly small, and we can therefore take into account only the interactions between the lone electron and molecule, and between the electron and the surface. We let V_p and V_s denote the respective interaction potentials.

We introduce the spherical coordinate system with the origin at the point of location of the molecule and with the polar axis perpendicular to the surface. The polar axis is directed toward the surface. The corresponding spherical coordinates are denoted by r and θ. The distance between the molecule and the surface is z_0. The half-space $r\xi < z_0$ ($\xi = \cos\theta$) is the vacuum, and the other half-space $r\xi \geq z_0$ is occupied by the solid; the surface is defined by the equation $r\xi = z_0$. At a sufficiently large distances from the molecule in the vacuum, $r \geq r_c$, where r_c is the molecule hard-core radius, the lone electron polarizes core electrons of the molecule, and the interaction can be described by the polarization potential

$$V_p = -\alpha r^{-4},$$

where α is the molecule polarization in the units of a_0^3 (a_0 is the Bohr radius) [22]; the length and energy are measured in the units of a_0 and in Ry, respectively (in contrast to atomic units, we measure the energy in Ry). At small distances $r < r_c$, due to the Pauli principle, a short-range repulsion occurs, and we can therefore set

$$V_p = +\infty \text{ at } r < r_c$$

(see [22]). Thus, we use a spherically symmetric pseudopotential to describe a lone electron. Because a real diatomic molecule is not spherically symmetric,

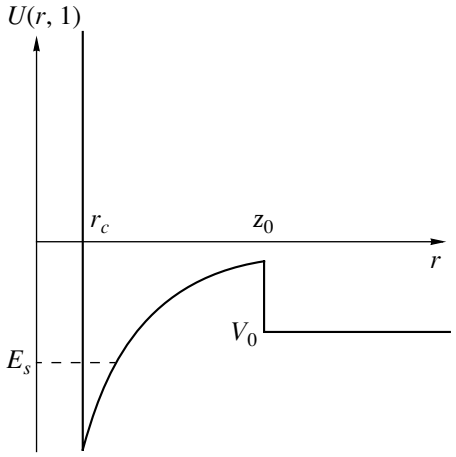

Fig. 1. Interaction potential for the attractive surface along the polar axis.

this assumption is made for simplicity. We also assume that the electron charge is screened inside the solid and therefore does not interact with the molecule, which allows us to set $V_p = 0$ for $r\xi \geq z_0$.

The lone electron polarizes the surface of the solid, and the arising electrostatic image force is responsible for the interaction between the electron and the surface at large distances. At short distances and inside the solid, the electron experiences attraction caused by polarization of surrounding molecules by its charge and the Pauli repulsion. For simplicity, we do not take the details of this interaction into account and use its simplest form instead.

We note that the characteristic length of variation of the image force potential is given by several a_0, whereas the characteristic length of the lone electron localization (the width of the electron wave function) is about $10a_0$. Obviously, V_s must be uniform inside the solid. This enables approximation of V_s by a simple "step" potential $V_s = 0$ for $r\xi < z_0$ and $V_s = V_0$ for $r\xi \geq z_0$. The interaction between the electron and the surface is therefore allowed by a single parameter V_0. This parameter characterizes the interaction as a whole and is essentially an effective one. Its sign depends on the ratio between the strengths of the attractive electrostatic image force outside the solid and the Pauli repulsion inside it. This ratio is defined by constants characterizing the molecule and the solid. In some cases (e.g., the surface of a metal), the image force predominates, and $V_0 < 0$. In this case, a potential well occurs, which causes attraction, at large distances at least. If $V_0 > 0$, the solid is represented by a potential barrier and the net effect is repulsion.

Although the details of the potential V_s can be included (as, e.g., in [21]), this can considerably complicate the analysis of the results obtained and mask the nature of the effects that we want to demonstrate. At the same time, any complicated form of the electron-surface interaction potential can be included in our formalism if necessary.

Thus, the total electron interaction potential

$$U = V_p + V_s$$

is given by

$$U(r, \xi) = \begin{cases} +\infty, & r < r_c \\ -\alpha r^{-4}, & \xi r_c \leq \xi r < z_0 \\ V_0, & \xi r \geq z_0. \end{cases} \quad (1)$$

It is shown in Fig. 1. Because potential (1) has axial symmetry, the electron energy is found from the 2D stationary Schrödinger equation

$$\hat{H}\psi = E_s\psi, \quad (2)$$

where $\psi(r, \xi)$ is the electron wave function and E_s is the energy. In spherical coordinates, the Hamiltonian $\hat{H}$ is given by

$$\hat{H} = -\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) - \frac{1}{r^2}\left[(1-\xi^2)\frac{\partial^2}{\partial \xi^2} - 2\xi\frac{\partial}{\partial \xi}\right] + U(r, \xi). \quad (3)$$

The boundary conditions for the wave function are

$$\psi(r_c, \xi) = \psi(\infty, \xi) = 0. \quad (4)$$

3. METHOD OF SOLUTION

To solve Eq. (2), we have to additionally assume that its solution has the same symmetry as the potential $U(r, \xi)$ in Eq. (1), i.e., axial symmetry. This is correct for the ground state, which is a single state for most diatomic anions in the approximation of the simplified electron–molecule interaction potential V_p.

A solution of Eq. (2) can be represented as an expansion in any complete set of functions of ξ. Similarly to the quantum scattering theory, we use the Legendre polynomials $P_l(\xi)$. However, in contrast to scattering, we seek a stationary bound state localized in a finite spatial region with the real energy E_s less than the minimum value of potential (1) at $r \longrightarrow \infty$. Thus, we can represent the solution as a series in the Legendre polynomials

$$\psi(r, \xi) = \frac{1}{r}\sum_{l=0}^{\infty}\varphi_l(r)P_l(\xi). \quad (5)$$

The wave function ψ is normalized to unity,

$$2\pi\int_0^\infty r^2 dr\int_{-1}^{1}|\psi(r,\xi)|^2 d\xi = 1. \quad (6)$$

It follows from the discussion above and condition (6) that $\varphi_l(r)$ are real functions. Substituting (5) into (6) and recalling the normalization and orthogonality condition for the Legendre polynomials,

$$\int_{-1}^{1}P_l^2(\xi)d\xi = \frac{2}{2l+1},\quad \int_{-1}^{1}P_l(\xi)P_k(\xi)d\xi = 0,\quad l\neq k,$$

we obtain the normalization condition for the functions $\varphi_l(r)$,

$$4\pi\sum_{l=0}^{\infty}\left[\frac{1}{2l+1}\int_0^\infty\varphi_l^2(r)dr\right] = 1. \quad (7)$$

Because the Legendre polynomials are eigenfunctions of the square angular momentum operator,

$$(\xi^2-1)\frac{d^2P_l}{d\xi^2}+2\xi\frac{dP_l}{d\xi} = l(l+1)P_l, \quad (8)$$

substitution of expansion (5) into Schrödinger equation (2) yields

$$-\sum_{k=0}^{\infty}P_k\frac{d^2\varphi_k}{dr^2}+\frac{1}{r^2}\sum_{k=0}^{\infty}k(k+1)P_k\varphi_k + U\sum_{k=0}^{\infty}P_k\varphi_k = E_s\sum_{k=0}^{\infty}P_k\varphi_k. \quad (9)$$

We multiply both sides of Eq. (9) with $P_l(\xi)$ and integrate over ξ from –1 to 1 to derive

$$\frac{d^2\varphi_l}{dr^2}+\left[E_s-\frac{l(l+1)}{r^2}\right]\varphi_l-\sum_{k=0}^{\infty}\bar{V}_{lk}\varphi_k = 0, \quad (10)$$

where the matrix elements

$$\bar{V}_{lk}(r) = \left(l+\frac{1}{2}\right)\int_{-1}^{1}U(r,\xi)P_l(\xi)P_k(\xi)d\xi \quad (11)$$

are the effective spherically symmetric potentials; they are related by the ratio

$$\bar{V}_{lk} = \frac{2l+1}{2k+1}\bar{V}_{kl}. \quad (12)$$

The boundary conditions for set of equations (10) are

$$\varphi_l(r_c) = \varphi_l(\infty) = 0. \quad (13)$$

Thus, we have reduced the 2D Schrödinger equation to an infinite set of ordinary differential equations (10), each of which corresponds to a certain value of the angular momentum.

Solution of Eqs. (10) makes sense only if series (5) converges fast. This means that some state must predominate in superposition (5). Because the increase in the hard-core parameter r_c of the potential V_p by a few percent leads to the disappearance of the bound state of an isolated anion, we can assume that if the lone electron is localized on the molecule, the deviation of its wave function from the spherically symmetric form is moderate. Hence, the *s*-state ($l = 0$) must predominate, and we can truncate the set of equations (10) at some finite value of l. The error involved in this truncation can easily be estimated by inclusion of a higher order equation. Thus, (10) can be regarded as a key to obtain an asymptotically exact solution.

In the zero approximation ($l = 0$), Eqs. (10) are reduced to the one-dimensional Schrödinger equation for the radial wave function $\varphi_0(r)/r$ in the spherically averaged interaction potential $\bar{V}_{00}(r)$ in Eq. (1),

$$\frac{d^2\varphi_0}{dr^2}+[E_s-\bar{V}_{00}(r)]\varphi_0 = 0, \quad (14)$$

where

$$\bar{V}_{00}(r) = \frac{1}{2}\int_{-1}^{1}U(r,\xi)d\xi = V_p(r)+\frac{1}{2}\int_{-1}^{1}V_s(r,\xi)d\xi = \begin{cases}+\infty, & r<r_c,\\ -\alpha r^{-4}, & r_{hc}\le r<z_0,\\ \dfrac{V_0}{2}\left(1-\dfrac{z_0}{r}\right)-\dfrac{\alpha}{2r^4}\left(1+\dfrac{z_0}{r}\right), & r\ge z_0.\end{cases} \quad (15)$$

In the zero approximation, nonsphericity is obviously not allowed. Far from the surface, the lone electron wave function can be approximated by that of an isolated anion $\varphi_\infty(r)/r$. By definition,

$$\frac{d^2\varphi_\infty}{dr^2}+[E_0-V_p(r)]\varphi_\infty = 0, \quad (16)$$

where

$$E_0 = \lim_{z_0\to\infty}E_s$$

is the electron energy of an isolated anion ($-E_0$ is the electron affinity of an isolated molecule), and we can therefore obtain the electron energy from (14) as

$$E_s \approx E_0 + 2\pi \int_{-1}^{1} \varphi_\infty^2(r) V_s(r, \xi) d\xi.$$

This coincides exactly with the result of perturbation theory if V_s is treated as a small perturbation.

At the surface (small distance z_0), V_s is not small and the first order of perturbation theory therefore does not provide high accuracy. Calculation of higher order corrections of perturbation theory is impossible, because the nonsphericity effect is not included. But we can handle small distances z_0 even in the first approximation on the basis of set of equations (10). In this approximation, (10) is truncated into two equations for $l = 0$ and 1,

$$\frac{d^2\varphi_0}{dr^2} + [E_s - \bar{V}_{00}(r)]\varphi_0 - \bar{V}_{01}(r)\varphi_1 = 0, \quad (17)$$

$$\frac{d^2\varphi_1}{dr^2} + \left[E_s - \frac{2}{r^2} - \bar{V}_{11}(r)\right]\varphi_1 - 3\bar{V}_{01}(r)\varphi_0 = 0, \quad (18)$$

where relation (12) is used,

$$\bar{V}_{01}(r) = \frac{1}{2}\int_{-1}^{1} U(r, \xi) d\xi$$

$$= \begin{cases} 0, & r < z_0 \\ \left(1 - \dfrac{z_0^2}{r^2}\right)\left(\dfrac{\alpha}{4r^4} + \dfrac{V_0}{4}\right), & r \geq z_0, \end{cases} \quad (19)$$

$$\bar{V}_{11}(r) = \frac{3}{2}\int_{-1}^{1} U(r, \xi)\xi^2 d\xi$$

$$= \begin{cases} +\infty, & r < r_c \\ -\alpha r^{-4}, & r_c \leq r < z_0 \\ \dfrac{V_0}{2}\left(1 - \dfrac{z_0^3}{r^3}\right) - \dfrac{\alpha}{2r^4}\left(1 + \dfrac{z_0^3}{r^3}\right), & r \geq z_0, \end{cases} \quad (20)$$

and the wave function is the sum of two terms

$$\psi(r, \xi) \approx \frac{1}{r}[\varphi_0(r) + \xi\varphi_1(r)] \quad (21)$$

normalized by the condition

$$4\pi \int_0^\infty [\varphi_0^2(r) + \varphi_1^2(r)/3] dr = 1.$$

We can estimate the accuracy of this approximation by comparing the electron energy $E_s^{(0)}$ obtained from (14) with that calculated using (17) and (18), $E_s^{(1)}$.

Set of equations (17) and (18) can be solved numerically using the iteration method. First, we set

$$\varphi_1(r) \equiv 0,$$

and (17) becomes a one-dimensional equation. This equation is solved with the boundary conditions

$$\varphi_0(r_c) = 0, \quad \left(\frac{d\varphi_0}{dr}\right)_{r = r_c} = C_0,$$

where C_0 is an arbitrary number; the parameter E_s appearing in the equation is adjusted to satisfy the boundary condition

$$\varphi_0(\infty) = 0.$$

The obtained value of E_s and the calculated function $\varphi_0(r)$ are then substituted into Eq. (18), which is solved with the boundary conditions

$$\varphi_1(r_c) = 0, \quad \left(\frac{d\varphi_1}{dr}\right)_{r = r_c} = C_1.$$

The value of C_1 is then varied to satisfy the boundary condition

$$\varphi_1(\infty) = 0.$$

For the next iteration, the calculated function $\varphi_1(r)$ is substituted into Eq. (17), etc. This iteration procedure terminates when the difference in the values of E_s obtained from successive iterations, $E_s^{(0)} - E_s^{(1)}$, is sufficiently small. Obviously, the procedure described can also be used for a set containing more than two equations.

4. CALCULATION RESULTS

For numerical calculations, we considered two molecules, O_2 with a moderate polarizability and highly polarizable Br_2. The molecule O_2 was selected because it is widely used in experiments. In addition, in spite of the obvious internal asymmetry, the excess electron can be approximately treated as a lone electron, which makes it possible to solve the one-electron problem. Br_2 is an example of a dimer with a very high polarizability. Polarizabilities α and electron affinities $-E_0$ for these molecules [23] are listed in table. The values of hard-core radii r_c of the potential V_p were adjusted to fit the corresponding experimental electron affinities by the values of $-E_0$ obtained from the ground state solution of Eq. (16).

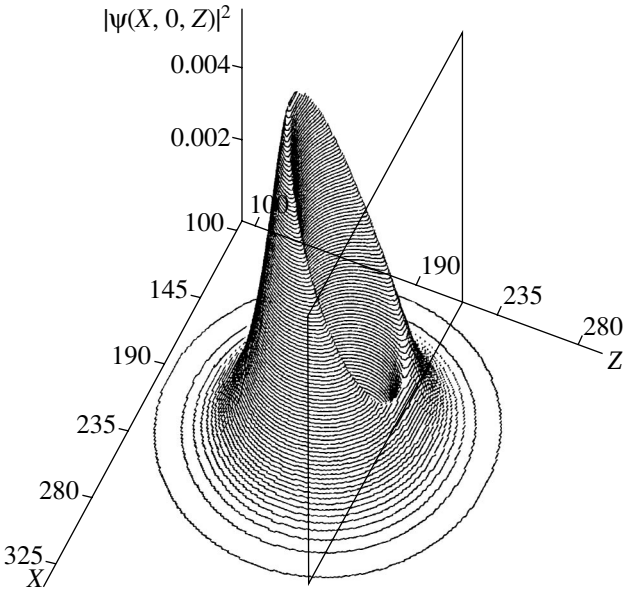

Fig. 2. Wave function for Br_2^- in the *XZ* plane (repulsive surface, $V_0 = 1$ eV). The plane $Z = 227$ ($z_0 = 2a_0$) indicates the location of the surface of the solid.

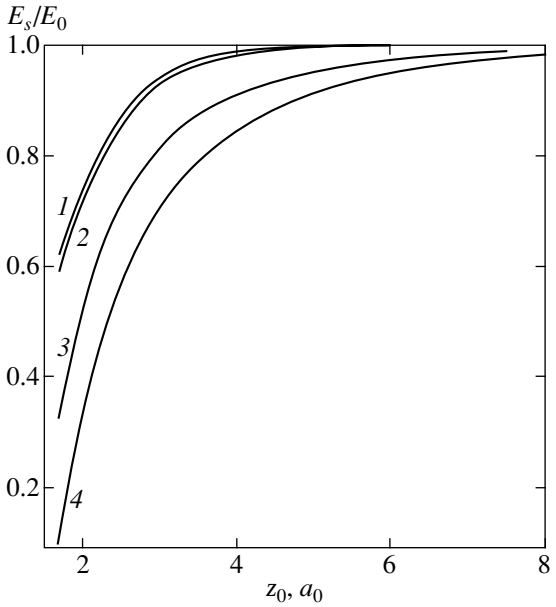

Fig. 3. Lone electron energy as a function of the distance from a repulsive surface. *1*—Br_2^- for $V_0 = 0.5$ eV; *2*—Br_2^- for $V_0 = 1$ eV; *3*—O_2^- for $V_0 = 0.5$ eV; *4*—O_2^- for $V_0 = 1$ eV.

Set of equations (17) and (18) was solved numerically using the procedure discussed in the previous section for different values of the interaction potential parameter V_0. We first consider positive V_0. Figure 2 illustrates the solution of Eqs. (17) and (18) for Br_2^- at $V_0 = 1$ eV and $z_0 = 2a_0$. In this figure, wave function (21) is plotted in the *XZ* plane ($Y = 0$) of the Cartesian coordinate system whose *Z* axis is parallel to the polar axis of the spherical coordinate system used in the foregoing. The spherical coordinates are related to the Cartesian ones as follows:

$$r = \sqrt{x^2 + z^2}, \quad \xi = z/r,$$

$$x = 0.075(X - 200), \quad z = 0.075(Z - 200).$$

It is seen in Fig. 2 that the front of the wave function is lowered near the repulsive surface. This is indicative of a considerable repulsion of the wave function from the surface. The ratio

$$\frac{E_s^{(1)} - E_s^{(0)}}{E_0} = 0.17$$

is moderate, however.

The lone electron energy as a function of the distance from the solid surface is shown in Fig. 3. It is seen that the results are similar for O_2^- and Br_2^-, although the differences in polarizabilities and electron affinities for these molecules are about an order of magnitude. Each curve drops abruptly at some small distance, and E_0 vanishes, which is indicative of the existence of a minimum distance at which a stable state of the anion is possible (at this distance, $E_s = 0$). At shorter distances, the electron is detached from the anion and is removed to infinity in the vacuum, which means electron detachment. Fast vanishing of the effect of the surface as the anion moves away from the surface is also seen in this figure. We note that the electron energy at a short distance from the surface cannot be calculated on the basis of perturbation theory; at moderate distances, the results of the latter are of no interest due to this vanishing.

In the case of an attractive surface, the form of the wave function is qualitatively similar to the previous case, but the front of the wave function is raised near the attractive surface, which is indicative of the attraction of the wave function. The potential curve $E_s(z_0)$ is more complicated, however. In Fig. 4, the relative energy shifts are shown for the values of V_0 that coincide with the corresponding energies E_0 of isolated anions. For both molecules, potential curves have the form of the

Parameters of diatomic molecules

Molecule	Polarizability α, a_0^3	Electron affinity E_0, eV	Cutoff radius r_c, a_0
O_2	10.6	0.46	0.909
Br_2	43.6	2.6	1.457

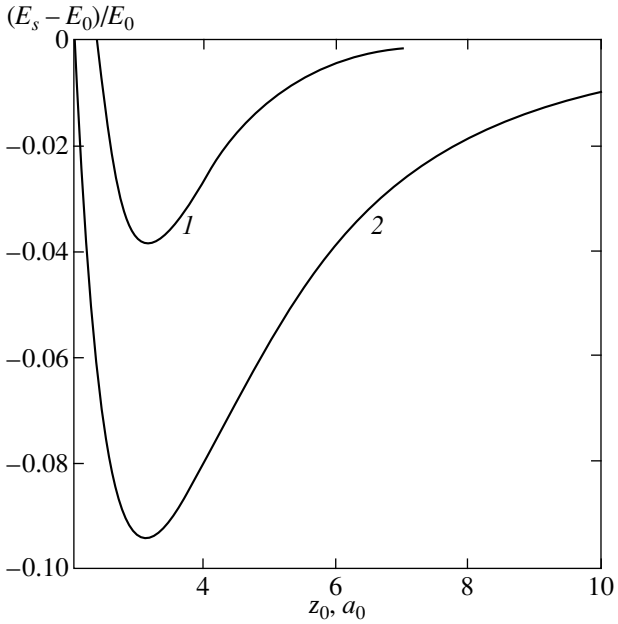

Fig. 4. Relative electron energy shifts as functions of the distance from an attractive surface at $V_0 = E_0$. *1*—Br_2^-; *2*—O_2^-.

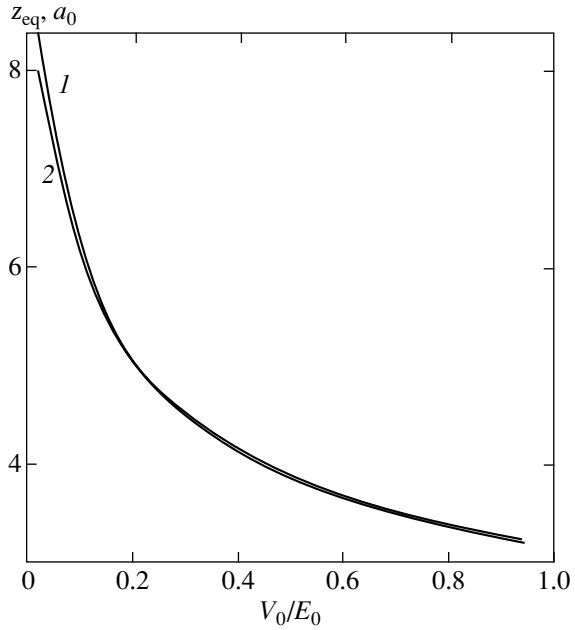

Fig. 5. Equilibrium distance as a function of V_0 for an attractive surface. *1*—Br_2^-; *2*—O_2^-.

Lennard–Jones interatomic potential with a short-range repulsion and vanishing long-range attraction. This behavior of a potential curve holds for any negative value of V_0. The reason of such behavior is as follows. At a large separation z_0 from the surface, $|V_p(z_0)| < |V_0|$, and the effect of the surface amounts to lowering the energy E_s. Because $|V_p(r_c)| \gg |E_0|$, the inequality $|V_p(z_0)| > |V_0|$ holds at sufficiently short distances (interaction of the electron with the molecule is stronger than with the surface), which results in the increase of E_s due to the confinement of space available for the electron and in the consequent increase in its kinetic energy.

A stable electron state is realized only if the energy E_s is less than the minimum value $\lim_{r \to \infty} U(r, \xi)$; otherwise, the tunneling to the region of the lower potential must lead to electron detachment. In the previously considered case of a repulsive potential ($V_0 > 0$), this minimum is equal to zero at $\xi = -1$. If $V_0 < 0$, it is equal to V_0 at $\xi = 1$. In the latter case, the region of stable states is therefore limited by the condition $E_s < V_0$.

In the case of an attractive surface, the existence of a stable state depends on the ratio of parameters E_0 and V_0. If $|E_0| > |V_0|$, a bound state is possible at $z_0 \longrightarrow \infty$ at least, where $E_s = E_0$. Therefore, a bound state exists at a sufficiently large distance $z_0 > z_{min}$, where z_{min} is a single root of the equation $E_s(z_0) = V_0$. If $|E_0| < |V_0|$, electron detachment occurs at large distances and the region of possible bound states is finite: $z_{min} < z_0 < z_{max}$, where z_{min} and z_{max} are two roots of the equation $E_s(z_0) = V_0$. As V_0 increases, $z_{min} \longrightarrow z_{max}$, and at some value $V_0 = V_{min}$, $z_{min} = z_{max}$. Obviously, this case corresponds to the bottom of the potential curve $E_s(z_0)$. At $|V_0| > |V_{min}|$, no stable state is realized; in this case, only electron detachment is possible.

We note that independently of z_0 and the sign of V_0, the interaction with a surface leads to a decrease in anion stability, $|E_s(V_{min}) - V_{min}| < |E_0|$, due to broken spherical symmetry of the lone electron state in an isolated anion.

If we neglect the interaction between the molecule and the surface, which is most likely a hard-core attraction at small distances, then the energy $E_s - E_0$ is that of the anion as a whole. The condition

$$\left(\frac{dE_s}{dz_0}\right)_{z = z_{eq}} = 0$$

defines some distance z_{eq} at which the energy $E_s(z_0)$ reaches the minimum and an equilibrium bound state of the anion at the solid surface is realized. We note that the interaction between the molecule and the metal surface may noticeably contribute to the total energy of the anion, especially in the region of short distances z_0, but we can expect that this does not change the situation qualitatively. Equilibrium distances z_{eq} and the corresponding equilibrium state energies $E_{eq} = E_s(z_{eq})$ are shown in Figs. 5 and 6 as functions of the potential well

depth V_0. It is seen in Fig. 5 that the equilibrium distance increases sharply as $V_0 \longrightarrow 0$ and that the curves for O_2^- and Br_2^- almost coincide. The quantities $|E_{eq}(V_0)|$ increase with $|V_0|$ until the condition $E_s = V_0$ (dashed curve in Fig. 6) is satisfied. In Fig. 6, the intersection points of solid curves with the dashed curve indicate the maximum well depths at which bound states can occur. For O_2^-, the maximum value of V_0/E_0 is 1.04; for Br_2^-, it is 1.12.

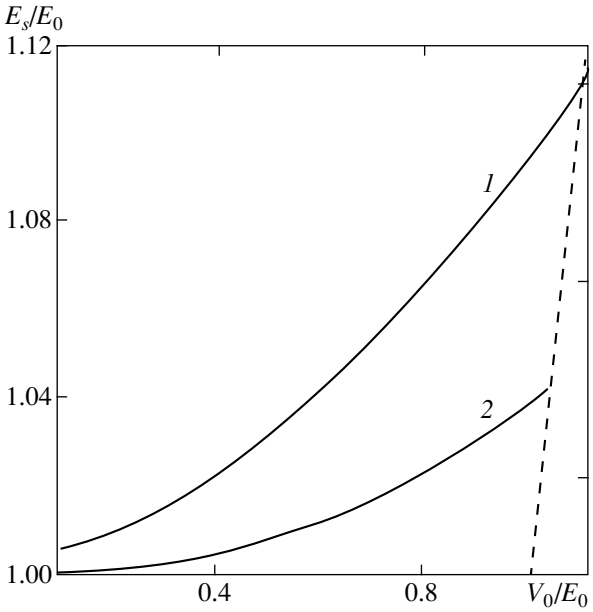

Fig. 6. Equilibrium energies as functions of V_0 for an attractive surface. *1*—Br_2^-; *2*—O_2^-; dashed curve indicates the line $E_s = V_0$.

5. DISCUSSION

In this paper, we have solved the 2D Schrödinger equation for a lone electron interacting with a diatomic molecule and the surface of a solid. Our solution makes it possible to calculate electron energies of an anion at the surface. In the case of an attractive surface, we predict molecule-like bound states of the anion, which are realized if the potential well depth $|V_0|$ characterizing the interaction with the surface does not exceed some threshold value. We have determined this value, above which the electron detachment occurs and the lone electron is removed either into the vacuum or into the solid.

Molecule-like states of the system under consideration can exist only if the vibrational energy quantum $\hbar\omega_v$ is much smaller than the well depth $E_0 - E_s$. For the potential curves shown in Fig. 4, the estimate

$$\hbar\omega_v \sim \frac{\hbar}{z_{eq} - z_{min}}\sqrt{\frac{2(E_0 - E_{eq})}{M}}$$

is valid, where M is the mass of a diatomic molecule. Because the ratio m/M is small, the resulting condition

$$\frac{4m}{M}a_0^2(z_{eq} - z_{min})^{-2}\mathrm{Ry}(E_0 - E_{eq})^{-1} \ll 1$$

is satisfied for both Br_2^- and O_2^- (the product in the left-hand side is on the order of 10^{-3} for Br_2^- and 10^{-2} for O_2^-). We note that a similar phenomenon of levitation above the surface is known for liquid helium [24].

The objective of introducing potential (1) was to qualitatively include all possible cases of the interaction between the lone electron and the solid. For some particular anion and a surface, the interaction potential may not be reduced to its simplest form (1). A rigorous answer to the question concerning the existence of an equilibrium state for the given experimental conditions implies calculation of a real interaction potential between the lone electron and the surface. This complicated problem (see, e.g., [21]) requires additional investigation.

It is natural to discuss the lifetimes of transient anion states near the surface. The lifetime of an anion above the electron detachment threshold can be estimated as the time of tunneling under the barrier with the height $-E_0$ and width z_0,

$$\tau \sim \frac{2\pi\hbar}{E_0}\exp\left(\frac{z_0}{a_0}\sqrt{\frac{E_0}{\mathrm{Ry}}}\right).$$

For typical values used in our calculations, τ reaches the minimum time necessary for experimental detection of an anion (10^{-3} s) at $z_0 > 60a_0$, i.e., at mesoscopic distances, where its interaction with the surface is negligibly small. Therefore, it is impossible to observe an anion in the instability region: the electron detachment is very fast.

It is clear that the higher the value of l is at which set of equations (10) is truncated, the higher the accuracy of the calculated electron energy. Because (5) is an exact solution of (10), the sequence of approximations $E_s^{(0)}$, $E_s^{(1)}$, $E_s^{(2)}$, … converges to the exact energy E_s. From the standpoint of the variational principle, the higher the approximation order, the closer the "trial" wave function

$$\psi^{(k)}(r, \xi) = r^{-1}\sum_{l=0}^{k}\varphi_l(r)P_l(\xi)$$

is to the exact solution for which the energy reaches the minimum. Therefore,

$$E_s \le E_s^{(k+1)} \le E_s^{(k)},$$

and the accuracy of $E_s^{(k+1)}$ can be roughly estimated as $\left|E_s^{(k+1)} - E_s^{(k)}\right|$. In our calculations, this accuracy depends primarily on the relative energy shift $(E_s - E_0)/E_0$, and the accuracy is better than 4% for an attractive surface. If necessary, higher values of l can be included. The iteration method discussed above can also be applied for the solution of the corresponding set of equations.

It is interesting to note that for truncated set of equations (17) and (18) for an attractive potential, a stable solution disappears at some threshold value of E_s somewhat higher than V_0; that is, the limit condition $E_s = V_0$ does not automatically hold. However, the threshold value of E_s converges to V_0 as l increases. Thus, for O_2^- in the zero approximation, this threshold value is $E_s \approx 1.4$ eV; in the first approximation, $E_s \approx 0.8$ eV; the exact value is 0.52 eV.

One can expect that the accuracy of the method described in this paper is very high and that it is compatible with the accuracy of spectroscopic measurements. Thus, calculation results could be directly compared with, e.g., measurements of electron photodetachment threshold shifts.

We now discuss possible experimental realization of the effects proposed in this paper. Layers of noble gases are frequently used as simple model surfaces to study the effect of a condensed environment [25]. In a typical experiment, O_2^- anions are deposited on a metal surface covered with approximately ten monolayers of a noble gas (krypton). Because the energy of the electron inside the solid (relative to the vacuum) can range from –2 to –1 eV, the case of attraction would be realized in such an experiment. The case of repulsion could occur when the metal is coated with a polyethylene film. If this film is sufficiently thick, the energy of the electron inside the solid can vary from 0.5 to 1 eV. The problems of possible experimental investigation are also generally related to the photochemistry of adsorbed molecules via phototransfer of substrate electrons to the adsorbate molecules. Such problems can be treated experimentally by charging experiments.

We hope that the solution of the 2D Schrödinger equation proposed in this paper is sufficiently universal to be used in other applications where the effect of the wave function nonsphericity is not negligibly small.

A more detailed investigation of the stability of anions requires solution of the time-dependent Schrödinger equation, because electron detachment occurs as a tunnel process. In addition, the effect of the image force has to be included more accurately in the calculation of the overall interaction potential.

This paper was supported by a joint program of the Deutsche Forschungsgemeinschaft (DFG, grant no. 436 RUS 113/433/0-2(R)) and the Russian Foundation for Basic Research (project no. 01-02-04008).

REFERENCES

1. R. Azria, I. Le Coat, J.-P. Ziesel, *et al.*, Chem. Phys. Lett. **220**, 417 (1994).
2. K. Nagesha and L. Sanche, Phys. Rev. Lett. **78**, 4725 (1997).
3. F. Weik and E. Illenberger, J. Chem. Phys. **109**, 6079 (1998).
4. R. Palmer and J. P. Rous, Rev. Mod. Phys. **64**, 383 (1992).
5. J. W. Gadzuk, L. J. Richter, S. A. Buntin, *et al.*, Surf. Sci. **235**, 317 (1990).
6. St. J. Dixon-Warren, E. T. Jensen, and J. C. Polanyi, J. Chem. Phys. **98**, 5938 (1993).
7. T. Hertel, M. Wolf, and G. Ertl, J. Chem. Phys. **102**, 3414 (1995).
8. L. Sanche, J. Phys. B: At. Mol. Opt. Phys. **23**, 1597 (1990).
9. O. Ingólfsson, F. Weik, and E. Illenberger, Int. J. Mass Spectrom. Ion Processes **155**, 1 (1996).
10. F. Weik, E. Illenberger, K. Nagesha, and L. Sanche, J. Phys. Chem. B **102**, 824 (1998).
11. M. A. Huels, L. Parenteau, and L. Sanche, Phys. Rev. B **52**, 11343 (1995).
12. M. Michaud and L. Sanche, J. Electron Spectrosc. Relat. Phenom. **51**, 237 (1990).
13. P. Rowntree, L. Parenteau, and L. Sanche, Chem. Phys. Lett. **182**, 479 (1991).
14. E. Illenberger, in *Photoionization and Photodetachment*, Ed. by Cheuk–Yiu Ng (World Sci., Singapore, 2000), Part 2, Advanced Series in Physical Chemistry, Vol. 10B, p. 1063.
15. A. G. Khrapak, P. Tegeder, E. Illenberger, and W. F. Schmidt, Chem. Phys. Lett. **310**, 557 (1999).
16. J. R. Hiskes, A. Karo, and M. Gardner, J. Appl. Phys. **47**, 3888 (1976).
17. J. W. Gadzuk and S. Holloway, J. Chem. Phys. **84**, 3502 (1986).
18. S. Holloway, J. Vac. Sci. Technol. A **5**, 476 (1987).
19. R. T. Taha and M. J. Ablowitz, J. Comput. Phys. **55**, 203 (1984).
20. O. Bang, P. L. Christiansen, F. If, *et al.*, Appl. Anal. **57**, 3 (1995).
21. D. C. Marinica, C. Ramseyer, A. G. Borisov, *et al.*, Phys. Rev. Lett. **89**, 046802-1 (2002).
22. A. G. Khrapak, Teplofiz. Vys. Temp. **13**, 858 (1975).
23. A. A. Radtsig and B. M. Smirnov, *Reference Data on Atoms, Molecules, and Ions* (Atomizdat, Moscow, 1980; Springer, Berlin, 1985).
24. P. D. Grigor'ev and A. M. Dyugaev, Zh. Éksp. Teor. Fiz. **115**, 593 (1999) [JETP **88**, 325 (1999)].
25. K. Nagesha, I. I. Fabrikant, and L. Sanche, J. Chem. Phys. **114**, 4934 (2001).

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 615–623.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 680–690.

SOLIDS
Electronic Properties

On a Quantum Plateau of Magnetization in Metal–Organic Quasi-One-Dimensional Ferrimagnets

I. G. Bostrem, A. S. Boyarchenkov, A. A. Konovalov, A. S. Ovchinnikov*, and V. E. Sinitsyn

Ural State University, pr. Lenina 51, Yekaterinburg, 620083 Russia
**e-mail: alexander.ovchinnikov@usu.ru*

Received September 17, 2002

Abstract—The possibility of realizing a quantum plateau of magnetization in [Mn(*hfac*)$_2$BNO$_R$] metal–organic compounds is investigated theoretically. A model of a one-dimensional ferrimagnetic chain (5/2, 1) is used for calculating the magnetization as a function of an external field by the method of discrete path integral representation (DPIR). Within this model, the coexistence of classical and quantum plateaus of magnetization is revealed. It is shown that the critical field H_{c1} that destroys the classical plateau (ground-state magnetization) is determined by the optical gap in zero field, which is estimated by the matrix-product method and a numerical method of exact diagonalization (recursion method). © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

The quantization of magnetization is one of the most interesting phenomena in the physics of low-dimensional magnetic materials. Magnetization plateaus were observed in strong magnetic fields in many quasi-one-dimensional compounds: in spin chains with spin $S = 1$ [1], organic spin ladders [2], etc.

In [4], the authors applied the Lieb–Schultz–Mattis theorem [3] to derive a general condition for the existence of a magnetization plateau,

$$\tilde{S} - m \in \mathbb{Z}, \tag{1}$$

where $\tilde{S}$ and m are the sum of spins over all sites of a unit cell and the magnetization per period (magnetization of a unit cell), respectively.

Ferrimagnetic spin chains are of special interest among the above-mentioned systems. Real prototypes of such compounds include MCu(*pba*OH)$(H_2O)_3 \cdot 2H_2O$ bimetal chains [5], where M = Mn ($S = 5/2$), Fe ($S = 2$), Co ($S = 3/2$), and Ni ($S = 1$). Methods of modern synthesis have made it possible to realize a new class of quasi-one-dimensional ferrimagnets—metal–organic compounds with the general formula [Mn(*hfac*)$_2$BNO$_R$] (R = H, F, Cl, Br) [6].

One-dimensional ferrimagnets, including metal–organic compounds, exhibit many interesting thermodynamic properties. Due to the existence of an optical gap in the excitation spectrum, these materials are similar to antiferromagnets with integer spin at high temperatures, whereas, at low temperatures, they exhibit properties characteristic of ferromagnets due to the existence of spontaneous magnetization [7, 8].

The quantization of magnetization for one-dimensional ferrimagnets has been intensively studied via an example of model (1, 1/2). In this model, the magnetization curve of the ground state has a plateau. However, the existence of this plateau may have a classical explanation due to the presence of an optical gap in the spectrum of elementary oscillations (classical plateau) [9]. In [7], it was pointed out that quantum ferrimagnets consisting of two types of spins (S, s) each greater than 1/2 may exhibit other kinds of plateaus (quantum plateaus) with the magnetization

$$m = S - s + 1, S - s + 2, \ldots, S + s - 1.$$

The origin of these plateaus cannot be accounted for either in the classical Ising model or in the classical Heisenberg model. Bimetal compounds of the type MCu(*pba*OH)$(H_2O)_3 \cdot 2H_2O$ cannot be used in searching for quantum plateaus since the copper spin S(Cu) is equal to 1/2. Metal–organic compounds are more promising in this respect.

To describe the magnetic properties of [Mn(*hfac*)$_2$BNO$_R$] (Fig. 1), one can apply the model of a ferrimagnet consisting of two types of spins (5/2, 1). This approximation is justified when the ferromagnetic interaction between the spins of NO groups ($s = 1/2$) is sufficiently strong so that one can consider these spins as a single spin $s = 1$ [10].

According to condition (1), a plateau of magnetization M is realized under the condition that $S_{\max}(1 - M) \in \mathbb{Z}$, where $S_{\max}$ is the maximal value of spin per unit cell ($S_{\max} = 7/2$), and the magnetization M is normalized by its saturation value ±1. Thus, one can expect that the magnetization has a plateau at $M = 3/7$, which corresponds to the ground state (a classical plateau) at $M =$

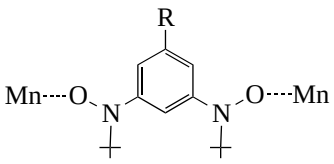

Fig. 1. The structure of the spin chain of [Mn($hfac$)$_2$BNO$_R$] compound.

5/7 (a partially polarized state or a quantum plateau), and at $M = 1$ (a completely polarized state).

The aim of this paper is to investigate the possibility of realizing a quantum plateau of magnetization within the model of a one-dimensional quantum ferrimagnet (5/2, 1). To take into consideration strongly developed quantum fluctuations when calculating magnetization, in Section 2 we apply the method of discrete path integral representation (DPIR) [11]. This method allows us to transform the original one-dimensional quantum spin Hamiltonian into a classical two-dimensional one according to the generalized Trotter formula [12], which is used for numerical investigations by the Monte Carlo method. In [7], it was demonstrated that it is the classical scenario (spin-wave excitations against the background of Néel ordering), rather than excitations, that destroys singlet bonds (valence-band states (VBSs)) [13], which is more appropriate for explaining the nature of the initial plateau. From the viewpoint of spin-wave theory, the ground-state magnetization per unit cell ($m = 3/2$) as a function of applied field h increases by one when the field value reaches the antiferromagnetic excitation gap Δ. Therefore, in Sections 3 and 4, we calculate this quantity by the recursion method [14, 15] and the matrix-product method [16].

2. CALCULATION OF MAGNETIZATION

A one-dimensional quantum ferrimagnetic chain consisting of two types of spins $S = 5/2$ and $s = 1$ and placed in an external magnetic field h is described by the Heisenberg Hamiltonian

$$H = J\sum_{i=1}^{N}[(\mathbf{S}_i\cdot\mathbf{s}_i)+(\mathbf{s}_i\cdot\mathbf{S}_{i+1})]+h\sum_{i=1}^{N}[S_i^z+s_i^z], \quad (2)$$

where N is the total number of unit cells in the system and $J > 0$ is the integral of antiferromagnetic exchange interaction. Concerning Hamiltonian (2), note that the basic feature of one-dimensional ferrimagnets is the presence of long-range ordering at zero temperature, which manifests itself under strongly developed quantum fluctuations. As applied to these systems, a spin-wave description of the properties of the ground state gives sufficiently good results, which, however, can be refined by the variational matrix-product method, which takes into account the formation of VBSs [10]. This fact specifies the difference between one-dimensional ferrimagnets and antiferromagnets in which fluctuations destroy long-range ordering so that the spin-wave theory becomes inapplicable. From a theoretical viewpoint, for antiferromagnets with integer spin, first of all, with $S = 1$, it is more convenient to consider a generalization of the Heisenberg Hamiltonian that takes into account biquadratic exchange terms,

$$H = \sum_i[(\mathbf{S}_i\cdot\mathbf{S}_{i+1})-\beta(\mathbf{S}_i\cdot\mathbf{S}_{i+1})^2]. \quad (3)$$

Although less realistic, this consideration explicitly demonstrates the possibility of realizing, as the ground state of an antiferromagnet, a VBS with singlet bonds (for example, an Affleck–Kennedy–Lieb–Tasaki (AKLT) model with $\beta = -1/3$) that does not break the translational symmetry and is characterized by the exponential behavior of correlation functions and an energy gap above the ground state. Moreover, for $\beta = 1$, such a model can be solved exactly [17]. A qualitative picture of the evolution of the ground state of an antiferromagnetic model with spin $S = 1$ and biquadratic exchange terms, proposed by Affleck [18], was confirmed in numerical analysis with the Monte Carlo method [19], by calculating the spectrum of low-energy excitations with the matrix-product method [20], and by the magnetization curves that exhibit a cusp singularity in moderate fields [21].

Since one does not face the above-mentioned fundamental difficulties when studying the properties of the ground state of a ferrimagnetic chain, the description of this chain within the Heisenberg Hamiltonian proves to be acceptable and realistic.

Note that the decomposition of interactions into intrablock ($H_{i,i}$) and interblock ($H_{i,i+1}$) ones is artificial. Therefore, for further analysis, it is convenient to rewrite the original Hamiltonian as

$$H = \sum_{i=1}^{2N}\frac{1}{2}(s_i^+S_{1+1}^- + s_i^-S_{i+1}^+)+s_i^zS_{i+1}^z$$
$$+\frac{h}{2}(s_i^z+S_{i+1}^z) = \sum_{i=1}^{2N}H_{i,i+1},$$

where the original Hamiltonian (2) is represented as a sum of "two-particle" operators $H_{i,i+1}$. When Hamiltonian (2) becomes a sum of operators $H_{i,i+1}$, two such operators are contained in a single unit cell.

Let us estimate the statistical sum of the original system:

$$Z = \mathrm{Tr}(e^{-\beta H}) = \sum_u\langle u|\exp(-\beta H)|u\rangle,$$

where $\beta = 1/k_B T$. Since the Hamiltonian cannot be exactly diagonalized, we apply the so-called discrete path integral representation (DPIR). The basic idea of this method is to convert a quantum spin of each site into a p-component vector $\mathbf{u} = (u^1, u^2, \ldots, u^p)$, where $p \longrightarrow \infty$. Each component u^t ($t = 1, 2, \ldots, p$) is a classical spin variable—a set of eigenvalues of appropriate operators, i.e., $u = 0, \pm 1$ for s^z and $u = \pm 5/2, \pm 3/2, \pm 1/2$ for S^z. The introduction of an additional dimension defined by the parameter p represents quantum indeterminacy by producing replicas of the original variable. The quantum spin Hamiltonian can be transformed into the classical one; then,

$$Z = \sum_u \langle u| \left[\prod_{i=1}^{2N} \exp\left((-\beta)\frac{H_{i,i+1}}{p}\right)\right]^p |u\rangle$$

$$= \sum_{u^{(1)}}\sum_{u^{(2)}}\cdots\sum_{u^{(p)}}\prod_{t=1}^{p}\prod_{i=1}^{2N}\langle u_i^{(t)} u_{i+1}^{(t)}|$$

$$\times \exp\left(-\beta\frac{H_{i,i+1}}{p}\right)|u_i^{(t+1)} u_{i+1}^{(t+1)}\rangle.$$

The matrix

$$Q = \langle u_i^{(t)} u_{i+1}^{(t)}| \exp\left(-\beta\frac{H_{i,i+1}}{p}\right)|u_i^{(t+1)} u_{i+1}^{(t+1)}\rangle$$

is decomposed into subblocks that relate the states with identical full projections $u_i^{(t)} + u_{i+1}^{(t)}$ per unit cell. Calculating the eigenvalues of these matrices, in the limit as $p \longrightarrow \infty$, we obtain

$$\lambda_1 = 1 + 7sz - 5yz, \quad \lambda_2 = 1 - 7sz - 5yz,$$

$$u_i^{(t)} + u_{i+1}^{(t)} = \pm\frac{7}{2},$$

$$\lambda_3 = 1 - 5sz + 5yz, \quad \lambda_4 = 1 - 5sz + 2yz,$$

$$u_i^{(t)} + u_{i+1}^{(t)} = \frac{5}{2},$$

$$\lambda_5 = 1 + 5sz - 5yz, \quad \lambda_6 = 1 + 5sz + 2yz,$$

$$u_i^{(t)} + u_{i+1}^{(t)} = -\frac{5}{2},$$

$$\lambda_7 = 1 - 3sz - 5yz, \quad \lambda_8 = 1 - 3sz + 2yz,$$

$$\lambda_9 = 1 - 3sz + 7yz, \quad u_i^{(t)} + u_{i+1}^{(t)} = \frac{3}{2},$$

$$\lambda_{10} = 1 + 3sz - 5yz, \quad \lambda_{11} = 1 + 3sz + 2yz,$$

$$\lambda_{12} = 1 + 3sz + 7yz, \quad u_i^{(t)} + u_{i+1}^{(t)} = -\frac{3}{2},$$

$$\lambda_{13} = 1 - sz - 5yz, \quad \lambda_{14} = 1 - sz + 2yz,$$

$$\lambda_{15} = 1 - sz + 7yz, \quad u_i^{(t)} + u_{i+1}^{(t)} = \frac{1}{2},$$

$$\lambda_{16} = 1 + sz - 5yz, \quad \lambda_{17} = 1 + sz + 2yz,$$

$$\lambda_{18} = 1 + sz + 7yz, \quad u_i^{(t)} + u_{i+1}^{(t)} = -\frac{1}{2},$$

where $s = \beta h/4$, $y = \beta/2$, and $z = 1/p$. Since the original quantum model has been reduced to the classical two-dimensional model, one can apply the double-chain approximation (DCA) [22], where a two-dimensional system is represented by parallel double chains in an effective external field $\tilde{h}$. In this case, the statistical sum is calculated exactly along a "fictitious" direction p and, in the mean-field approximation, along a direction in a real space

$$Z = Z_{\mathrm{DCA}}^{2N}, \quad Z_{\mathrm{DCA}}^{2N} = \lim_{p\to\infty}\sum_{i=1}^{18}\lambda_i^p.$$

A direct calculation yields

$$Z_{\mathrm{DCA}} = \left(\exp\left(-\frac{5}{2T}\right) + \exp\left(\frac{1}{T}\right) + \exp\left(\frac{7}{2T}\right)\right)$$

$$\times\left(\cosh\left(\frac{\tilde{h}}{4T}\right) + \cosh\left(\frac{3\tilde{h}}{4T}\right)\right)$$

$$+\left(\exp\left(-\frac{5}{2T}\right) + \exp\left(\frac{1}{T}\right)\right)\cosh\left(\frac{5\tilde{h}}{4T}\right)$$

$$+ \exp\left(-\frac{5}{2T}\right)\cosh\left(\frac{7\tilde{h}}{4T}\right).$$

The free energy per unit cell is $f = F/N = -2T\ln Z_{\mathrm{DCA}}$. The magnetization per unit cell is given by the expression

$$m = -\frac{\partial f}{\partial h} = \frac{1}{2}[(\exp(-5/2T) + \exp(1/T) + \exp(7/2T))(\sinh(\tilde{h}/4T) + 3\sinh(3\tilde{h}/4T))$$

$$+ 5(\exp(-5/2T) + \exp(1/T))\sinh(5\tilde{h}/4T) + 7\exp(-5/2T)\sinh(7\tilde{h}/4T)]$$

$$\times[(\exp(-5/2T) + \exp(1/T) + \exp(7/2T))(\cosh(\tilde{h}/4T) + \cosh(3\tilde{h}/4T))$$

$$+ (\exp(-5/2T) + \exp(1/T))\cosh(5\tilde{h}/4T) + \exp(-5/2T)\sinh(7\tilde{h}/4T)]^{-1}, \quad (4)$$

where $\tilde{h} = h + m$.

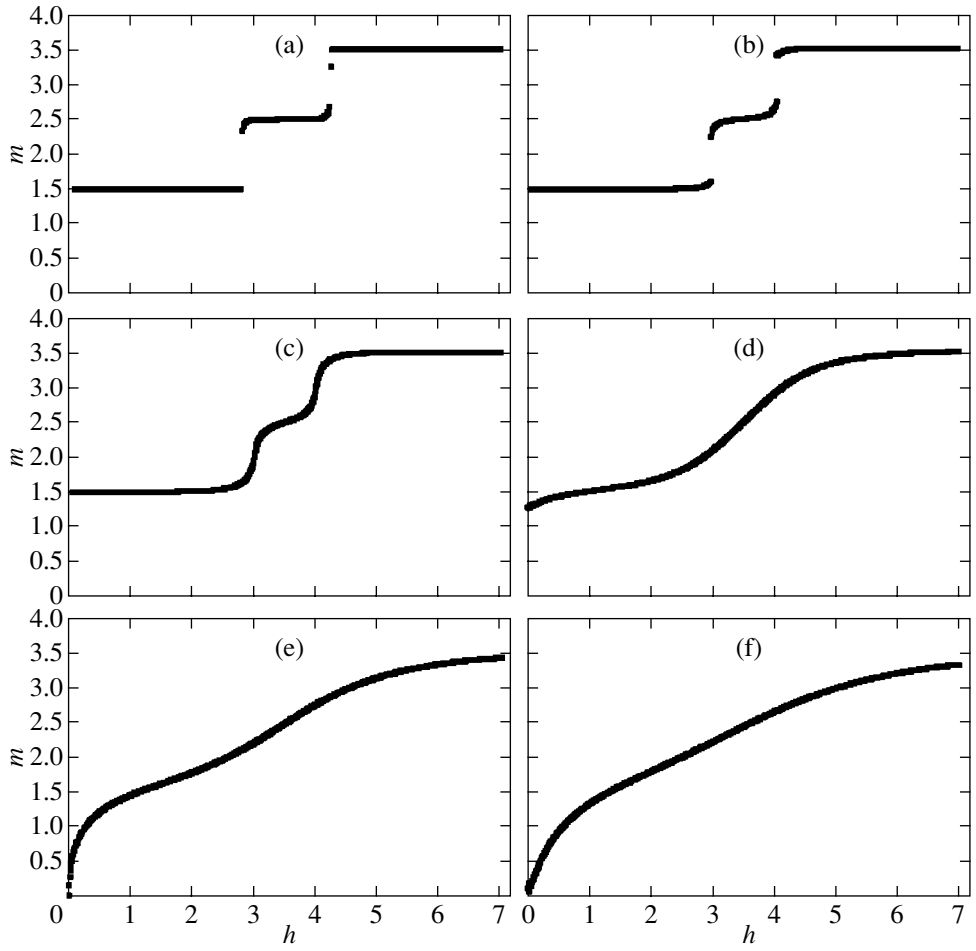

Fig. 2. The curves of magnetization m of the chain (5/2, 1) versus applied external field h for various temperatures T: (a) 0.05, (b) 0.1, (c) 0.15, (d) 0.4, (e) 0.7, and (f) 1.0J.

The dependence $m(h)$ calculated by formula (4) for different temperatures is shown in Fig. 2. As expected, the calculation demonstrates the coexistence of classical and quantum plateaus in this model.

One can propose the following qualitative interpretation for this dependence relying on the analysis of the ground state of mixed ferrimagnetic chains for the general case (S, s) [23] and the particular cases (3/2, 1) and (2, 1) [24], where each spin is greater than 1/2, as a function of the external field h, as well as using the analysis of the ground state of the chain (5/2, 1) [10]. The process of magnetization saturation corresponds to the scheme represented in Fig. 3. The initial plateau with $m = 3/2$ corresponds to the state with a broken translational symmetry with two singlet bonds and a noncompensated spin $S = 3/2$, which is responsible for the spin-wave "ferromagnetic" branch of excitations [8] (Fig. 3a). The initial plateau begins to destruct at a certain critical value H_{c1} of the field due to a break of one singlet bond in a block and the formation of a triplet. This fact gives rise to a partially polarized phase with one singlet bond that corresponds to the magnetization plateau with $m = 5/2$ (Fig. 3b). A transition to the saturation magnetization $m = 7/2$ takes place under full destruction of the remaining singlet bonds (Fig. 3c).

Such an evolution of the ground state accounts for the fact that the linear spin-wave theory provides a successful explanation for the properties of the ground state and low-energy excitations. A classical picture—spin-wave excitations above the Néel ordering—proves to be adequate [25]; at the same time, it fails to describe the optical mode of excitations, where local singlet–triplet excitations play the key role. The magnitude of the optical gap $\Delta = 3J$ proves to be strongly understated compared with the numerical result and the value obtained by the matrix-product method, which takes into account local quantum fluctuations. The optical excitation mode, called an antiferromagnetic mode in what follows, is formed in the same way as that in one-dimensional ferrimagnets with integer spin (in particular, for $S = 1$) with the Haldane gap, which are different from antiferromagnets of half-integer spin with gapless excitations and linear k dependence [20].

Thus, from a theoretical point of view, investigation of the magnetization curve is very important since the curve $m(h)$, while reflecting the structure of the ground state in finite fields, contains information about excited states in zero field.

3. NUMERICAL ANALYSIS

To obtain the acoustic and optical excitation branches in zero external field, we calculate the dynamic structural factor

$$S^{-+}(k, \omega) = \sum_n \left|\langle n|S_k^- + s_k^-|0\rangle\right|^2 \delta(\omega - (E_n - E_0)),$$

$$S^{+-}(k, \omega) = \sum_n \left|\langle n|S_k^+ + s_k^+|0\rangle\right|^2 \delta(\omega - (E_n - E_0))$$

for the ferro- and antiferromagnetic excitation branches, respectively. Here, $|n\rangle$ is the eigenstate of the Hamiltonian corresponding to energy E_n and E_0 is the energy of the ground state, which, according the Lieb–Mattis theorem [26], has the spin $(S - s)N$. To obtain this state, one has to diagonalize the original Hamiltonian (2) of a finite chain with periodic boundary conditions applying, for example, the restarted Lanczos method.

Since the total magnetization $M = \sum_j S_j^z + s_j^z$ is invariant, one can diagonalize Hamiltonian (2) in each subspace with fixed M. For instance, when dealing with the ground state of a chain of length $N = 7$ sites, instead of approximately 6.12×10^8 states of the total Hamiltonian, we have states of the subspace $N(S - s)$ with an approximate dimension of 5.54×10^6. The subspace corresponding to the acoustic mode of excitations has a magnetization of $M = N(S - s) - 1$, while the optical mode corresponds to the total magnetization of the chain $M = N(S - s) + 1$.

Let us express the dynamic structural factor in terms of the appropriate Green function as

$$S^{\sigma\bar{\sigma}}(k, \omega) = -\frac{1}{\pi}\mathrm{Im}G^{\sigma\bar{\sigma}}(k, \omega),$$

where $\sigma = \pm 1$ and $\bar{\sigma} = -\sigma$. This Green function is defined as a continued fraction,

$$G^{\sigma\bar{\sigma}}(k, \omega) = \frac{\langle 0|(S_{-k}^{\bar{\sigma}} + s_{-k}^{\bar{\sigma}})(s_k^{\sigma} + S_k^{\sigma})|0\rangle}{\omega - a_0 - \dfrac{b_1^2}{\omega - a_1 - \dfrac{b_2^2}{\omega - a_2 - \ldots}}},$$

where the coefficients $a_n = \langle f_n|H|f_n\rangle/\langle f_n|f_n\rangle$, $b_n^2 = \langle f_n|f_n\rangle/\langle f_{n-1}|f_{n-1}\rangle$, and $b_0 = 0$ are determined by the set of orthogonal states

$$|f_{n+1}\rangle = H|f_n\rangle - a_n|f_n\rangle - b_n^2|f_{n-1}\rangle,$$

$$|f_0\rangle = (s_k^{\sigma} + S_k^{\sigma})|0\rangle.$$

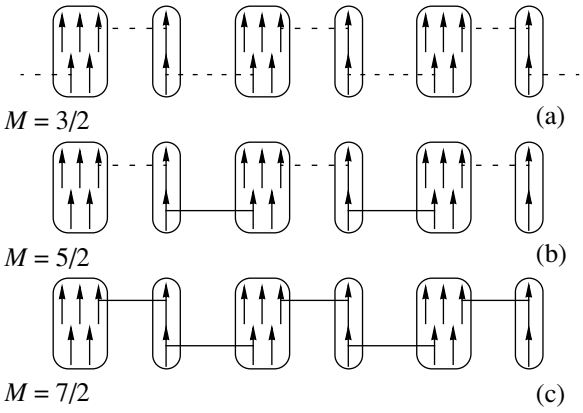

Fig. 3. Schematic representation of the ground state of the spin ferrimagnetic chain (5/2, 1) versus external magnetic field. The arrow denotes a spin of 1/2. The dashed (solid) lines correspond to the formation of a singlet (triplet) pair. The ovals represent the construction of spin 5/2 or 1 by symmetrizing five (one) spins inside them.

When implementing the method, the initial state is chosen as the lowest energy state with the quantum number $M = N(S - s)$. The Fourier images of spin operators are determined by the relations $S_k^{\pm} = N^{-1/2}\sum_j e^{ik(j-1/4)}S_j^{\pm}$ and $s_k^{\pm} = N^{-1/2}\sum_j e^{ik(j+1/4)}s_j^{\pm}$.

The energy of elementary excitations is determined from the position of the maximum of spectral intensity. The results obtained by the recursion method are displayed in Fig. 4 for selected values of the wave vector $k = 2\pi m/N$, $m = 0, 1, \ldots, n - 1$. The dots on the diagram are obtained on chains of length $N = 4, 5, 6, 7$.

As applied to the given system, the recursion method rapidly converges to the thermodynamic limit. The table presents the ground-state energy per unit cell as a function of the number N of unit cells. Such fast convergence is attributed to the small correlation length, which is on the order of the lattice constant; correlation lengths in various approximations were estimated in [10]. In numerical calculations, the shape of the density of states of spin excitations stabilizes sufficiently rapidly even for small values of n. The optical gap proves to be equal to $\Delta \approx 3.793J$, which is much

The ground-state energy per unit cell, E_0/JN, as a function of chain length N, calculated by the recursion method

N	3	4	5	6	7
E_0/J	–5.9281	–5.9161	–5.9141	–5.9135	–5.9133

greater than the result obtained by spin-wave theory but is in excellent agreement with the variant of this theory that takes into account the spin–wave interaction [27].

4. MATRIX-PRODUCT METHOD

The method of the density matrix renormalization group (DMRG) proposed by White [28] turns out to be very fruitful for investigating the properties of one-dimensional quantum systems. Based on the idea of renormgroup in a real space, the DMRG method allows one to construct quantum states of an infinite chain subject to periodic conditions in the form of special products (matrix products) composed of the states of a unit cell. The numerical implementation of the DMRG algorithm by an iterative method involves the construction of a projector from the states of a single cell to the basis of the new lowest energy effective states of the cell; the calculation of observables on these states yields the results of the thermodynamic limit. Later, it was shown in [20] that these states can be obtained via a simple variational ansatz without using the renormgroup procedure. This is the idea of the matrix-product method.

As applied to quantum ferrimagnets, this method has been applied mainly to the spin chain (1, 1/2). Unfortunately, this model does not illustrate the potentials of the matrix-product method because there are a relatively small number of quantum states in a magnetic cell. This applies, for example, to the choice of the number of effective states. Therefore, we first consider in detail the calculation of the ground-state energy of the spin chain (5/2, 1) by using matrix products of different dimensions.

4.1. The Ground-State Energy

The translation-invariant wave function of the ground state of a chain of length N with cyclic boundary conditions can be represented as [16]

$$|\Omega\rangle = \mathrm{Tr}(\hat{g}_1^{jj}\hat{g}_2^{jj}\ldots\hat{g}_N^{jj}).$$

The site matrices $\hat{g}_i^{jj}$ are constructed as follows:

$$\hat{g}_i^{jj} = \sum_k C_j^{ks}\sum_{q,M}\begin{bmatrix} k & s & j \\ q & M & j\end{bmatrix}\hat{X}_q^k|SM\rangle_i,$$

where [...] are the Clebsch–Gordan coefficients, $|SM\rangle_i$ is the spin function of the ith magnetic cell, and $\hat{X}_q^k = (\hat{T}_q^k)^\dagger$ are basis matrices. The matrices $\hat{T}_q^k$ represent a matrix representation of irreducible tensor operators of the rotation group of rank k. The quantity j is the "hyperspin" of the ith magnetic cell in the "expanded" space, which represents a direct product of the matrix space and the Hilbert space of the cell. The value of j should be specified when constructing $|\Omega\rangle$. If all the cells are assumed to be identical, then $j = S - s$, or 3/2 as in our case. The variational coefficients C_j^{ks} are determined from the condition that the ground-state energy is minimal. The wave functions of the magnetic cell are constructed in a standard way:

$$|SM\rangle \equiv \left|\frac{5}{2}1SM\right\rangle = \sum_{M_1M_2}\begin{bmatrix} 5/2 & 1 & S \\ M_1 & M_2 & M\end{bmatrix}\left|\frac{5}{2}M_1\right\rangle|1M_2\rangle,$$

where possible values of S are 3/2, 5/2, and 7/2. The problem of choosing the matrix basis $\hat{X}_q^k$ does not have a unique solution. The rank of the basis matrices of dimension $(N+1)\times(N+1)$ may take values from 0 to N. Apparently, an optimal size of matrices should be specified for each particular problem. We have calculated the ground-state energy for the spin chain (5/2, 1) by using the basis of matrices of dimensions 2×2 and 3×3. On the basis of 2×2 matrices [10], we have

$$\hat{g}^{\frac{3}{2}\frac{3}{2}} = \begin{pmatrix} \left(u-\sqrt{\frac{3}{5}}v\right)\left|\frac{3}{2}\frac{3}{2}\right\rangle - \frac{1}{\sqrt{15}}w\left|\frac{5}{2}\frac{3}{2}\right\rangle & \frac{2}{\sqrt{3}}w\left|\frac{5}{2}\frac{5}{2}\right\rangle \\ -\frac{2}{\sqrt{5}}v\left|\frac{3}{2}\frac{1}{2}\right\rangle - \sqrt{\frac{2}{15}}w\left|\frac{5}{2}\frac{1}{2}\right\rangle & \left(u+\sqrt{\frac{3}{5}}v\right)\left|\frac{3}{2}\frac{3}{2}\right\rangle + \frac{2}{\sqrt{15}}w\left|\frac{5}{2}\frac{3}{2}\right\rangle \end{pmatrix}, \quad (5)$$

where the following notations are used: $u = C_{3/2}^{0,3/2}$, $v = C_{3/2}^{1,3/2}$, and $w = C_{3/2}^{1,5/2}$. For such a choice of the basis, we take into account only the states $\left|\frac{3}{2}M\right\rangle$ of the magnetic cell and the lowest state $\left|\frac{5}{2}M\right\rangle$. Introducing a transfer matrix $\hat{G} = \left(\hat{g}^{\frac{3}{2}\frac{3}{2}}\right)^\dagger \otimes \hat{g}^{\frac{3}{2}\frac{3}{2}}$, we can represent the ground-state energy per unit cell as

$$\frac{E_0}{JN} = \frac{\mathrm{Tr}\left[\left(\hat{g}^{\frac{3}{2}\frac{3}{2}} \otimes \mathbf{S}_1\mathbf{S}_2\hat{g}^{\frac{3}{2}\frac{3}{2}}\right)\hat{G}^{N-1}\right] + \mathrm{Tr}\left[\left(\hat{g}^{\frac{3}{2}\frac{3}{2}} \otimes \mathbf{S}_2\hat{g}^{\frac{3}{2}\frac{3}{2}}\right)\left(\hat{g}^{\frac{3}{2}\frac{3}{2}} \otimes \mathbf{S}_1\hat{g}^{\frac{3}{2}\frac{3}{2}}\right)\hat{G}^{N-2}\right]}{\mathrm{Tr}[\hat{G}^N]}, \quad (6)$$

where the symbol $\otimes$ stands for the external product of matrices. Since formula (6) contains the powers of the matrix $\hat{G}$, it is convenient to reduce this matrix to the Jordan form ($\hat{G}$ is nonsymmetric). Fortunately, the Jordan form of $\hat{G}$ proves to be diagonal; this fact substantially simplifies the calculations and allows us to represent the ground-state energy in the thermodynamic limit $N \longrightarrow \infty$ in an analytic form. Formula (6) also shows that the number of independent variable parameters equals two. The numerical minimization of the expression for the energy by the simplex method yields the following result:

$$E_0/N \approx -5.806J$$

with variational coefficients

$$\frac{v}{u} = -0.5798, \quad \frac{w}{u} = -0.9131.$$

This value of energy proves to be somewhat greater than that obtained by the recursion method. On the basis of 3×3 matrices, the matrix $\hat{g}$ is expressed as

$$g^{\frac{3}{2}\frac{3}{2}} = \begin{pmatrix} \left(1 - \sqrt{\frac{3}{10}}\alpha + \frac{1}{\sqrt{30}}\gamma\right)\left|\frac{3}{2}\frac{3}{2}\right\rangle & \left(-\frac{1}{\sqrt{5}}\alpha + \frac{1}{\sqrt{5}}\gamma\right)\left|\frac{3}{2}\frac{1}{2}\right\rangle & \sqrt{\frac{2}{5}}\gamma\left|\frac{3}{2}-\frac{1}{2}\right\rangle + \frac{2}{\sqrt{35}}\delta\left|\frac{5}{2}-\frac{1}{2}\right\rangle \\ +\left(-\sqrt{\frac{2}{15}}\beta + \sqrt{\frac{2}{35}}\delta\right)\left|\frac{5}{2}\frac{3}{2}\right\rangle + \frac{1}{\sqrt{42}}\eta\left|\frac{7}{2}\frac{3}{2}\right\rangle & +\left(-\frac{1}{\sqrt{10}}\beta + \frac{3}{\sqrt{70}}\delta\right)\left|\frac{5}{2}\frac{1}{2}\right\rangle + \frac{1}{\sqrt{35}}\eta\left|\frac{7}{2}\frac{1}{2}\right\rangle & +\frac{1}{\sqrt{70}}\eta\left|\frac{7}{2}-\frac{1}{2}\right\rangle \\ \left(\frac{1}{\sqrt{3}}\beta - \frac{1}{\sqrt{7}}\delta\right)\left|\frac{5}{2}\frac{5}{2}\right\rangle - \frac{1}{\sqrt{7}}\eta\left|\frac{7}{2}\frac{5}{2}\right\rangle & \left(1 + \frac{2}{\sqrt{30}}\gamma\right)\left|\frac{3}{2}\frac{3}{2}\right\rangle & \left(-\frac{1}{\sqrt{5}}\alpha - \frac{1}{\sqrt{5}}\gamma\right)\left|\frac{3}{2}\frac{1}{2}\right\rangle \\ & -\frac{2\sqrt{2}}{\sqrt{35}}\delta\left|\frac{5}{2}\frac{3}{2}\right\rangle - \frac{2}{\sqrt{42}}\eta\left|\frac{7}{2}\frac{3}{2}\right\rangle & +\left(-\frac{1}{\sqrt{10}}\beta - \frac{3}{\sqrt{70}}\delta\right)\left|\frac{5}{2}\frac{1}{2}\right\rangle - \frac{1}{\sqrt{35}}\eta\left|\frac{7}{2}\frac{1}{2}\right\rangle \\ \frac{1}{2}\eta\left|\frac{7}{2}\frac{7}{2}\right\rangle & \left(\frac{1}{\sqrt{3}}\beta + \frac{1}{\sqrt{7}}\delta\right)\left|\frac{5}{2}\frac{5}{2}\right\rangle & \left(1 + \sqrt{\frac{3}{10}}\alpha + \frac{1}{\sqrt{30}}\gamma\right)\left|\frac{3}{2}\frac{3}{2}\right\rangle \\ & +\frac{1}{\sqrt{7}}\eta\left|\frac{7}{2}\frac{5}{2}\right\rangle & +\left(\sqrt{\frac{2}{15}}\beta + \sqrt{\frac{2}{35}}\delta\right)\left|\frac{5}{2}\frac{3}{2}\right\rangle + \frac{1}{\sqrt{42}}\eta\left|\frac{7}{2}\frac{3}{2}\right\rangle \end{pmatrix}.$$

Although this matrix contains a much greater number of variable parameters

$$\alpha = \frac{C_{3/2}^{1,3/2}}{C_{3/2}^{0,3/2}}, \quad \beta = \frac{C_{3/2}^{2,3/2}}{C_{3/2}^{0,3/2}}, \quad \gamma = \frac{C_{3/2}^{1,5/2}}{C_{3/2}^{0,3/2}},$$

$$\delta = \frac{C_{3/2}^{2,5/2}}{C_{3/2}^{0,3/2}}, \quad \eta = \frac{C_{3/2}^{2,7/2}}{C_{3/2}^{0,3/2}},$$

the expression for $g^{\frac{3}{2}\frac{3}{2}}$ contains the whole family of wave functions of the magnetic cell.

An attempt to calculate the ground-state energy by reducing the matrices $\hat{G}$ to the Jordan form gives an incorrect result in this case. The analysis of the matrix $\hat{G}$ shows that it has two nearly coinciding eigenvalues; therefore, the transfer matrix that reduces $\hat{G}$ to the Jordan form proves to be degenerate. An alternative is a direct calculation by formula (6) for a chain of finite length ($N \approx 200$) without reducing matrix $\hat{G}$ to the diagonal form. This yields the value of the ground-state energy close to that obtained by the recursion method,

$$E_0/N \approx -5.903J.$$

The values of the variable coefficients are as follows:

$$\alpha \approx 1.90, \quad \beta \approx 1.32, \quad \gamma \approx 0.84, \quad \delta \approx 0.12, \quad \eta \approx 0.72.$$

Note that the contribution of the state $\left|\frac{5}{2}1\frac{7}{2}M\right\rangle$ cannot be assumed negligible. We have also calculated the ground-state energy on the basis of 3×3 matrices with-

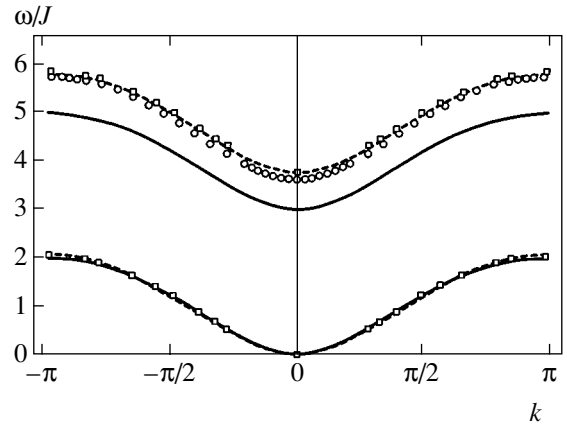

Fig. 4. Dispersion curves of the ferrimagnetic and antiferromagnetic branches of elementary excitations. The results obtained by spin-wave theory without taking into account the interaction of spin waves [25] and with regard to their interaction [27] are shown by solid and dashed lines, respectively. Calculation by the matrix-product method on the basis of 2×2 matrices is illustrated by circles, while the squares represent the results obtained by the recursion method.

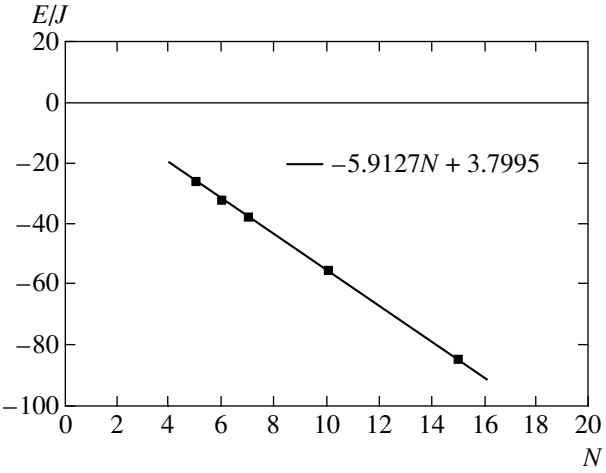

Fig. 5. Total energy of the spin chain (5/2, 1) with excitation $E_{k=0}$ as a function of its length N calculated by the matrix-product method on the basis of 3×3 matrices. The solid curve represents a linear approximation of the results obtained.

out taking into account the state $\left|\frac{5}{2}1\frac{7}{2}M\right\rangle$ (the coefficient η was assumed to be zero; the minimization was carried out with respect to the remaining parameters). The value thus obtained is

$$E_0/N \approx -5.840J,$$

which is closer to the value obtained on the basis of 2×2 matrices. One can conclude that the dimension of the basis matrices must be so that it allows one to take into account the whole family of wave functions of the magnetic cell. For the chains (1, 1/2) that have been intensively studied in the literature, the basis of matrices 2×2 is quite sufficient. For the chain (5/2, 1) considered in this paper, the choice of such a basis makes it possible to obtain quite reasonable results. The use of the extended 3×3 basis allows one to obtain the ground-state energy with an accuracy comparable to that of numerical calculations, although at the expense of considerable complication of calculations.

4.2. Calculation of the Optical Gap

The simplest form of the wave function of an optical magnon with momentum k was proposed in [29]:

$$|k\rangle = \sum_n e^{ikn}|n\rangle,$$

$$|n\rangle = \mathrm{Tr}\{g_1 \ldots g_{n-1}\tilde{g}_n g_{n+1} \ldots g_N\}.$$

The matrix $\tilde{g}$ is constructed for the state of a magnetic cell with the hyperspin $j + 1$ ($\frac{5}{2}$ in our case).

The total energy of the chain with excitation is given by the expression

$$E_k = \frac{\langle k|\hat{H}|k\rangle}{\langle k|k\rangle}, \quad (7)$$

and the value of the optical gap is given by

$$w(0) = E_{k=0} - NE_0, \quad (8)$$

where E_0 is the ground-state energy. A detailed account of the calculation technique in the one-magnon approximation is given in [20]. Our calculation on the basis of 2×2 matrices yields the following value for the optical gap: $w(0) \approx 3.61527$ (in the thermodynamic limit $N \longrightarrow \infty$).

The calculation of the antiferromagnetic gap on the basis of 3×3 matrices has been carried out on chains with the lengths ranging from 5 to 15 unit cells. The results obtained are presented in Fig. 5. The linear dependence of the total energy $E_{k=0}$ on N, which follows from (7), is clearly displayed even for chains of length greater than five sites. Application of the least-squares method to the linear dependence yields the following value for the optical gap:

$$w(0) \approx 3.7995J,$$

which coincides with the result of the recursion method up to three decimal places.

5. CONCLUSIONS

As predicted in many publications, the analysis carried out has confirmed that metal–organic compounds

are suitable objects for searching a quantum plateau of magnetization. The model of a quantum ferrimagnetic chain (5/2, 1) proposed for describing such materials reveals the coexistence of classical and quantum plateaus. The value of the first critical field H_{c1} (at which a transition form the ground-state plateau to the plateau with magnetization 5/2 occurs) obtained by the discrete path integral representation in the double-chain approximation is closer to the result of linear spin-wave theory ($3J$) and smaller than the values of the optical gap obtained by the recursion method ($3.793J$) and the method of matrix products ($3.7995J$). The values of the latter quantities are closer to the center of the quantum plateau (about 3.5–3.6J).

In conclusion, note that the fields that enable one to observe quantum plateaus of magnetization are sufficiently high, on the order of the exchange integral, and amount to about 10^6–10^7 Oe.

ACKNOWLEDGEMENTS

We are grateful to the Department of Parallel Computations, Institute of Mathematics and Mechanics, Ural Division, Russian Academy of Sciences, and personally to S. Sharf for affording us the opportunity to carry out numerical calculations. We also thank N.B. Baranov and K. Inoue for discussing a number of questions considered in this paper.

This work was supported by the program "Universities of Russia," grant UR.01.01.005. One of the authors (V.E.S.) gratefully acknowledges the support from the Noncommercial Program Foundation "Dynasty" and from the International Center for Basic Physics (Moscow).

REFERENCES

1. M. Hagiwara, K. Koboyashi, and T. Chihara, J. Phys. Soc. Jpn. **66**, 1702 (1997).
2. K. Katoh, Y. Hosokoshi, K. Inoue, and T. Goto, J. Phys. Soc. Jpn. **66**, 1702 (1997); T. Goto, M. I. Bartashevich, Y. Hosokoshi, *et al.*, Physica B (Amsterdam) **294**–**295**, 43 (2001).
3. E. Lieb, T. Schultz, and D. Mattis, Ann. Phys. **16**, 407 (1961).
4. M. Oshikawa, M. Yamanaka, and I. Affleck, Phys. Rev. Lett. **78**, 1984 (1997).
5. D. J. Van Koningsbruyjhnggen, O. Kahn, K. Nakatani, *et al.*, Inorg. Chem. **29**, 3325 (1990).
6. K. Inoue, F. Iwahori, A. S. Markosyan, and H. Iwamura, Coord. Chem. Rev. **198**, 219 (2000).
7. S. Yamamoto, Phys. Rev. B **61**, 842 (2000).
8. A. S. Ovchinnikov, I. G. Bostrem, V. E. Sinitsyn, *et al.*, J. Phys.: Condens. Matter **14**, 6188 (2002).
9. T. Sakai and S. Yamamoto, J. Phys.: Condens. Matter **12**, 9787 (2000).
10. A. S. Ovchinnikov, I. G. Bostrem, V. E. Sinitsin, *et al.*, J. Phys.: Condens. Matter **13**, 5221 (2001).
11. R. M. Stratt, Phys. Rev. B **33**, 1921 (1986); Z. Y. Li and Q. Jiang, Phys. Lett. A **138**, 247 (1989); X. Y. Chen, Q. Jiang, and Y.-Z. Wu, Solid State Commun. **121**, 641 (2002).
12. M. Suzuki, Progr. Theor. Phys. **56**, 1454 (1976).
13. I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Phys. Rev. Lett. **59**, 799 (1987).
14. S. Haas, J. Riera, and E. Dagotto, Phys. Rev. B **48**, 3281 (1994).
15. E. R. Gagliano and C. A. Balseiro, Phys. Rev. Lett. **59**, 2999 (1987).
16. A. K. Kolezhuk, H. J.-Mikeska, and S. Yamamoto, Phys. Rev. B **55**, R3336 (1997).
17. L. A. Takhtajan, Phys. Lett. A **87**, 479 (1982); H. M. Babujian, Nucl. Phys. B **215**, 317 (1983).
18. I. Affleck, T. Kennedy, E. H. Lieb, and H. Tasaki, Comm. Math. Phys. **115**, 477 (1988).
19. M. Uchinami, Phys. Rev. B **39**, 4554 (1989).
20. S. Östlund and S. Rommer, Phys. Rev. Lett. **75**, 3537 (1995); S. Rommer and S. Östlund, Phys. Rev. B **55**, 2164 (1997).
21. K. Okunichi, Phys. Rev. B **60**, 4043 (1999).
22. J. A. Plascak and N. A. Siwa, Phys. Status Solidi B **110**, 669 (1982).
23. S. Yamamoto and T. Sakai, Phys. Rev. B **62**, 3795 (2000).
24. T. Sakai and S. Yamamoto, J. Phys.: Condens. Matter **12**, 9787 (2000).
25. S. Brehmer, H.-J. Mikeska, and S. Yamamoto, J. Phys.: Condens. Matter **9**, 3921 (1997).
26. E. Lieb and D. Mattis, J. Math. Phys. **3**, 749 (1962).
27. S. Yamamoto, T. Fukui, and T. Sakai, Eur. Phys. J. B **15**, 211 (2000).
28. S. R. White, Phys. Rev. Lett. **69**, 2863 (1992); Phys. Rev. B **48**, 10345 (1993).
29. A. K. Kolezhuk, H.-J. Mikeska, K. Maisinger, and U. Schollwöck, Phys. Rev. B **59**, 13565 (1999).

Translated by I. Nikitin

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 624–631.
Translated from Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 691–699.

SOLIDS
Electronic Properties

The Dynamics of Transients in an Overdamped Josephson Junction Biased by Trains of Short Current Pulses with Controllable Parameters

V. M. Buchstaber, O. V. Karpov*, and S. I. Tertychnyi

Research Institute of Physicotechnical and Radio Engineering Measurements, Mendeleevo, Moscow oblast, 141570 Russia

**e-mail: mera@vniiftri.ru*

Received August 19, 2002

Abstract—The paper discusses the development of the method for studying the dynamics of an overdamped Josephson junction biased by a periodic current with controlled parameters, suggested earlier by the authors. The results obtained allow several known experimental observations to be explained. Analytic expressions for the rate of transient process damping and energy expenditures necessary for switching a Josephson junction from one quantum state to another are derived. Problems of interactions between the bias current and supercurrent and the influence of $\cos\varphi$-type terms on the rates of transient process damping and energy conversion are discussed. The results for junction biasing by short (uni- or bipolar) δ function-shaped pulses are obtained in the form of exact analytic expressions. Diagrams illustrating the dependence of the Shapiro step width on the shape of biasing pulses are given. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Recently, an important practical application of the Josephson effect is its use in DC voltage standards. The Volt unit is reproduced by apparatus based on this effect in nearly all countries having national metrological centers. The success in the development and wide practical use of DC voltage standards has stimulated works on creating quantum AC voltage synthesizers [1–4].

Instantaneous electric voltage $V(\theta)$ applied to a Josephson junction is known [5, 6] to be related to changes in the $\varphi(\theta)$ function by the Josephson equation

$$V(\theta) = \frac{\Phi_0}{2\pi}\frac{d\varphi}{d\theta}, \tag{1}$$

where $\varphi(\theta)$ is the phase difference of order parameter functions in weakly coupled Josephson junction superconductors (in what follows, junction phase function or merely junction phase), θ is the current dimensional time, $\Phi_0 = h/2e$ is the magnetic flux quantum, h is the Planck constant, and e is the charge of the electron. Note that, in metrology, the electric current voltage unit is reproduced using the Josephson constant $K_{J-90} \equiv K_J = 483597.9\ \text{GHz/V} \approx \Phi_0^{-1}\ \text{GHz/V} \approx \Phi_0^{-1}$ rather than the Φ_0 constant, which may be refined as new knowledge is accumulated [7, 8].

The principle of the generation (synthesis) of AC voltage using the Josephson effect is based on controlling the dynamics of the $\varphi(\theta)$ phase in such a way that the alternating "step" voltage (1) has the desired shape with fundamental accuracy. The dynamics of the $\varphi \equiv \varphi(\theta)$ phase of a Josephson junction with a finite electric capacitance C when the junction is biased by an external current is described by the solution to a second-order differential equation (e.g., see Eq. (2.42) in [5]). Under the conditions of phase φ locking by an external alternating bias current with frequency f, constant-voltage regions, so-called Shapiro steps [5], appear in the I–V characteristics of Josephson junctions. Within a certain time, the reproducible (observed) voltage defined by the equation

$$U(\theta) = \frac{1}{\Delta\theta}\int_{\theta}^{\theta+\Delta\theta} V(\theta')d\theta' = \frac{1}{2\pi K_J}\frac{\varphi(\theta+\Delta\theta)-\varphi(\theta)}{\Delta\theta} \tag{1'}$$

becomes equal to kf/K_J to within the width of these steps. Here, $k = 0, \pm1, \pm2, \pm3, \ldots$ is the ordinal number of the Shapiro step and $\Delta\theta \gg f^{-1}$.

The nonzero capacitance C leads to a hysteresis of the junction I–V characteristics [6] and, as a consequence, a nonunique dependence of the $U(\theta)$ on bias current parameters. In particular, at equal bias current parameters, voltage $U(\theta)$ can take on values corre-

sponding to different numbers k depending on the prehistory of phase evolution. This makes it difficult to the use of such Josephson structures for synthesizing AC voltage.

To minimize the undesirable effect of capacitance C on the dynamics of the Josephson junction phase, it is necessary that the condition [5, 6]

$$\omega_c \ll (R_N C)^{-1} \tag{2}$$

be satisfied, where R_N is the normal resistance, $\omega_c = 2\pi I_c R_N K_J$ is the critical frequency, and I_c is the critical junction current. This condition is met in so-called overdamped Josephson structures. In these structures, the dynamics of the $\varphi(\theta)$ phase can be described by the first-order differential equation [6], which, in dimensionless variables, has the form

$$\dot{\varphi}(t) + \sin\varphi(t) = i(t). \tag{3}$$

Here, $t = \omega_c\theta$ is the dimensionless time; $i(t) = i_0 + i_1(t)$ is the dimensionless external bias current with $\langle i_1(t)\rangle = 0$; $i_0 = I_0/I_c$ and I_0 are the dimensionless and dimensional constant bias current components, respectively; $i_1(t) = I_1(t)/I_c$ and $I_1(\theta)$ are the dimensionless and dimensional high-frequency bias current components, respectively; and the dot denotes the differentiation with respect to time.

Equation (3) was obtained based on the RSJ (Resistively Shunted Junction) model of the Josephson junction on the assumption that $\omega_c R_N C \ll 1$. As far as we know, no rigorous theoretical proof of the applicability of this model to real Josephson structures has been found. However, the results of numerous experimental studies of overdamped Josephson junctions [6], including studies of the *I–V* characteristics of two-barrier SINIS structures [9–11], are in good agreement with this model. What is more, currently, the quality of SINIS- and SNS-type Josephson structures is sometimes estimated by the correspondence of their *I–V* characteristics to that derived from the model based on Eq. (3).

At present, there are several competing scientific and technical approaches to the development of quantum AC synthesizers based on model (3). One of these is via controlling the dynamics of the Josephson junction phase by biasing the junction with a sinusoidal [1, 12–14] or short-pulse [2, 15–17] high-frequency current having variable parameters (amplitude and pulse repetition rate).

We used model (3) of the Josephson junction, in particular, to show [16, 17] that, for the biasing current of the form

$$i(t) = i_0^* + 2\pi i_{01}\sum_n \delta(t - nT) \tag{4a}$$

(unipolar pulses) or

$$i(t) = i_0 + 2\pi i_{01} \times \sum_n \left\{\delta\left(t - \frac{T+\tau}{2} - nT\right) - \delta\left(t - \frac{T-\tau}{2} - nT\right)\right\} \tag{4b}$$

(bipolar pulses) [17], the phase φ locking condition can be expressed as the inequality $|D| > 1$. Here, $\delta(t)$ is the Dirac δ function, $T = \omega_c f^{-1}$ is the dimensionless pulse repetition period, τ is the dimensionless lapse between pulses of different polarities within one period ($0 < \tau \le T/2$), $i_0^* = i_0 - 2\pi i_{01} T^{-1}$, and $i_{01} = I_c^{-1}\langle I_1(t)\rangle$ is the dimensionless "integral" amplitude of a separate pulse. The function $D \equiv D(i_0, i_{01}, T)$ is defined by the equations

$$D = \cosh\left(\frac{T}{2}\sqrt{1 - i_0^{*2}}\right)\cos(\pi i_{01}) - \frac{i_0^*}{\sqrt{1 - i_0^{*2}}}\sinh\left(\frac{T}{2}\sqrt{1 - i_0^{*2}}\right)\sin(\pi i_{01}) \tag{5a}$$

for bias current (4a) at $|i_0^*| < 1$ and

$$D = \cosh\left(\frac{1}{2}T\sqrt{1 - i_0^2}\right) - \frac{\sin^2(\pi i_{01})}{1 - i_0^2} \times \left[\cosh\left(\frac{T}{2}\sqrt{1 - i_0^2}\right) - \cosh\left(\left(\frac{T}{2} - \tau\right)\sqrt{1 - i_0^2}\right)\right] \tag{5b}$$

for bias current (4b) at $|i_0| < 1$. [If $|i_0^*| > 1$ in (5a) or $|i_0| > 1$ in (5b), we must use the analytic continuations of the hyperbolic functions to trigonometric functions, which corresponds to the substitution $\sqrt{1 - i_0^2} \longrightarrow \sqrt{-1}\sqrt{i_0^2 - 1}$.] Note that the use of δ functions as a mathematical model of a high-frequency junction bias current pulse is admissible if the pulse width is much smaller than the time ω_c^{-1} (a short pulse or shock pulse excitation).

Equations (5) determine the ranges of i_0, i_{01}, and T bias current parameter variations within which phase locking conditions is satisfied. A complete description of $U(t)$, however, requires knowledge of the evolution of the phase difference in the last multiplier in the right-hand side of (1'). For phase locking, Eq. (1') for the phase difference can be written in the dimensionless form [16]

$$\varphi(t + \Delta t) - \varphi(t) = 2\pi k\frac{\Delta t}{T} + \mu(t, \Delta t), \tag{6}$$

where $\Delta t = NT$, $N = 1, 2, 3, \ldots$; the $\mu(t, \Delta t)$ term describes the additional phase contribution related to

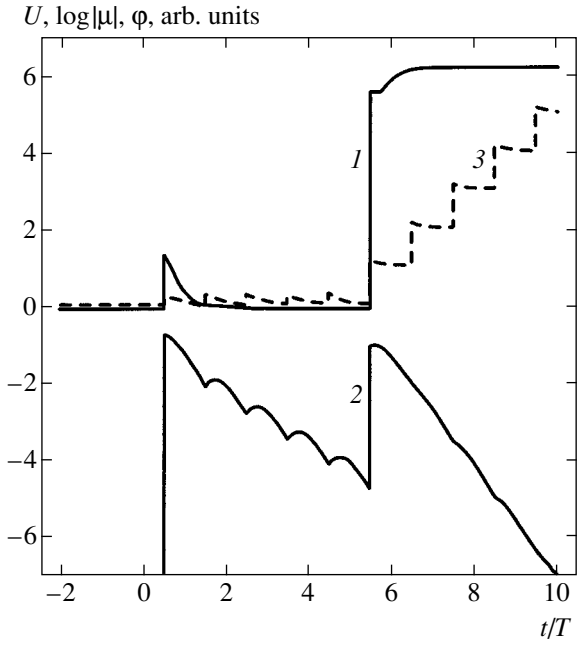

Fig. 1. Plots of (*1*) $U(t)$, (*2*) $\log|\mu(t, T)|$, and (*3*) $\varphi(t)$ functions.

transient processes, $\mu(t, \Delta t)$ tending to zero as $t \longrightarrow \infty$ at fixed bias current parameters.

To illustrate the aforesaid, the functions $U(t)$ and $\varphi(t)$ plotted in arbitrary units are shown in Fig. 1 (curves *1* and *3*, respectively). These dependences were obtained by the direct numerical integration of (3) with the right-hand side of (4a) for $T = 5.3$. Bias current $i(t)$ was set as follows. For $t < 0$, $i(t) = 0.6$ and $\varphi(t) = \text{const}$; for $0 \le t < 5T$ (phase locking conditions are met for the zeroth step),

$$i(t) = 0.6 + 0.22 \times 2\pi \sum_{n=0}^{4} \delta(t - 0.5T - nT),$$

and for $t \ge 5T$ (phase locking conditions are met for the first step),

$$i(t) = 0.6 + (0.22 + 0.9)2\pi \sum_{n=5}^{10} \delta(t - 0.5T - nT).$$

The $\log|\mu(t, T)|$ function is also shown in Fig. 1 (curve *2*). This function characterizes the accuracy of the voltage generated. The figure shows that the $\mu(t, T)$ function can be comparatively large for parameter t variation intervals on the order of several periods T.

The purpose of this work was to further develop the method suggested in [16, 17] for studying the dynamics of the phase and transient processes in an overdamped Josephson junction based on model (3), to obtain new results, substantiate the most important and fundamental old results, and determine particular analytic dependences for exactly solvable bias current models (4).

2. PROBLEM STATEMENT AND ANALYTIC SOLUTION

2.1. Problem Statement

The dynamics of transient processes in an overdamped Josephson junction will be studied by the method suggested by us in [17]. In place of phase $\varphi(t)$, we introduce two real functions $x(t)$ and $y(t)$ according to the equation

$$e^{j\varphi} = \frac{x - jy}{x + jy}. \tag{7}$$

This equation defines the functions $x(t)$ and $y(t)$ to within some factor, which is a nonzero real function. This function can be selected in such a way that $x(t)$ and $y(t)$ related to $\varphi(t)$ that satisfies (3) be solutions to the following system of equations [17]:

$$\begin{aligned} 2\dot{x}(t) &= x(t) + i(t)y(t), \\ -2\dot{y}(t) &= i(t)x(t) + y(t). \end{aligned} \tag{8}$$

Note that an arbitrary nonzero solution to system (8) determines the solution to (3) (see Section 3) to within a summand multiple to 2π.

Two points should be mentioned. First, system (8) can be reduced to a second-order Hill-type equation if the $i(t)$ function is periodic. Solutions to such equations have been well studied and proved to play a key role in many practically important applications. In particular, the solution to the Schrödinger equation that determines the form of the allowed energy bands for the electron in a rectangular periodic field [18] leads to a dependence formally similar to (5).

Secondly, system (8) with the replacement $t \longrightarrow j\omega_c\theta$ becomes formally similar to the system of quantum-mechanical equations first used by Josephson to describe the Josephson effect. Hopefully, this similarity can be used as a basis for rigorously substantiating mathematical model (3) of some overdamped Josephson structures.

Next, note that system (8) of linear differential equations satisfies the conditions of the Lyapunov theorem [19] if $i(t) = i_0 + i_1(t)$ is periodic. According to this theorem, provided the criterion for phase φ locking by periodic external current $i(t)$ is satisfied, any solution to (8) can be represented in the form

$$\begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = L(t) \begin{bmatrix} x_1(t - t_0) \\ y_1(t - t_0) \end{bmatrix}, \tag{9a}$$

where $L(t)$ is a periodic 2×2 matrix, that is,

$$L_{p,q}(t+T) = L_{p,q}(t), \quad p = 1, 2, \quad q = 1, 2,$$

and the $x_1(t)$ and $y_1(t)$ functions generally have the form $x_1(t) = \exp(\lambda_1 t)$ and $y_1(t) = \exp(\lambda_2 t)$ or $x_1(t) = [t + x_1(t_0)]\exp(\lambda t)$ and $y_1(t) = \exp(\lambda t)$ in the degenerate case, where λ_1, λ_2, λ, and $x_1(t_0)$ are some real constants and t_0 is some initial time. Note that the degeneracy takes place at the boundary of the regions of the i_0, i_{01}, and T parameter values which satisfy the conditions of phase locking and does not correspond to the phase locking conditions.

Moreover, it can be shown that $\lambda_1 = -\lambda_2 = \lambda_0$, because the trace of the matrix of system (8) is zero. Under phase locking conditions, $\lambda_0 \neq 0$. Further, for definiteness we assume that $\lambda_0 > 0$. Equation (9a) is then simplified and can be rewritten in the form

$$x(t) = L_{11}(t)\exp[\lambda_0(t-t_0)] + L_{12}(t)\exp[-\lambda_0(t-t_0)], \tag{9b}$$

$$y(t) = L_{21}(t)\exp[\lambda_0(t-t_0)] + L_{22}(t)\exp[-\lambda_0(t-t_0)].$$

2.2. Phase Locking Regions and Exactly Solvable Models

If an overdamped Josephson junction is biased by external current $i(t)$ (4a) or (4b), the λ_0 value can be found from the equation [17]

$$\kappa = 2\lambda_0 T = 2\ln(|D| + \sqrt{D^2 - 1}), \tag{10}$$

where $D \equiv D(i_0, i_{01}, T)$ is, as previously, defined by (5).

Examples of the regions of the presence (absence) of phase locking calculated from the condition $|D| > 1$ ($|D| \leq 1$) by (5) are shown in Fig. 2. The same figure shows levels of equal κ values given by thin lines within phase locking regions (zones). Such graphic maps can be used to construct any *I–V* characteristic of an overdamped Josephson junction corresponding to fixed parameters i_0, i_{01}, and T of the bias current (4).

Note in particular that the *I–V* characteristic of a Josephson junction biased by unipolar pulses (4a) is asymmetric with respect to the axis of an direct bias current component i_0. This is related to the presence of an additional constant component equal to $2\pi i_{01} T^{-1}$ in bias current (4a).

If a Josephson junction is biased by bipolar pulses (4b), the τ parameter becomes important. The projections of the φ phase locking region [Eq. (3)] onto the direct bias current i_0–high-frequency bias current $2\pi i_{01}$ plane are shown in Figs. 2b and 2c for $\tau = T/2$ and $\tau/T \ll 1$, respectively, at $T = 5.3$. The series of graphic maps shown in Fig. 2 demonstrates the dependence of the

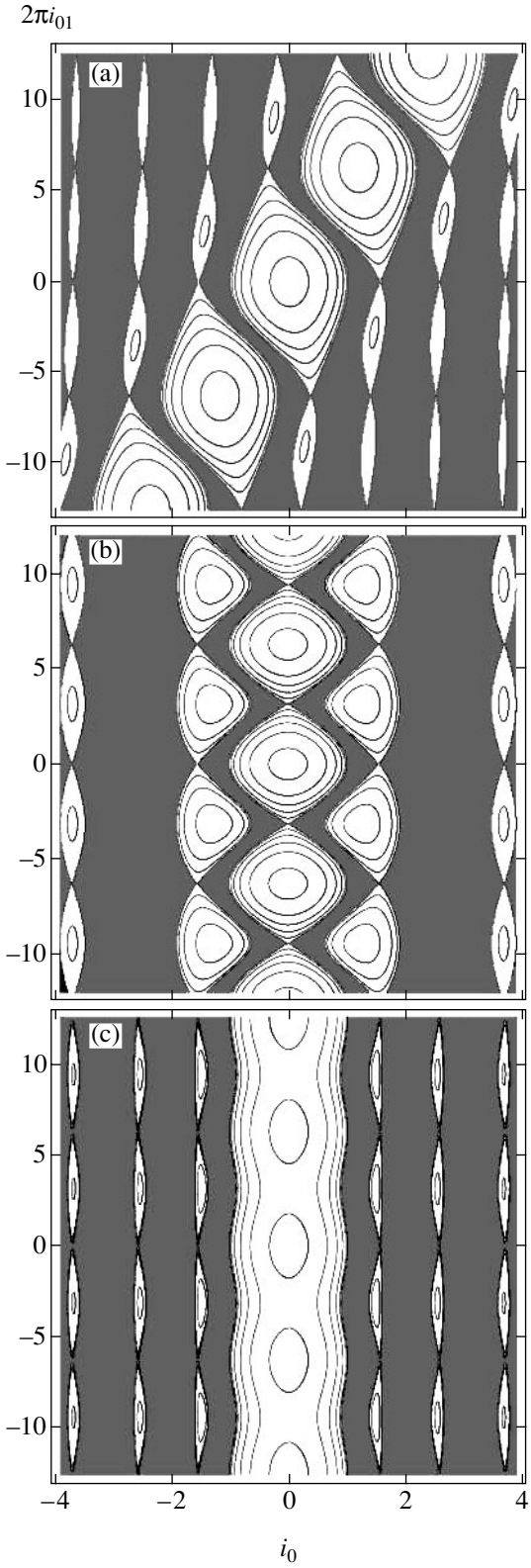

Fig. 2. Influence of the shape of bias current pulses on phase locking zones. Phase locking zones are shown by light regions; zones without phase locking are shaded. The constant κ = 1, 2, 3, 4, 5 value levels (κ increases from the boundaries to the center) are shown by lines. (a) Unipolar pulse, (b) bipolar "centered" pulse ($\tau = 0.5T$), and (c) bipolar "noncentered" pulse ($\tau = 0.005T$).

configuration of the phase locking regions of system (3) on the shape of high-frequency bias current pulses.

2.3. Analytic Solution

Equation (7) determines the $x(t)$ and $y(t)$ entering into system (8) of linear equations to within some factor, which is a nonzero real function. We will now complete the definition of these functions by representing them in the form

$$\begin{aligned} x(t) &= \exp\left[\lambda_0 t + \frac{1}{2}\rho(t)\right]\cos\frac{\varphi(t)}{2}, \\ -y(t) &= \exp\left[\lambda_0 t + \frac{1}{2}\rho(t)\right]\sin\frac{\varphi(t)}{2}. \end{aligned} \tag{11a}$$

Here and throughout, the $\rho(t) \equiv \rho(t, \lambda_0)$ function is defined by the equation

$$\begin{aligned} \exp[\rho(t)] = \{&[L_{11}(t) + L_{12}(t)\exp(-2\lambda_0 t)]^2 \\ &+ [L_{21}(t) + L_{22}(t)\exp(-2\lambda_0 t)]^2\}, \end{aligned} \tag{11b}$$

and the $\varphi(t)$ phase function, by (7).

Substituting (11a) into (8) and performing certain transformations yields a system of linear homogeneous equations for $\sin(\varphi(t)/2)$ and $\cos(\varphi(t)/2)$, namely,

$$\begin{aligned} [1 - 2\lambda_0 - \dot\rho(t)]\cos(\varphi/2) + [i(t) - \dot\varphi(t)]\sin(\varphi/2) &= 0, \\ [i - \dot\varphi(t)]\cos(\varphi/2) + [1 + 2\lambda_0 + \dot\rho(t)]\sin(\varphi/2) &= 0. \end{aligned} \tag{12}$$

The compatibility condition for (12) can be written in the form

$$1 - [2\lambda_0 + \dot\rho(t)]^2 - [i(t) - \dot\varphi(t)]^2 = 0, \tag{13a}$$

or, taking into account the initial equation $i(t) - \dot\varphi(t) = \sin\varphi(t)$ [Eq. (3)], in the form of two equations

$$\sin\varphi(t) = i(t) - \dot\varphi(t), \quad \cos\varphi(t) = 2\lambda_0 + \dot\rho(t). \tag{13b}$$

Two remarks should be made. First, the sign "+" is chosen for $\cos\varphi$ term in the second equation in (13b) because the equation with $-\cos\varphi(t)$ does not correspond to a solution of (8). Secondly, the compatibility conditions for (13) mean that the λ_0 constant and the "new" $\dot\rho(t)$ function can be used to describe the dynamics of the phase φ in (3) if they are given an explicit physical interpretation. It should at the same time be emphasized that the quantities λ_0 and $\dot\rho(t)$ are not independent parameters.

Next, note that, for phase locking at $N \longrightarrow \infty$, the $\rho(t + NT)$ term in (13b) becomes a periodic function of time t; therefore, $\lim_{N\to\pm\infty}\langle\dot\rho(t + NT)\rangle = 0$, where the angle brackets denote averaging over time interval T. Taking this and (13b) into account, we obtain

$$\lambda_0 = \frac{1}{2}\lim_{N\to\infty}\langle\cos[\varphi(t + NT)]\rangle$$

and, therefore, $\lambda \le 1/2$. Similarly, we obtain

$$-\lambda_0 = \frac{1}{2}\lim_{N\to\infty}\langle\cos[\varphi(t - NT)]\rangle.$$

Further, let us express $\cos[\varphi(t \pm NT)]$ via the $L_{p,q}(t \pm NT)$ functions. The equations for λ_0 then take the form

$$\lambda_0 = \frac{1}{2}\lim_{N\to\infty}\left\langle\frac{L_{11}^2 - L_{21}^2}{L_{11}^2 + L_{21}^2}\right\rangle = \frac{1}{2}\lim_{N\to-\infty}\left\langle\frac{L_{22}^2 - L_{12}^2}{L_{22}^2 + L_{12}^2}\right\rangle. \tag{14}$$

3. THE DYNAMICS OF TRANSIENT PROCESSES AND THE ASYMPTOTIC PROPERTIES OF THE PHASE FUNCTION

3.1. The Dynamics of the Phase

Although Eqs. (9)–(14) do not yield a complete analytic solution of system (8) of linear differential equations, they nevertheless can be used for analysis of the evolution of the $\varphi(t)$ phase determined by (3) and therefore transient processes if the conditions of the phase locking by a periodic external bias current $i(t)$ are satisfied. Indeed, using the transformation

$$\varphi(t) = -2\arctan\frac{y(t)}{x(t)} + 2\pi\chi(t), \tag{15}$$

inverse to (7), and the solution to system (8) in form (9b), we obtain

$$\varphi(t) = -2\arctan\frac{L_{21}(t)\exp[\lambda_0(t - t_0)] + L_{22}(t)\exp[-\lambda_0(t - t_0)]}{L_{11}(t)\exp[\lambda_0(t - t_0)] + L_{12}(t)\exp[-\lambda_0(t - t_0)]} + 2\pi\chi(t), \tag{16}$$

where the $\chi(t)$ function in (15) is some step function experiencing jumps of ± 1 at the points at which the $x(t)$ denominator vanishes. As previously, the $L_{p,q}(t) \equiv L_{p,q}(t; t_0, T, i_0, i_{01})$ functions in (16) are periodic functions of time t, which depend on parameter t_0 and bias current $i(t)$.

As t changes in the range $-\infty < t < \infty$, the $\varphi(t)$ phase function in (16) changes from

$$\varphi_{-\infty}(t) = -2\arctan\frac{L_{22}(t)}{L_{12}(t)} + 2\pi\chi(t) \quad (17)$$

to

$$\varphi_{\infty}(t) = -2\arctan\frac{L_{21}(t)}{L_{11}(t)} + 2\pi\chi(t). \quad (18)$$

It follows that, for $-\infty < t < \infty$, the solution to (16) describes the natural (that is, depending on the prehistory) transition of system (3) to the only possible "stable" asymptotic state given by (18) [20]. The rate of the transition of system (3) from one state to another and the rate of transient process damping in (16) are characterized by the λ_0 exponent. This means that, no matter what the initial conditions at time $t = t_0$ for the phase $0 \le \varphi(t_0) < 2\pi$, system (3) sooner or later (depending on the $\lambda_0 \neq 0$ value) experiences the transition to a stable state with a priori known properties determined exclusively by bias current parameters.

So far, we assumed that

$$\lambda_0 \equiv \lambda_0(t_0; T, i_0, i_{01}) > 0.$$

If $\lambda_0 < 0$, we similarly obtain the second solution to (3) to which (17) corresponds when $t \longrightarrow \infty$. In these conditions, (16) describes the natural transition of system (3) in the $-\infty < t_0 < t < \infty$ interval to the only possible stable asymptotic state given by (17).

3.2. The Asymptotic Properties of the Phase Function

Consider the behavior of $\langle\dot{\varphi}(t + NT)\rangle$ as $N \longrightarrow \infty$. Taking into account the periodicity of $L_{p,q}(t)$ and the properties of the function $\chi(t)$ introduced in (16), we can write

$$\lim_{N\to\infty}\langle\dot{\varphi}(t + NT)\rangle = \lim_{N\to\infty}\frac{\varphi(t + (N+1)T) - \varphi(t + NT)}{T} = \frac{2\pi k}{T}, \quad (19)$$

where, as previously, $k \equiv k(i_0, i_{01}, T)$ is the number of the Shapiro step.

Further, (19) and (3) can be used to show that, in the stationary (asymptotic) state, the equation

$$\lim_{N\to\infty}\langle\sin[\varphi(t + NT)]\rangle = i_2 = i_0 - \frac{2\pi}{T}k \quad (20)$$

is valid. By its physical meaning, $\imath_2 \equiv i_2(k, \varphi_0)$ corresponds to the mean superconduction current (supercurrent) that flows through the Josephson junction.

Indeed, using (13b), we obtain

$$\tan\varphi_0 = \frac{\langle\sin[\varphi_{\infty}(t)]\rangle}{\langle\cos[\varphi_{\infty}(t)]\rangle} = \frac{i_2}{2\lambda_0} = \frac{i_0 - 2\pi kT^{-1}}{2\lambda_0}, \quad (21a)$$

where, for the phase function in the form $\varphi_{\infty}(t) = \varphi_1(t) + \varphi_0$, the following equations are valid:

$$\lim_{N\to\infty}\langle\sin[\varphi_1(t + NT)]\rangle = 0, \quad \lim_{N\to\infty}\langle\cos[\varphi_1(t + NT)]\rangle = \Delta i_{Sh}(k). \quad (21b)$$

We then have $i_2 = \Delta i_{Sh}(k)\sin\varphi_0$. Here and throughout, φ_0 is some constant defined only under phase locking conditions that characterizes the state of system (3) and is equal to the phase difference between the external bias current with frequency $f\omega_c^{-1}$ and the supercurrent with mean frequency $\langle\dot{\varphi}(t)\rangle$. If $|i_0| < 1$ and $i_{01} = 0$, we obtain $k = 0$ and $\Delta i_{Sh}(k) = 1$ and arrive at the well-known Josephson relation $i_2 = \sin\varphi_0$ [5, 6].

4. ENERGY CONVERSION AND BALANCE

Let us consider the physical meaning of the functions $\dot{\varphi}(t)$, $\dot{\rho}(t)$, $\sin[\varphi(t)]$, and $\cos[\varphi(t)]$ functions and the constant λ_0. For this purpose, we will use the RSJ model of Josephson junctions [5, 6]. According to this model, an overdamped Josephson junction can be represented by two functional elements. The first element is a supercurrent generator; the generated supercurrent direction and value are determined by $\sin[\varphi(t)]$. The second element is the "Ohmic" normal resistance of the Josephson junction. The voltage drop across the normal Josephson junction resistance [according to (3)] and the current and supercurrent that flow through this resistance are described by $\dot{\varphi}(t)$.

Next, differentiating the second equation in (13b) yields the $\ddot{\rho}(t) = -\sin[\varphi(t)]\dot{\varphi}(t)$ equation for the power spent (acquired) by the supercurrent of unit strength in overcoming the $\dot{\varphi}(t)$ potential of the barrier to the transition. In other words, $\ddot{\rho}(t)$ determines the interaction power between the supercurrent of unit strength and the bias current (high-frequency current absorption or emission, bias current transformation into a direct or high-frequency supercurrent, etc.). Then $\dot{\rho}(t) = \int\ddot{\rho}(t')dt' + \text{const}$ describes the work performed by the supercurrent of unit strength in some time interval.

It follows (also see [6, p. 31]) that changes in the $E(\varphi)$ barrier energy of a Josephson junction related to

the work of the supercurrent of unit strength with phase $\varphi(t)$ can be described by the equation

$$E(\varphi(t)) = E_0^{(1)}[2\lambda_0 - \cos\varphi(t)], \quad (22a)$$

where $E_0^{(1)} = \hbar\omega_c = 2eV_c$ and $V_c = R_N I_c$ is a constant, the so-called critical voltage determined by the gap width of the superconductor.

By way of example, consider the case of $|i_0| < 1$ and $i_{01} = 0$. The supercurrent that flows through the junction is then $i_2 = \sin\varphi_0$, and the barrier potential is zero. Changes in the $E(\varphi)$ barrier energy caused by the interaction with the supercurrent, should also be zero. Indeed, as $\dot{\varphi}(t) = 0$ for $|i_0| < 1$ and $i_{01} = 0$ and the exponent $2\lambda_0 = +\sqrt{1 - i_0^2} = \cos\varphi_0$, we find from (22a) that $E(\varphi_0) = 0$ for all φ_0 in the interval $-\pi/2 < \varphi_0 < \pi/2$.

Next, note that the $2\lambda_0 = 2\Delta i_{Sh}(k)\cos\varphi_0$ value corresponds to the current flowing through the barrier by its physical meaning [see (20) and (21a)]. For this reason, the equality $E_0^{(2)} = hI_c/2e$ should preferably be used in Eq. (22a) for the kinetic energy of the supercurrent. Taking this into account, let us rewrite (22a) in the form

$$E(\varphi(t)) = \beta E_0^{(2)}[2\lambda_0 - \cos\varphi(t)], \quad (22b)$$

where value $\beta = E_0^{(1)}/E_0^{(1)} = R_N/(h/4e^2) = R_N/R_0$ plays the role of the "viscosity" parameter of the medium and $R_0 \approx 6.4$ kΩ is the fundamental constant, the so-called quantum Hall resistance [7, 8]. In our problem, the β parameter characterizes (according to the widely used but, probably, not quite adequate, terminology) the degree of "strong shunting" (or screening) of the junction.

To summarize, we showed that the $\dot{\rho}(t)$ function corresponds to the $E(\varphi)$ energy of the interaction between the Josephson junction barrier and the supercurrent by its physical meaning. Comparing the second equation in (13b) with (22a), we then eventually obtain

$$\dot{\rho}(t) = -\frac{2\pi\beta}{T}\frac{E(\varphi)}{hf}. \quad (23)$$

For real overdamped Josephson junctions produced by the SNS or SINIS technology, the R_N normal junction resistance is in the range of 0.01–0.1 Ω. We therefore have $\beta \sim 10^{-6}$–10^{-5}. This means that the duration of transient processes, which has the same order of magnitude as the $1/2\lambda_0$ value, can amount to $(0.01–0.5)f^{-1}$ under real operation conditions of overdamped Josephson junctions.

5. CONCLUSIONS

The approach to studying the dynamics of the phase and transient processes suggested by us in [16, 17] was developed based on model (3) of an overdamped Josephson junction. The results explained several most important and fundamental experimental facts that determined the dynamics and asymptotic properties of the phase of the Josephson junction biased by an external current.

In particular, (16) can be used to describe the dynamics of the Josephson junction when it is switched from one quantum state into another by an external bias current and to determine the rate of transient process damping and the rate at which asymptotic operation conditions are reached. Simultaneously using (16) and (23), we can determine energy expenditures necessary for establishing such conditions.

In addition, we touched upon problems of the interaction of the bias current with the supercurrent flowing through the junction. It was shown that the rate of energy conversion was determined by the barrier potential of the junction, which was, in turn, determined (controlled) by bias current parameters. Several questions, however, remain open, for instance, the question of the influence of the function $\varphi_1(t)$ in (21b) on the $\Delta i_{Sh}(k)$ function, which requires additional study.

The results obtained in this work can find applications in the physics of processes whose dynamics is described by (3) and in designing Josephson chips for quantum AC voltage synthesizers for telecommunication purposes and metrology.

ACKNOWLEDGMENTS

The authors thank A.A. Rukhadze and the participants of the seminar that he heads for useful discussions and to J. Niemeyer (PTB, Germany) for attention to this work. The work was financially supported by an international contract with PTB (contract no. Transform 13N7259).

REFERENCES

1. C. A. Hamilton, C. J. Burroughs, and R. L. Kautz, IEEE Trans. Instr. Meas. **44**, 223 (1995).
2. J. X. Przybysz, A. H. Worshman, S. P. Benz, *et al.*, U.S. Patent No. 5812078 (1998).
3. J. Niemeyer, Appl. Supercond. **167**, 707 (2000).
4. European Project, www.jaws-project.nl.
5. K. K. Likharev and B. T. Ul'rikh, *Systems with Josephson Contacts* (Mosk. Gos. Univ., Moscow, 1978).
6. A. Barone and G. Paterno, *Physics and Applications of the Josephson Effect* (Wiley, New York, 1982; Mir, Moscow, 1984).
7. B. N. Taylor and T. Y. Witt, Metrologia **26**, 47 (1989).
8. T. J. Quin, Metrologia **26**, 69 (1989).

9. J. Niemeyer, Transform Project BMBF 13N7259 (Physikalisch–Technische Dundesanstalt, Braunschweig, 2000), p. 48.
10. O. V. Karpov, V. D. Koutovoi, S. V. Sherstobitov, and J. Niemeyer, Metrologia **38**, 471 (2001).
11. O. V. Karpov, V. D. Kutovoĭ, S. V. Sherstobitov, and J. Niemeyer, Prib. Tekh. Éksp., No. 5, 91 (2001).
12. C. A. Hamilton, S. P. Benz, C. J. Burroughs, *et al.*, IEEE Trans. Instr. Meas. **46**, 224 (1997).
13. O. V. Karpov, S. V. Sherstobitov, and Yu. F. Verkhovykh, Prib. Tekh. Éksp., No. 5, 83 (2002).
14. O. V. Karpov, V. D. Koutovoi, S. V. Sherstobitov, *et al.*, in *Conference Digest CPEM-2002, Ottawa, Canada* (2002), p. 496.
15. S. P. Benz, C. J. Burroughs, T. E. Hawey, and C. A. Hamilton, IEEE Trans. Appl. Supercond. **9** (2), 3306 (1999).
16. V. M. Buchstaber, O. V. Karpov, and S. I. Tertychnyĭ, Zh. Éksp. Teor. Fiz. **120**, 1478 (2001) [JETP **93**, 1280 (2001)].
17. V. M. Buchstaber, O. V. Karpov, and S. I. Tertychniy, in *Conference Digest CPEM-2002, Ottawa, Canada* (2002), p. 502.
18. S. V. Vonsovskiĭ and M. I. Katsnel'son, *Quantum Solid-State Physics* (Nauka, Moscow, 1983; Springer, Berlin, 1989), p. 182.
19. F. R. Gantmakher, *The Theory of Matrices* (Fizmatgiz, Leningrad, 1953; Chelsea, New York, 1959).
20. S. I. Tertychnyĭ, Usp. Mat. Nauk **55**, 195 (2000).

Translated by V. Sipachev

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 632–644.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 700–713.

SOLIDS
Electronic Properties

Strong Interaction of Correlated Electrons with Phonons: Exchange of Phonon Clouds by Polarons¶

V. A. Moskalenko[a], P. Entel[b], M. Marinaro[c], and D. F. Digor[d]

[a]*Joint Institute for Nuclear Research, Dubna, Moscow oblast, 141980 Russia*
e-mail: moskalen@thsunl.jinr.ru
[b]*Theoretische Physik, Gerhard-Mercator-Universität Duisburg 47048, Duisburg, Germany*
e-mail: entel@thp.uni-duisburg.de
[c]*Dipartimento di Scienze Fisiche E. R. Caianiello, Universita degli Studi di Salerno 84081, Baronissi, Italy*
e-mail: marinaro@sa.infn.it
[d]*Institute of Applied Physics, Moldova Academy of Sciences, Chisinau, 2028 Moldova*
e-mail: statphys@asm.md
Received January 24, 2003

Abstract—The interaction of strongly correlated electrons with phonons in the framework of the Hubbard–Holstein model is investigated. The electron–phonon interaction is considered to be strong and is an important parameter of the model, in addition to the Coulomb repulsion of electrons and the band filling. This interaction with nondispersive optical phonons is transformed to the problem of mobile polarons using the canonical transformation of Lang and Firsov. We discuss the case where the on-site Coulomb repulsion is exactly canceled by the phonon-mediated attractive interaction. It is suggested that polarons exchanging phonon clouds can lead to polaron pairing and superconductivity. The fact that the frequency of the collective mode of phonon clouds is larger than the bare frequency then determines the superconducting transition temperature. © *2003 MAIK "Nauka/Interperiodica".*

1. INTRODUCTION

Since the discovery of high-temperature superconductivity by Bednorz and Müller [1], the Hubbard model and related models such as RVB and *t–J* have been widely used to discuss the physical properties of the normal and superconducting states [2–6]. However, a unanimous explanation for the origin of the condensate in high-temperature superconductors has not emerged so far. One of the unsolved questions is how far phonons can be involved in the formation of the superconducting state. In experimental and theoretical works, the change in phonon frequencies and phonon lifetimes associated with the superconducting transition was mostly discussed. For example, the decrease in frequencies of Raman-active phonons at the transition [7], observation of the isotope effect for nonoptimally doped superconductors [8], and observation of a phonon-induced structure in the tunnel characteristics [9] are evidence in favor of strong electron–phonon coupling in the cuprates.

The aim of the present paper is to gain further insight into the mutual influence of strong on-site Coulomb repulsion and strong electron–phonon interaction using the single-band Hubbard–Holstein model and a recently developed diagram approach [10–14].

For simplicity, we consider coupling to dispersionless phonons only, although this might not be the most interesting case as regards superconductivity. However, previous investigations [15–17] have shown that the Hubbard–Holstein model [18, 19] constitutes a formidable problem of its own. Other authors have also intensively studied this model Hamiltonian [20–23].

Because the interactions between electrons and between electrons and phonons are strong, we include the Coulomb repulsion in the zero-order Hamiltonian and apply the canonical transformation of Lang and Firsov [24] to eliminate the linear electron–phonon interaction. In the strong electron–phonon coupling limit, the resulting Hamiltonian of hopping polarons (i.e., hopping electrons surrounded by clouds of phonons) can lead to an attractive interaction among electrons mediated by the phonons. In this limit, the chemical potential, the on-site Coulomb energy, and the frequency of the collective mode of phonon clouds (which is much larger than the bare frequency of the Einstein oscillators) are strongly renormalized [17, 25, 26], which affects the dynamical properties of the polarons and the character of the superconducting transition. In our discussion of this, we assume that the renormalized on-site Coulomb repulsion and attractive electron–electron interaction completely cancel each other. We suggest that the resulting superconducting state with polaronic Cooper pairs is mediated by the

¶This article was submitted by the authors in English.

exchange of phonon clouds during the hopping processes of the electrons.

2. THEORETICAL APPROACH

2.1. Lang–Firsov Transformation of the Hubbard–Holstein Model

The initial Hamiltonian of correlated electrons coupled to optical phonons with bare frequency ω_0 is given by

$$\mathcal{H} = \mathcal{H}_e + \mathcal{H}^0_{\text{ph}} + \mathcal{H}_{e\text{–ph}}, \tag{1}$$

$$\mathcal{H}_e = \sum_{i,j,\sigma}\{t(j-i)-\epsilon_0\delta_{ij}\}a^\dagger_{j\sigma}a_{i\sigma} + U\sum_i n_{i\uparrow}n_{i\downarrow}, \tag{2}$$

$$\mathcal{H}^0_{\text{ph}} = \sum_i \hbar\omega_0\left(b^\dagger_i b_i + \frac{1}{2}\right), \tag{3}$$

$$\mathcal{H}_{e\text{–ph}} = g\sum_i n_i q_i,$$

$$n_i\sum_\sigma n_{i\sigma}, \quad n_{i\sigma} = a^\dagger_{i\sigma}a_{i\sigma}, \tag{4}$$

$$q_i = \frac{1}{\sqrt{2}}(b_i + b^\dagger_i),$$

where $a^\dagger_{i\sigma}$ ($a_{i\sigma}$) and $b^\dagger_i$ (b_i) are creation (annihilation) operators of electrons and phonons, respectively; i refers to the lattice site, and σ to the spin; q_i is the phonon coordinate; g is the electron–phonon interaction constant; U is the on-site Coulomb repulsion; $t(j-i)$ is the two-center transfer integral; and $\epsilon_0 = \bar{\epsilon}_0 - \mu$, with a local energy of $\bar{\epsilon}_0$ and a chemical potential of μ. The Fourier representation of $t(j-i)$ is related to the tight-binding dispersion $\varepsilon(\mathbf{k})$ of bare electrons,

$$t(j-i) = \frac{1}{N}\sum_{\mathbf{k}}\varepsilon(\mathbf{k})\exp\{-i\mathbf{k}\cdot(\mathbf{R}_j - \mathbf{R}_i)\},$$

with a bandwidth of W. The energy scale of this model is fixed by the parameters W, U, g, and $\hbar\omega_0$. An additional parameter is given by the band filling.

After applying the Lang–Firsov transformation [24]

$$\mathcal{H}_p = e^S\mathcal{H}e^{-S}, \quad c_{i\sigma} = e^S a_{i\sigma}e^{-S}, \quad c^\dagger_{i\sigma} = e^S a_{i\sigma}e^{-S} \tag{5}$$

with

$$S = -i\bar{g}\sum_i n_i p_i, \quad \bar{g} = \frac{g}{\hbar\omega_0}, \quad p_i = \frac{i}{\sqrt{2}}(b^\dagger_i - b_i), \tag{6}$$

where p_i is the phonon momentum and $\bar{g}$ is the dimensionless interaction constant, we obtain the polaron Hamiltonian:

$$\mathcal{H}_p = \mathcal{H}^0_p + \mathcal{H}^0_{\text{ph}} + \mathcal{H}_{\text{int}}, \tag{7}$$

$$\mathcal{H}^0_p = \sum_i \mathcal{H}^0_{ip}, \quad \mathcal{H}^0_{ip} = \epsilon\sum_\sigma n_{i\sigma} + \bar{U}n_{i\uparrow}n_{i\downarrow}, \tag{8}$$

$$\mathcal{H}_{\text{int}} = \sum_{i,j,\sigma} t(j-i)c^\dagger_{j\sigma}c_{i\sigma}, \tag{9}$$

where

$$c^\dagger_{i\sigma} = a^\dagger_{i\sigma}\exp(-i\bar{g}p_i), \quad c_{i\sigma} = a_{i\sigma}\exp(i\bar{g}p_i), \tag{10}$$

$$\epsilon = \bar{\epsilon}_0 - \bar{\mu}, \quad \bar{\mu} = \mu + \alpha\hbar\omega_0, \quad \bar{U} = U - 2\alpha\hbar\omega_0, \quad \alpha = \frac{1}{2}\bar{g}^2. \tag{11}$$

To derive the polaron Hamiltonian, it was necessary to include the shift of the phonon coordinate q_i of the form

$$e^S q_i e^{-S} = q_i - \bar{g}n_i,$$

which is responsible for the elimination of the linear electron–phonon interaction. The polaron Hamiltonian is a polaron–phonon operator by its nature; i.e., the creation operator $c^\dagger_{i\sigma}$ and the destruction operator $\mathcal{H}_p$ entering $c_{i\sigma}$ must be interpreted as creation and destruction operators of polarons (electrons dressed with displacements of ions) that couple dynamically to the momentum of the optical phonon. In the zero-order approximation (omitting $\mathcal{H}_{\text{int}}$), polarons and phonons are localized with the strongly renormalized chemical potential $\bar{\mu}$ and on-site Coulomb interaction $\bar{U}$. The operator $\mathcal{H}_{\text{int}}$ describes tunneling of polarons between lattice sites, i.e., tunneling of electrons surrounded by clouds of phonons.

2.2. Expansion around the Atomic Limit

The problem is now to deal properly with the impact of electronic correlations on the polaron problem. This can be done best using Green's functions provided one finds a key for dealing with the spin and charge degrees

of freedom. In the general case where $\bar{U}$ is different from zero, the Coulomb interaction must be included in the zero-order Hamiltonian. As a consequence, conventional perturbation theory of quantum statistical mechanics is an inadequate tool because it relies on the expansion of the partition function around the noninteracting state (achieved using the traditional Wick theorem and conventional Feynman diagrams). A similar situation occurs for composite particles such as polarons,

$$c_{i\sigma} = a_{i\sigma}\exp(i\bar{g}p_i),$$

involving operators for the electron and phonon subsystems.

Hubbard [27] proposed a graphical expansion for correlated electrons about the atomic limit in powers of hopping integrals. This diagram approach was systematically reformulated for the single-band Hubbard model by Slobodyan and Stasyuk [28] and independently by Zaitsev [29], and further developed by Izyumov [30]. In these approaches, the complicated algebraic structure of the projection or Hubbard operators was used. It therefore appeared to be more appropriate to develop a diagram technique involving simpler creation and annihilation operators for electrons at all intermediate stages of the theory (see [10, 11] for details). In the latter approach, the averages of chronological products of interactions are reduced to the n-particle Matsubara Green's functions of the atomic system. These functions can be factorized into independent local averages using a generalization of the Wick theorem (GWT), which takes strong local correlations into account (details are given in [10, 11, 25]). Application of the GWT yields new irreducible on-site many-particle Green's functions, or Kubo cumulants. These new functions contain all local spin and charge fluctuations. A similar linked-cluster expansion for the Hubbard model around the atomic limit was recently reformulated by Metzner [31].

2.3. Averages of Phonon Operators

We define the temperature Green's function for polarons in (7) in the interaction representation by

$$\mathcal{G}(\mathbf{x}, \sigma, \tau|\mathbf{x}', \sigma', \tau') = -\langle Tc_{\mathbf{x}\sigma}(\tau)\bar{c}_{\mathbf{x}'\sigma'}(\tau')U(\beta)\rangle_0^c \quad (12)$$

with

$$c_{\mathbf{x}\sigma}(\tau) = \exp(\mathcal{H}^0\tau)c_{\mathbf{x}\sigma}\exp(-\mathcal{H}^0\tau),$$

$$\bar{c}_{\mathbf{x}\sigma}(\tau) = \exp(\mathcal{H}^0\tau)c^\dagger_{\mathbf{x}\sigma}\exp(-\mathcal{H}^0\tau),$$

where $\mathcal{H}^0 = \mathcal{H}^0_p + \mathcal{H}^0_{\text{ph}}$ and the evolution operator is given by

$$U(\beta) = T\exp\left(-\int_0^\beta d\tau H_{\text{int}}(\tau)\right), \quad (13)$$

x, **x**' are site indices; τ, τ' stand for the imaginary time with $0 < \tau < \beta$; T is the time-ordering operator; and β is the inverse temperature. The statistical average $\langle\ldots\rangle_0^c$ is evaluated with respect to the zero-order density matrix of the grand canonical ensemble of localized polarons and phonons,

$$\frac{\exp(-\beta\mathcal{H}^0)}{\mathrm{Tr}\exp(-\beta\mathcal{H}^0)} = \prod_i \frac{\exp(-\beta\mathcal{H}^0_{ip})}{\mathrm{Tr}\exp(-\beta\mathcal{H}^0_{ip})}\frac{\exp(-\beta\mathcal{H}^0_{i\text{ph}})}{\mathrm{Tr}\exp(-\beta\mathcal{H}^0_{i\text{ph}})}. \quad (14)$$

The superscript c in (12) indicates that only connected diagrams must be taken into account. Density matrix (14) is factorized with respect to the lattice sites. The phonon part is easily diagonalized using the free phonon operators b_i and $b_i^\dagger$, while the on-site polaron Hamiltonian contains the polaron–polaron interaction proportional to the renormalized parameter $\bar{U}$, which can only be diagonalized using Hubbard operators [18]. At this stage, no special assumption is made about the quantity $\bar{U}$ and its sign; we set up the equations of motion for the dynamical quantities in this general case, but investigate the equations in detail only in the special case where $\bar{U} = 0$.

The Wick theorem of weakly coupled quantum field theory can be used in evaluating statistical averages of phonon operators; e.g., the propagator of the phonon cloud,

$$\Phi(\tau_1|\tau_2) = \Phi(\tau_1 - \tau_2) \equiv \langle T\exp\{i\bar{g}[p(\tau_1) - p(\tau_2)]\}\rangle_0 \equiv \exp\left(-\frac{1}{2}\bar{g}^2\langle T[p(\tau_1) - p(\tau_2)]^2\rangle_0\right) = \exp(-\sigma(\beta) + \sigma(|\tau_1 - \tau_2|)), \quad (15)$$

$$\Phi(\tau_1, \tau_2|\tau_3, \tau_4) \equiv \langle T\exp\{i\bar{g}[p(\tau_1) + p(\tau_2) - p(\tau_3) - p(\tau_4)]\}\rangle_0 = \exp\left(-\frac{1}{2}\bar{g}^2\langle T[p(\tau_1) + p(\tau_2) - p(\tau_3) - p(\tau_4)]^2\rangle_0\right) = \exp\{\sigma(|\tau_1 - \tau_3|) + \sigma(|\tau_1 - \tau_4|) + \sigma(|\tau_2 - \tau_3|) + \sigma(|\tau_2 - \tau_4|) - \sigma(|\tau_1 - \tau_2|) - \sigma(|\tau_3 - \tau_4|) - 2\sigma(\beta)\}, \quad (16)$$

where

$$\sigma(|\tau_1 - \tau_2|) = \bar{g}^2 \langle T p(\tau_1) p(\tau_2) \rangle_0 = \alpha \frac{\cosh\left(\hbar\omega_0 \left\{\frac{\beta}{2} - |\tau_1 - \tau_2|\right\}\right)}{\sinh\left(\frac{\beta\hbar\omega_0}{2}\right)}. \quad (17)$$

We now discuss the problem of calculating chronological averages of combinations of polaron operators. Here, we use the above-mentioned new diagram technique and the GWT [10, 11]. The many-particle on-site irreducible Green's functions are the main element of diagrams in this approach.

3. POLARON AND PHONON GREEN'S FUNCTIONS

In the zero-order approximation, the one-polaron Green's function is given by

$$\mathcal{G}_p^0(x|x') = -\langle T c_{\mathbf{x}\sigma}(\tau) c_{\mathbf{x}'\sigma'}(\tau') \rangle_0 = -\langle T a_{\mathbf{x}\sigma}(\tau) \bar{a}_{\mathbf{x}'\sigma'}(\tau') \rangle_0 \Phi(\tau|\tau') = \mathcal{G}^{(0)}(x|x') \Phi(\tau|\tau'), \quad (18)$$

where $x = (\mathbf{x}, \sigma, \tau)$. The simplest new element of the diagram technique is the two-particle irreducible Green's function, or Kubo cumulant, which is equal to

$$\mathcal{G}_2^{(0)ir}(x_1, x_2|x_3, x_4) = \delta_{\mathbf{x}_1,\mathbf{x}_2}\delta_{\mathbf{x}_1,\mathbf{x}_3}\delta_{\mathbf{x}_1,\mathbf{x}_4} \times \mathcal{G}_2^{(0)ir}(\sigma_1, \tau_1; \sigma_2, \tau_2|\sigma_3, \tau_3; \sigma_4, \tau_4), \quad (19)$$

where

$$\begin{aligned}\mathcal{G}_2^{(0)ir}(\sigma_1, \tau_1; \sigma_2, \tau_2|\sigma_3, \tau_3; \sigma_4, \tau_4) &= \langle T c_{\sigma_1}(\tau_1) c_{\sigma_2}(\tau_2) \bar{c}_{\sigma_3}(\tau_3) \bar{c}_{\sigma_4}(\tau_4) \rangle_0 \\ &- \langle T c_{\sigma_1}(\tau_1) \bar{c}_{\sigma_4}(\tau_4) \rangle_0 \langle T c_{\sigma_2}(\tau_2) \bar{c}_{\sigma_3}(\tau_3) \rangle_0 \\ &+ \langle T c_{\sigma_1}(\tau_1) \bar{c}_{\sigma_3}(\tau_3) \rangle_0 \langle T c_{\sigma_2}(\tau_2) \bar{c}_{\sigma_4}(\tau_4) \rangle_0.\end{aligned} \quad (20)$$

The first term in the right-hand side of Eq. (20) is

$$\langle T c_{\sigma_1}(\tau_1) c_{\sigma_2}(\tau_2) \bar{c}_{\sigma_3}(\tau_3) \bar{c}_{\sigma_4}(\tau_4) \rangle_0 = \langle T a_{\sigma_1}(\tau_1) a_{\sigma_2}(\tau_2) \bar{a}_{\sigma_3}(\tau_3) \bar{a}_{\sigma_4}(\tau_4) \rangle_0 \Phi(\tau_1, \tau_2|\tau_3, \tau_4). \quad (21)$$

As the number of polaron operators increases, more complicated irreducible Green's functions like $\mathcal{G}_n^{(0)ir}(x_1 \ldots x_n|x_1' \ldots x_n')$ with $n \geq 3$ and all possible terms of their products appear. The sum of all strongly connected diagrams (i.e., those that cannot be divided into two parts by cutting a single hopping line) containing all kinds of irreducible Green's functions in the perturbation expansion of the evolution operator defines the special function $Z(x|x')$ (see [10, 11] for details). This function contains all contributions from charge and spin fluctuations. Together with the mass operator (which is the hopping matrix element in our case), it allows us to formulate a Dyson-type equation for the one-polaron Green's function [10–14],

$$\mathcal{G}(x|x') = \Lambda(x|x') + \sum_{1,2} \Lambda(x|1) t(1-2) \mathcal{G}(2|x'), \quad (22)$$

where

$$\Lambda(x|x') = \mathcal{G}_p^{(0)}(x|x') + Z(x|x'), \quad (23)$$

$$t(x - x') = \delta_{\sigma,\sigma'} \delta(\tau - \tau') t(\mathbf{x} - \mathbf{x}'). \quad (24)$$

Here, x again denotes $\mathbf{x}$, σ, τ, and the sum is over the discrete indices and includes integration over τ. Using the Fourier representation for these quantities,

$$\begin{aligned}\mathcal{G}_\sigma(\mathbf{x}|\tau) &= \frac{1}{N}\sum_{\mathbf{k}} \frac{1}{\beta} \\ &\times \sum_{\omega_n} \exp(-i\mathbf{k}\cdot\mathbf{x} - i\omega_n\tau) \mathcal{G}_\sigma(\mathbf{k}|i\omega_n), \\ \Lambda_\sigma(\mathbf{x}|\tau) &= \frac{1}{N}\sum_{\mathbf{k}} \frac{1}{\beta} \\ &\times \sum_{\omega_n} \exp(-i\mathbf{k}\cdot\mathbf{x} - i\omega_n\tau) \Lambda_\sigma(\mathbf{k}|i\omega_n), \\ \mathcal{G}_\sigma(\mathbf{x}|i\omega_n) &= \frac{1}{2}\int_{-\beta}^{\beta} d\tau \exp(i\omega_n\tau) \mathcal{G}_\sigma(\mathbf{x}|\tau), \\ \Lambda_\sigma(\mathbf{x}|i\omega_n) &= \frac{1}{2}\int_{-\beta}^{\beta} d\tau \exp(i\omega_n\tau) \Lambda_\sigma(\mathbf{x}|\tau),\end{aligned} \quad (25)$$

we obtain Dyson equation for the renormalized one-polaron Green's function,

$$\mathcal{G}_\sigma(\mathbf{k}|i\omega_n) = \frac{\Lambda_\sigma(\mathbf{k}|i\omega_n)}{1 - \varepsilon(\mathbf{k})\Lambda(\mathbf{k}|i\omega_n)}, \quad (26)$$

where

$$\omega_n = \frac{(2n+1)\pi}{\beta}$$

is the odd Matsubara frequency.

To discuss $\mathcal{G}_\sigma(k|i\omega)$ further, we need the Fourier representation of the zero-order one-polaron Green's function $\mathcal{G}_p^{(0)}$ defined in (18). In order to facilitate the investigation, we have evaluated the propagator of the phonon cloud (16) in the strong-coupling limit $\alpha \gg 1$ [15, 16, 26]:

$$\Phi(\tau) = \frac{1}{\beta}\sum_{\Omega_n} \exp(-i\Omega_n(\tau))\bar{\Phi}(i\Omega_n), \tag{27}$$

$$\bar{\Phi}(i\Omega_n) = \frac{\exp(-\sigma(\beta))}{2}\int_{-\beta}^{\beta} d\tau \exp(i\Omega_n\tau + \sigma(|\tau|)), \tag{28}$$

where $\Omega_n = 2n\pi/\beta$. To find $\bar{\Phi}(i\Omega_n)$, we use the Laplace approximation [32] for integral (28), which contains an exponential function with the parameter α. In the strong-coupling limit $\alpha \gg 1$, we obtain

$$\bar{\Phi}(i\Omega_n) \approx \frac{2\omega_c}{\Omega_n^2 + \omega_c^2}, \quad \omega_c = \hbar\alpha\omega_0 = \frac{g^2}{2\hbar\omega_0}. \tag{29}$$

This term is the harmonic propagator of the collective mode of phonons belonging to the polaron clouds. There are further terms describing anharmonic deviations. For $\alpha \gg 1$, these terms can be omitted because they are small compared with the harmonic contribution. Using the Laplace approximation [32] and

$$\Phi(\tau_1, \tau_2|\tau_3, \tau_4) = \frac{1}{\beta^4}\sum_{\Omega_1\ldots\Omega_4} \bar{\Phi}(i\Omega_1, i\Omega_2|i\Omega_3, i\Omega_4) \times \exp(-i\Omega_1\tau_1 - i\Omega_2\tau_2 + i\Omega_3\tau_3 + i\Omega_4\tau_4), \tag{30}$$

$$\bar{\Phi}(i\Omega_1, i\Omega_2|i\Omega_3, i\Omega_4) = \int_0^\beta \ldots \int_0^\beta d\tau_1 \ldots d\tau_4 \times \exp(i\Omega_1\tau_1 + i\Omega_2\tau_2 - i\Omega_3\tau_3 - i\Omega_4\tau_4)\Phi(\tau_1, \tau_2|\tau_3, \tau_4), \tag{31}$$

we then obtain the Fourier representation of the phonon correlation function,

$$\bar{\Phi}(i\Omega_1, i\Omega_2|i\Omega_3, i\Omega_4) \approx [\delta_{\Omega_1,\Omega_3}\delta_{\Omega_2,\Omega_4} + \delta_{\Omega_1,\Omega_4}\delta_{\Omega_2,\Omega_3}]\bar{\Phi}(i\Omega_1)\bar{\Phi}(i\Omega_2), \tag{32}$$

which corresponds to

$$\Phi(\tau_1, \tau_2|\tau_3, \tau_4) \approx \Phi(\tau_1|\tau_3)\Phi(\tau_2|\tau_4) + \Phi(\tau_1|\tau_4)\Phi(\tau_2|\tau_3). \tag{33}$$

This implies that in what follows, we can keep only the free collective oscillations of phonon clouds (29) surrounding the polarons and use the Hartree–Fock approximation: (32) and (33) for their two-particle correlation functions. In particular, we investigate the influence of the absorption and emission of this collective mode by polarons on the superconducting phase transition.

With the harmonic mode given by (29), the Fourier representation of the local polaron Green's function

$$\bar{\mathcal{G}}_{p\sigma}^{(0)}(i\omega_n) = \frac{1}{2}\int_{-\beta}^{\beta} d\tau \exp(i\omega_n\tau)\bar{\mathcal{G}}_{p\sigma}^{(0)}(\tau) \tag{34}$$

becomes

$$\bar{\mathcal{G}}_{p\sigma}^{(0)}(i\omega_n) \approx \frac{1}{Z_0} \times \left(\frac{\exp(-\beta E_0) + \bar{N}(\omega_c)(\exp(-\beta E_0) + \exp(-\beta E_\sigma))}{i\omega_n + E_0 - E_\sigma - \omega_c} \right.$$

$$+ \frac{\exp(-\beta E_\sigma) + \bar{N}(\omega_c)(\exp(-\beta E_0) + \exp(-\beta E_\sigma))}{i\omega_n + E_0 - E_\sigma + \omega_c} \tag{35}$$

$$+ \frac{\exp(-\beta E_{-\sigma}) + \bar{N}(\omega_c)(\exp(-\beta E_\sigma) + \exp(-\beta E_2))}{i\omega_n + E_{-\sigma} - E_2 - \omega_c}$$

$$\left. + \frac{\exp(-\beta E_2) + \bar{N}(\omega_c)(\exp(-\beta E_\sigma) + \exp(-\beta E_2))}{i\omega_n + E_{-\sigma} - E_2 + \omega_c} \right),$$

where

$$Z_0 = 1 + \exp(-\beta E_\sigma) + \exp(-\beta E_{-\sigma}) + \exp(-\beta E_2), \tag{36a}$$

$$E_0 = 0, \quad E_{\pm\sigma} = \epsilon, \quad E_2 = \bar{U} + 2\epsilon, \tag{36b}$$

$$\bar{n}(\epsilon) = (\exp(\beta\epsilon) + 1)^{-1}, \quad \bar{N}(\omega_c) = (\exp(\beta\omega_c) - 1)^{-1}. \tag{37}$$

Equation (35) shows that the on-site transition energies of polarons are changed by the collective-mode energy $\pm\omega_c$ of the phonon clouds. The delocalization of polarons due to their hopping between lattice sites

causes the broadening of the polaron energy levels. Equation (35) can be further simplified for a small on-site interaction energy $\bar{U}$ of polarons. For $\bar{U} = 0$, we obtain

$$\begin{aligned}\bar{\mathcal{G}}^{(0)}_{p\sigma}(i\omega_n|\epsilon) &= \frac{\bar{N}(\omega_c)+1-\bar{n}(\epsilon)}{i\omega_n-\epsilon-\omega_c}+\frac{\bar{N}(\omega_c)+\bar{n}(\epsilon)}{i\omega_n-\epsilon+\omega_c}\\ &= \frac{(i\omega_n-\epsilon)\coth(\beta\omega_c/2)+\omega_c\tanh(\beta\epsilon/2)}{(i\omega_n-\epsilon)^2-\omega_c^2}.\end{aligned} \tag{38}$$

This function has the antisymmetry property

$$\bar{\mathcal{G}}^{(0)}_{p\sigma}(-i\omega_n|-\epsilon) = -\bar{\mathcal{G}}^{(0)}_{p\sigma}(i\omega_n|\epsilon) \tag{39}$$

which also holds for the renormalized polaron quantities,

$$\begin{aligned}\Lambda_\sigma(-\mathbf{k}, -i\omega_n|-\epsilon) &= -\Lambda_\sigma(\mathbf{k}, i\omega_n|\epsilon),\\ \mathcal{G}_\sigma(-\mathbf{k}, -i\omega_n|-\epsilon) &= -\mathcal{G}_\sigma(\mathbf{k}, i\omega_n|\epsilon).\end{aligned} \tag{40}$$

Setting $\bar{U} \approx 0$, we assume that the strong on-site Coulomb repulsion of polarons can be canceled by the attraction induced by the strong electron–phonon interaction. We consider this as a model case that allows us a transparent discussion of the polarons exchanging phonon clouds during hopping between lattice sites.

4. TWO-PARTICLE IRREDUCIBLE CORRELATION FUNCTIONS

In what follows, we discuss the influence of a strong electron–phonon interaction on the two-particle irreducible Green's function. For $\bar{U} = 0$, the electronic correlation function in (22) is given by

$$\begin{aligned}&\langle Ta_{\sigma_1}(\tau_1)a_{\sigma_2}(\tau_2)\bar{a}_{\sigma_3}(\tau_3)\bar{a}_{\sigma_4}(\tau_4)\rangle_0\\ &= \langle Ta_{\sigma_1}(\tau_1)\bar{a}_{\sigma_4}(\tau_4)\rangle_0\langle Ta_{\sigma_2}(\tau_2)\bar{a}_{\sigma_3}(\tau_3)\rangle_0\\ &\quad -\langle Ta_{\sigma_1}(\tau_1)\bar{a}_{\sigma_3}(\tau_3)\rangle_0\langle Ta_{\sigma_2}(\tau_2)\bar{a}_{\sigma_4}(\tau_4)\rangle_0\end{aligned} \tag{41}$$

because the standard Wick theorem is now applicable. Using (33), we obtain the relation

$$\begin{aligned}&\mathcal{G}^{(0)ir}_2(\sigma_1,\tau_1;\sigma_2,\tau_2|\sigma_3,\tau_3;\sigma_4,\tau_4)\\ &= \delta_{\sigma_1,\sigma_4}\delta_{\sigma_2,\sigma_3}\mathcal{G}^{(0)}_{\sigma_1}(\tau_1-\tau_4)\mathcal{G}^{(0)}_{\sigma_2}(\tau_2-\tau_3)\\ &\quad\times\Phi(\tau_1-\tau_3)\Phi(\tau_2-\tau_4)\\ &\quad-\delta_{\sigma_1,\sigma_3}\delta_{\sigma_2,\sigma_4}\mathcal{G}^{(0)}_{\sigma_1}(\tau_1-\tau_3)\mathcal{G}^{(0)}_{\sigma_2}(\tau_2-\tau_4)\\ &\quad\times\Phi(\tau_1-\tau_4)\Phi(\tau_2-\tau_3)\end{aligned} \tag{42}$$

for the two-particle irreducible Green's function (21). In the absence of exchange of phonon clouds by polarons, this quantity must vanish. Indeed, if the electrons keep their initial phonon clouds during the time of propagation of two polarons, then the irreducible two-polaron Green's function (21) vanishes for $\bar{U} = 0$. However, because two electrons can be exchanged (independently of the exchange of phonon clouds), we obtain new contributions corresponding to two polarons with the exchanged phonon clouds. Alternatively, we can say that for $\bar{U} = 0$, the Wick theorem applies separately to free electrons and free phonons; however, it does not apply to polarons as composite particles and their cumulants do not therefore vanish.

The Fourier representation of (42),

$$\begin{aligned}&\mathcal{G}^{(0)ir}_2(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_3,i\omega_3;\sigma_4,i\omega_4)\\ &= \int_0^\beta\!\!\dots\int_0^\beta d\tau_1\dots d\tau_4\\ &\quad\times\mathcal{G}^{(0)ir}_2(\sigma_1,\tau_1;\sigma_2,\tau_2|\sigma_3,\tau_3;\sigma_4,\tau_4)\\ &\quad\times\exp(i\omega_1\tau_1+i\omega_2 t_2-i\omega_3\tau_3-i\omega_4\tau_4),\end{aligned} \tag{43}$$

is given by

$$\begin{aligned}&\mathcal{G}^{(0)ir}_2(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_3,i\omega_3;\sigma_4,i\omega_4)\\ &= \beta\delta_{\omega_1+\omega_2,\omega_3+\omega_4}\\ &\quad\times\bar{\mathcal{G}}^{(0)ir}_2(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_3,i\omega_3;\sigma_4,i\omega_4)\\ &= \beta\delta_{\omega_1+\omega_2,\omega_3+\omega_4}\{\delta_{\sigma_1,\sigma_4}\delta_{\sigma_2,\sigma_3}\\ &\quad\times A_{\sigma_1,\sigma_2}(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_2,i\omega_3;\sigma_1,i\omega_4)\\ &\quad-\delta_{\sigma_1,\sigma_3}\delta_{\sigma_2,\sigma_4}\\ &\quad\times A_{\sigma_1,\sigma_2}(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_1,i\omega_4;\sigma_2,i\omega_3)\},\end{aligned} \tag{44}$$

where

$$A_{\sigma_1,\sigma_2}(\sigma_1,i\omega_1;\sigma_2,i\omega_2|\sigma_2,i\omega_3;\sigma_1,i\omega_4) = \frac{1}{\beta}\sum_\Omega\frac{(2\omega_c)^2}{[i\omega_1-i\Omega-\epsilon][i\omega_3-i\Omega-\epsilon][\Omega^2+\omega_c^2][(\Omega+\Omega_1)^2+\omega_c^2]} \tag{45}$$

with $\Omega_1 = \omega_2 - \omega_3$. The summation leads to

$$A_{\sigma_1,\sigma_2}(\sigma_1, i\omega_1; \sigma_2, i\omega_2|\sigma_2, i\omega_3; \sigma_1, i\omega_4)$$
$$= 2(\omega_c)^2 \left\{ \frac{\tanh(\beta\epsilon/2)[i\omega_1 + i\omega_2 - 2\epsilon][2\omega_c^2 - (i\omega_1 - \epsilon)(i\omega_4 - \epsilon) - (i\omega_2 - \epsilon)(i\omega_3 - \epsilon)]}{[(i\omega_1 - \epsilon)^2 - \omega_c^2][(i\omega_2 - \epsilon)^2 - \omega_c^2][(i\omega_3 - \epsilon)^2 - \omega_c^2][(i\omega_4 - \epsilon)^2 - \omega_c^2]} \right.$$
$$- \frac{2\omega_c[i\omega_1 + i\omega_2 - 2\epsilon]^2 \coth(\beta\omega_c/2)[2\omega_c^2 - (i\omega_1 - \epsilon)(i\omega_3 - \epsilon) - (i\omega_2 - \epsilon)(i\omega_4 - \epsilon)]}{[(i\Omega_1)^2 - (2\omega_c)^2][(i\omega_1 - \epsilon)^2 - \omega_c^2][(i\omega_2 - \epsilon)^2 - \omega_c^2][(i\omega_3 - \epsilon)^2 - \omega_c^2][(i\omega_4 - \epsilon)^2 - \omega_c^2]} \quad (46)$$
$$\left. - \frac{\coth(\beta\omega_c/2)}{\omega_c[(i\Omega_1)^2 - (2\omega_c)^2]} \left[\frac{(i\omega_1 - \epsilon)(i\omega_3 - \epsilon) + 3\omega_c^2}{[(i\omega_1 - \epsilon)^2 - \omega_c^2][(i\omega_3 - \epsilon)^2 - \omega_c^2]} + \frac{(i\omega_2 - \epsilon)(i\omega_4 - \epsilon) + 3\omega_c^2}{[(i\omega_2 - \epsilon)^2 - \omega_c^2][(i\omega_4 - \epsilon)^2 - \omega_c^2]} \right] \right\}.$$

The function A_{σ_1,σ_2} contains contributions of the different spin channels to the two-particle on-site Green's function. The spin structure in Eq. (44) is due to the conservation law for the spins of the polarons.

5. SUPERCONDUCTING PHASE TRANSITION

In what follows, we check whether the polaronic system can have a superconducting instability in the absence of a direct attractive interaction for the polarons, i.e., for $\bar{U} = 0$. In this case, the attraction is only induced dynamically by polarons exchanging phonon clouds. To describe superconductivity, we need the anomalous propagators [33] in addition to the normal state Green's function (13). For simplicity, we limit the discussion to the *s*-wave superconductivity as in previous investigations of superconducting instabilities in the Hubbard model [13, 14] and in the Hubbard–Holstein model in the strong-coupling limit $\alpha \gg 1$ [26].

To describe the superconducting state, we need three irreducible functions Λ_σ, $Y_{\sigma,-\sigma}$, and $\bar{Y}_{-\sigma,\sigma}$ that represent infinite sums of diagrams containing irreducible many-particle Green's functions. In order to obtain a closed set of equations, we restrict ourselves to a class of rather simple contributions, which nevertheless contain the most important charge, spin, and pairing correlations; see [26] for details. This class of diagrams is obtained by neglecting contributions for which the Fourier representation of the superconducting order parameters $Y_{\sigma,-\sigma}$ and $\bar{Y}_{-\sigma,\sigma}$ depend on the polaron momentum **k**. In this approximation, $Y_{\sigma,-\sigma}$ is to be obtained from

$$Y_{\sigma,-\sigma}(i\omega) = -\frac{1}{\beta N}\sum_{\mathbf{k},\omega_l} \frac{\varepsilon(\mathbf{k})\varepsilon(-\mathbf{k})Y_{\sigma,-\sigma}(i\omega_l)}{D_\sigma(\mathbf{k}, i\omega_l)} \quad (47)$$
$$\times \bar{\mathcal{G}}_2^{(0)ir}(\sigma, i\omega; -\sigma, -i\omega|\sigma, i\omega_l; -\sigma, -i\omega_l).$$

In the same approximation, Λ_σ is to be computed from

$$\Lambda_\sigma(i\omega) = \mathcal{G}_{p\sigma}^{(0)}(i\omega) - \frac{1}{\beta N}\sum_{\mathbf{k},\omega_l} \frac{\varepsilon^2(\mathbf{k})}{D_\sigma(\mathbf{k}, i\omega_l)}$$
$$\times \{\Lambda_\sigma(i\omega_l)[1 - \varepsilon(-\mathbf{k})\Lambda_{-\sigma}(-i\omega_l)]$$
$$- \varepsilon(\mathbf{k})Y_{\sigma,-\sigma}(i\omega_l)\bar{Y}_{-\sigma,\sigma}(i\omega_l)\}$$
$$\times \bar{\mathcal{G}}_2^{(0)ir}(\sigma, i\omega; \sigma, i\omega_l|\sigma, i\omega_l; \sigma, i\omega) \quad (48)$$
$$- \frac{1}{\beta N}\sum_{\mathbf{k},\omega_l} \frac{\varepsilon^2(\mathbf{k})}{D_\sigma(\mathbf{k}, i\omega_l)}\{\Lambda_{-\sigma}(-i\omega_l)[1 - \varepsilon(\mathbf{k})\Lambda_\sigma(i\omega_l)]$$
$$- \varepsilon(-\mathbf{k})Y_{\sigma,-\sigma}(i\omega_l)\bar{Y}_{-\sigma,\sigma}(i\omega_l)\}$$
$$\times \bar{\mathcal{G}}_2^{(0)ir}(\sigma, i\omega; -\sigma, -i\omega_l|-\sigma, -i\omega_l; \sigma, i\omega),$$

with

$$D_\sigma(\mathbf{k}, i\omega) = [1 - \varepsilon(\mathbf{k})\Lambda_\sigma(i\omega)][1 - \varepsilon(-\mathbf{k})\Lambda_{-\sigma}(-i\omega)]$$
$$+ \varepsilon(\mathbf{k})\varepsilon(-\mathbf{k})Y_{\sigma,-\sigma}(i\omega)\bar{Y}_{-\sigma,\sigma}(i\omega). \quad (49)$$

The corresponding equation for $\bar{Y}_{-\sigma,\sigma}(i\omega)$ can be obtained from the expression for $Y_{\sigma,-\sigma}(i\omega)$.

Together with the equations for the one- and two-particle Green's functions, the above equations completely determine the properties of the superconducting phase, provided it exists. In order to gain further insight into the physics contained in (47) and (48), we linearize the equations in terms of the order parameter $Y_{\sigma,-\sigma}(i\omega)$ that determines the critical temperature T_c. The resulting equation for the order parameter is

$$Y_{\sigma,-\sigma}(i\omega) = -\frac{1}{\beta N}$$
$$\times \sum_{\mathbf{k},\omega_l} \frac{\varepsilon(\mathbf{k})\varepsilon(-\mathbf{k})Y_{\sigma,-\sigma}(i\omega_l)}{[1 - \varepsilon(\mathbf{k})\Lambda_\sigma(i\omega_l)][1 - \varepsilon(-\mathbf{k})\Lambda_{-\sigma}(-i\omega_l)]} \quad (50)$$
$$\times \bar{\mathcal{G}}_2^{(0)ir}(\sigma, i\omega; -\sigma, -i\omega|\sigma, i\omega_l; -\sigma, -i\omega_l).$$

This equation must be solved together with the equation for $\Lambda_\sigma(i\omega)$, which can be approximated by setting the order parameters to zero, with the result

$$\Lambda_\sigma(i\omega) = \mathcal{G}^{(0)}_{p\sigma}(i\omega) - \frac{1}{\beta N}\sum_{\mathbf{k},\,\omega_l}\frac{\varepsilon^2(\mathbf{k})\Lambda_\sigma(i\omega_l)}{1-\varepsilon(\mathbf{k})\Lambda_\sigma(i\omega_l)} \times \bar{\mathcal{G}}^{(0)ir}_2(\sigma, i\omega; \sigma, i\omega_l|\sigma, i\omega_l; \sigma, i\omega) - \frac{1}{\beta N}\sum_{\mathbf{k},\,\omega_l}\frac{\varepsilon^2(\mathbf{k})\Lambda_{-\sigma}(i\omega_l)}{1-\varepsilon(-\mathbf{k})\Lambda_{-\sigma}(i\omega_l)} \times \bar{\mathcal{G}}^{(0)ir}_2(\sigma, i\omega; -\sigma, i\omega_l|-\sigma, i\omega_l; \sigma, i\omega). \quad (51)$$

To determine T_c, we must solve (51) for Λ_σ and insert the result in (50). The irreducible functions in (50) and (51) can be written as

$$\bar{\mathcal{G}}^{(0)ir}_2(\sigma, i\omega; \sigma, i\omega_l|\sigma, i\omega_l; \sigma, i\omega) = \frac{(\omega-\omega_l)^2}{\Delta^2\Delta_l^2}\{2\omega_c^2(x+x_l)\tanh(\beta\epsilon/2) - \frac{\coth(\beta\omega_c/2)}{\omega_c[(i\omega-i\omega_l)^2-4\omega_c^2]} \times [(xx_l+\omega_c^2)(\Delta\Delta_l+8\omega_c^4) - 2\omega_c^2(\Delta+\Delta_l)(xx_l-\omega_c^2)]\}, \quad (52)$$

$$\bar{\mathcal{G}}^{(0)ir}_2(\sigma, i\omega; -\sigma, i\omega_l|-\sigma, i\omega_l; \sigma, i\omega) = -\frac{2\omega_c}{\Delta^2\Delta_l^2} \times \{\omega_c(x+x_l)(\Delta+\Delta_l)\tanh(\beta\epsilon/2) + \coth(\beta\omega_c/2)(x+x_l)^2(xx_l-\omega_c^2)\} + \frac{\coth(\beta\omega_c/2)(xx_l+3\omega_c^2)}{\omega_c\Delta\Delta_l}, \quad (53)$$

$$\bar{\mathcal{G}}^{(0)ir}_2(\sigma, i\omega; -\sigma, -i\omega|\sigma, i\omega_l; -\sigma, -i\omega_l) = -\frac{2\epsilon(2\omega_c)^2\tanh(\beta\epsilon/2)[i\omega i\omega_l+\epsilon^2-\omega_c^2]}{[\omega^2+(\epsilon+\omega_c)^2][\omega^2+(\epsilon-\omega_c)^2][\omega_l^2+(\epsilon+\omega_c)^2][\omega_l^2+(\epsilon-\omega_c)^2]} + \frac{2\omega_c\coth(\beta\omega_c/2)[i\omega i\omega_l+2\omega_c(\epsilon-\omega_c)-(\epsilon-\omega_c)^2]}{[\omega^2+(\epsilon-\omega_c)^2][\omega_l^2+(\epsilon-\omega_c)^2][(\omega-\omega_l)^2+(2\omega_c)^2]} + \frac{2\omega_c\coth(\beta\omega_c/2)[i\omega i\omega_l-2\omega_c(\epsilon+\omega_c)-(\epsilon+\omega_c)^2]}{[\omega^2+(\epsilon+\omega_c)^2][\omega_l^2+(\epsilon+\omega_c)^2][(\omega-\omega_l)^2+(2\omega_c)^2]}, \quad (54)$$

where

$$x = i\omega-\epsilon, \quad \Delta = (i\omega-\epsilon)^2-\omega_c^2, \quad (55a)$$

$$x_l = i\omega_l-\epsilon, \quad \Delta_l = (i\omega_l-\epsilon)^2-\omega_c^2. \quad (55b)$$

To analyze (50) and (51) further, we introduce the notation

$$\phi_\sigma(i\omega) = \frac{1}{N}\sum_{\mathbf{k}}\frac{\varepsilon^2(\mathbf{k})\Lambda_\sigma(i\omega)}{1-\varepsilon(\mathbf{k})\Lambda_\sigma(i\omega)} = \frac{1}{N}\sum_{\mathbf{k}}\frac{\varepsilon(\mathbf{k})}{1-\varepsilon(\mathbf{k})\Lambda_\sigma(i\omega)}, \quad (56)$$

$$g_\sigma(i\omega) = \mathcal{G}_\sigma(\mathbf{x}=\mathbf{x}'|i\omega) = \frac{1}{N}\sum_{\mathbf{k}}\frac{\Lambda_\sigma(i\omega)}{1-\varepsilon(\mathbf{k})\Lambda_\sigma(i\omega)}, \quad (57)$$

$$\phi^{sc}_\sigma(i\omega) = \frac{1}{N}\sum_{\mathbf{k}}\frac{\varepsilon(\mathbf{k})\varepsilon(-\mathbf{k})}{[1-\varepsilon(\mathbf{k})\Lambda_\sigma(i\omega)][1-\varepsilon(-\mathbf{k})\Lambda_{-\sigma}(-i\omega)]} = \frac{\phi_\sigma(i\omega)-\phi_{-\sigma}(-i\omega)}{\Lambda_\sigma(i\omega)-\Lambda_{-\sigma}(-i\omega)}. \quad (58)$$

We also assume that $\varepsilon(\mathbf{k}) = \varepsilon(-\mathbf{k})$ holds with

$$\sum_{\mathbf{k}}\varepsilon(\mathbf{k}) = \sum_{\mathbf{k}}\varepsilon^3(\mathbf{k}) = 0.$$

We replace sums with integrals,

$$\frac{1}{N}\sum_{\mathbf{k}} = \int d\varepsilon \rho_0(\varepsilon), \tag{59}$$

$$\rho_0(\varepsilon) = \frac{4}{\pi W}\sqrt{1-\left(\frac{2\varepsilon}{W}\right)^2}\begin{cases}1, & |\varepsilon| < \frac{W}{2}\\ 0, & |\varepsilon| > \frac{W}{2},\end{cases} \tag{60}$$

where W is the bandwidth and ρ_0 is the model density of states of a semielliptic form. Since we do not consider magnetic states here, the spin subscript can be omitted in the paramagnetic phase,

$$\Lambda_\sigma(i\omega) = \Lambda_{-\sigma}(i\omega) = \Lambda(i\omega), \tag{61a}$$

$$\phi_\sigma(i\omega) = \phi_{-\sigma}(i\omega) = \phi(i\omega), \tag{61b}$$

$$\phi^{sc}_\sigma(i\omega) = \phi^{sc}_{-\sigma}(i\omega) = \phi^{sc}(i\omega). \tag{61c}$$

However, the spin subscript is essential for the superconducting order parameter $Y_{\sigma,-\sigma}(i\omega)$,

$$Y_{\sigma,-\sigma}(i\omega) = g_{\sigma,-\sigma}Y(i\omega), \tag{62a}$$

$$g_{\sigma,-\sigma} = \delta_{\sigma,\uparrow} - \delta_{\sigma,\downarrow}, \tag{62b}$$

where $Y(i\omega)$ is an even function of the frequency,

$$Y(i\omega) = Y(-i\omega). \tag{63}$$

We finally add the equation that determines the chemical potential,

$$\frac{1}{\beta}\sum_{\omega_n}\sum_{\sigma}\mathcal{G}_\sigma(\mathbf{x}=\mathbf{x}'|i\omega)\exp(i\omega_n 0^+) = \frac{2}{\beta}\sum_{\omega_n} g_\sigma(i\omega)\exp(i\omega_n 0^+) = \frac{N_p}{N}, \tag{64}$$

where N_p is the number of polarons and N is the number of lattice sites. With (59) and (60), functions (56) and (57) can be written as

$$\phi(i\omega) = \frac{W}{2}\frac{\left(1-\sqrt{1-\lambda^2(i\omega)}\right)^2}{\lambda^3(i\omega)} = \frac{W}{2}\frac{\lambda(i\omega)}{\left(1+\sqrt{1-\lambda^2(i\omega)}\right)^2}, \tag{65}$$

$$g(i\omega) = \frac{4}{W}\frac{\left(1-\sqrt{1-\lambda^2(i\omega)}\right)^2}{\lambda(i\omega)} = \frac{4}{W}\frac{\lambda(i\omega)}{\left(1+\sqrt{1-\lambda^2(i\omega)}\right)^2}, \tag{66}$$

where

$$\lambda(i\omega) = (W/2)\Lambda(i\omega).$$

In order to check whether the state determined from Eq. (51) is metallic or dielectric, we must analyze the renormalized density of states given by

$$\rho(E) = -\frac{1}{\pi}\mathrm{Im}g(E+i0^+) = -\frac{1}{\pi}\mathrm{Im}\left(\frac{1-\sqrt{1-\lambda^2(E+i0^+)}}{\lambda(E+i0^+)}\right), \tag{67}$$

where $\lambda(E+i0^+)$ is the analytic continuation of $\lambda(i\omega)$.

6. ANALYTIC SOLUTIONS

The expressions for $\Lambda(i\omega)$ and $Y(i\omega)$ can be simplified using notation (56) and (57) and symmetry property (62),

$$Y(i\omega) = \frac{1}{\beta}\sum_{\omega_l}\phi^{sc}(i\omega_l) \times \bar{\mathcal{G}}^{(0)ir}(\sigma, i\omega; -\sigma, -i\omega|\sigma, i\omega_l; -\sigma_l, -i\omega_l)Y(i\omega_l), \tag{68}$$

$$\Lambda(i\omega) = \mathcal{G}^{(0)}_p(i\omega) - \frac{1}{\beta}\sum_{\omega_l}\phi(i\omega_l)[\bar{\mathcal{G}}^{(0)ir}(\sigma, i\omega; \sigma, i\omega_l|\sigma, i\omega_l; \sigma, i\omega) + \bar{\mathcal{G}}^{(0)ir}(\sigma, i\omega; -\sigma, i\omega_l|-\sigma, i\omega_l; \sigma, i\omega)]. \tag{69}$$

To find the solution of Eq. (68), we insert (54), replace $Y(i\omega_n)$ with

$$Y(i\omega_n) = \phi^{sc}(z_0)\chi(i\omega_n)Y(z_0), \quad z_0 \approx 0, \tag{70}$$

$$\chi(i\omega_n) = \frac{1}{\beta}\sum_{\omega_l}\bar{\mathcal{G}}^{(0)ir}(\sigma, i\omega_n; -\sigma, -i\omega_n|\sigma, i\omega_l; -\sigma, -i\omega_l) = \frac{2\omega_c[\omega_c - \epsilon\tanh(\beta\epsilon/2)\cosh(\beta\omega_c/2)] + \cosh(\beta\omega_c)(-\omega_c^2+\epsilon^2+\omega_n^2)(\cosh(\beta\omega_c)-1)^{-1}}{[\omega_n^2+(\omega_c+\epsilon)^2][\omega_n^2+(\omega_c-\epsilon)^2]}, \tag{71}$$

and use the Poisson summation formula

$$\frac{1}{\beta}\sum_{\omega_n} f(i\omega_n) = -\frac{1}{2\pi i}\int_C dz \frac{f(z)}{e^{\beta z}+1}, \tag{72}$$

where C denotes the usual counterclockwise contour of the imaginary axis. With the help of the analytically continued function $\chi(z)$ for $Z = Z_0 = 0$, we then obtain an equation for the critical temperature T_c from Eq. (68),

$$\chi(0|\epsilon)\phi^{sc}(0|\epsilon) = 1, \tag{73}$$

$$\chi(0|\epsilon) = \Big(2\omega_c[\omega_c - \epsilon\tanh(\beta_c\epsilon/2)\coth(\beta_c\omega_c/2)] + \frac{(\epsilon^2-\omega_c^2)\cosh(\beta_c\omega_c)}{\cosh(\beta_c\omega_c)-1}\Big)(\omega_c^2-\epsilon^2)^{-2}. \tag{74}$$

This quantity is even in ϵ, and therefore only the absolute value of $\epsilon = \bar{\epsilon}_0 - \bar{\mu}$ determines $k_B T_c = \beta_c^{-1}$. From (58) and (65), we can make a rough guess on the quantity $\phi^{sc}(0)$,

$$\phi^{sc}(0) \approx \left(\frac{W}{4}\right)^2 \frac{1}{\gamma^2}, \qquad \gamma = \frac{1}{2}(1+\sqrt{1-\lambda^2(0+i\delta)}), \tag{75}$$

where γ must satisfy the condition $\gamma(-\epsilon) = \gamma(\epsilon)$. This quantity can be obtained self-consistently from Eq. (64) for the chemical potential. For simplicity, we here replace $[1+\sqrt{1-\lambda^2(0+i\delta)}\,]$ with 2γ. Then (64) can be written as

$$\frac{2}{\gamma}\frac{1}{\beta}\sum_{\omega_n}\Lambda(i\omega_n)\exp(i\omega_n 0^+) = \frac{N_p}{N}. \tag{76}$$

Using (69) together with (52) and (53), we can express $\Lambda(i\omega_n)$ as

$$\Lambda(i\omega_n) = \frac{(i\omega_n-\epsilon)A_1(\epsilon)+\omega_c B_1(\epsilon)}{(i\omega_n-\epsilon)^2-\omega_c^2} + \frac{\omega_c^2[(i\omega_n-\epsilon)A_2(\epsilon)+\omega_c B_2(\epsilon)]}{[(i\omega_n-\epsilon)^2-\omega_c^2]^2}, \tag{77}$$

with unknown coefficients A_i and B_i. They can be found from Eq. (69) or more easily from the asymptotic behavior of (77) as $|\omega_n| \longrightarrow \infty$,

$$\underset{|\omega_n|\to\infty}{\Lambda(i\omega_n)} = \frac{A_1}{i\omega_n} + \frac{A_1\epsilon+\omega_c B_1}{(i\omega_n)^2} + \frac{A_1(\omega_c^2+\epsilon^2)+2\epsilon\omega_c B_1+\omega_c^2 A_2}{(i\omega_n)^3} + \frac{A_1(\epsilon^3+3\epsilon\omega_c^2)+B_1(\omega_c^3+3\epsilon^2\omega_c)+\omega_c^2(3\epsilon A_2+\omega_c B_2)}{(i\omega_n)^4}+\dots \tag{78}$$

If we compare this with the asymptotic behavior of the full one-polaron Green's function (see the Appendix) by invoking the methods of moments together with the asymptotic behavior of $g(i\omega_n)$ in (66), we obtain

$$A_1(\epsilon) = 1, \tag{79a}$$

$$B_1(\epsilon) = -\frac{1}{\omega_c}[M_1+\epsilon], \tag{79b}$$

$$A_2(\epsilon) = \frac{1}{\omega_c^2}\left[M_2+2\epsilon M_1+\epsilon^2-\omega_c^2-\left(\frac{W}{4}\right)^2\right], \tag{79c}$$

$$B_2(\epsilon) = \frac{1}{\omega_c^3}\times\left[-M_3-3\epsilon M_2+M_1\left(\omega_c^2-3\epsilon^2+3\left(\frac{W}{4}\right)^2\right)+\epsilon\omega_c^2-\epsilon^3+3\epsilon\left(\frac{W}{4}\right)^2\right], \tag{79d}$$

where M_i is the ith moment of the one-polaron Green's function. The results in (A.5) for the moments in the lowest order allow us to evaluate A_i and B_i, see (A.7). $A_1 = 1$ describes the asymptotic freedom of the polarons. $B_1 = \tanh(\beta\epsilon/2)$ is identical to its value in the zero-order polaron Green's function (38). The two new quantities A_2 and B_2 are small, being proportional to $\omega_0/\omega_c = 1/\alpha$.

Inserting (77) in the left-hand side of Eq. (76) and performing the summation, we obtain

$$\frac{1}{\beta}\sum_{\omega_n}\Lambda(i\omega_n)\exp(i\omega_n 0^+) = \bar{n}(\epsilon) + \frac{\tanh(\beta\epsilon/2)[\tanh(\beta\omega_c/2)-1][1-\tanh^2(\beta\epsilon/2)]}{2[1-\tanh^2(\beta\omega_c/2)\tanh^2(\beta\epsilon/2)]} + \frac{B_2(\epsilon)}{4}\tanh(\beta\omega_c/2)\frac{1-\tanh^2(\beta\epsilon/2)}{1-\tanh^2(\beta\epsilon/2)\tanh^2(\beta\omega_c/2)} \tag{80}$$

$$-\frac{\beta\omega_c}{16}\left[\frac{A_2(\epsilon)+B_2(\epsilon)}{\cosh^2[\beta(\omega_c+\epsilon)/2]}-\frac{A_2(\epsilon)-B_2(\epsilon)}{\cosh^2[\beta(\omega_c-\epsilon)/2]}\right],$$

which is equal to $(\gamma/2)(N_p/N)$ in accordance with Eq. (76). Because the collective frequency is large, $\beta\omega_c \gg 1$, we can omit exponentially small quantities like $\exp(-\beta\omega_c)$. Since we are interested in the results for electron numbers that are close to half-filling ($\epsilon = 0$), also $|\epsilon| \ll \omega_c$ holds. We also neglect contributions on the order of $1/\alpha$. Then the equation for chemical potential (74) is simply

$$\bar{n}(\epsilon) = \gamma n_p/2, \quad n_p = N_p/N. \tag{81}$$

If we set $\gamma = 1$ (free polarons), we obtain from (75) that

$$\phi^{sc}(0) = (W/4)^2, \tag{82}$$

which allows us to write the equation for the critical temperature T_c as

$$\epsilon^2+\omega_c^2-2\epsilon\omega_c\tanh(\beta\epsilon/2) = (\omega_c^2-\epsilon^2)^2(4/W)^2. \tag{83}$$

In this approximation, T_c depends only on the local parameters, but we expect that close to half-filling, this should give an indication of which of the local quantities is most important for superconductivity in the strong-coupling limit of the Hubbard–Holstein model. Precisely at half-filling, Eq. (83) can only be satisfied if $\omega_c = W/4$. This might be an unphysically large value for a renormalized quantity. It also shows that the specific limit $\bar{U} = 0$ is probably the critical value for the occurrence of superconductivity in the framework of the Hubbard–Holstein model. It is clear that superconductivity is possible for $\bar{U} < 0$, but in this case, it would have to compete in energy with the energies of charge-ordered states.

For the special case where $\omega_c = W/4$, we obtain

$$\left(\frac{\epsilon}{\omega_c}\right)^2\left[3-\left(\frac{\epsilon}{\omega_c}\right)^2\right] = \frac{2|\epsilon|}{\omega_c}\tanh\left(\frac{\beta_c|\epsilon|}{2}\right). \tag{84}$$

Because $\epsilon/\omega_c < 3$ holds (which we do not discuss in detail), we can seek solutions in the case where $|\epsilon| \ll \omega_c$, leading to

$$\begin{aligned} k_B T_c &= \frac{\omega_c}{3}\left[1-\frac{5}{12}\left(\frac{\epsilon}{\omega_c}\right)^2+\ldots\right] \\ &= \frac{W}{12}\left[1-\frac{20}{3}\left(\frac{\epsilon}{W}\right)^2+\ldots\right]. \end{aligned} \tag{85}$$

In spite of the many approximations used (all of which are reasonable, however) the result for T_c is remarkable because it shows that the critical temperature depends on the bandwidth (corresponding to the largest cutoff energy of the model) and not on the effective mass of the ions. For small deviations from half-filling, T_c decreases and is independent of the sign of ϵ.

For different values of ω_c,

$$\omega_c = W/4 - y, \tag{86}$$

with $y \neq 0$, there are only solutions other than at half-filling. In this case, Eq. (84) can be written as

$$\beta_c|\epsilon| = \ln\frac{1+\kappa}{1-\kappa}, \tag{87}$$

$$\kappa = \frac{\epsilon^2+\omega_c^2-(4/W)^2(\omega_c^2-\epsilon^2)^2}{2|\epsilon|\omega_c}, \quad 0<\kappa<1. \tag{88}$$

The condition $\kappa < 1$ is equivalent to

$$|\epsilon| + \omega_c < W/4. \tag{89}$$

On the other hand, the condition $\kappa > 0$ is reformulated differently depending on the parameter y:

$$\omega_c < W/4, \quad y>0\text{: } (W/4-\omega_c)^2 < \epsilon^2 < \epsilon_{\max}^2, \tag{90}$$

$$\omega_c > W/4, \quad y<0\text{: } \epsilon_{\min}^2 < \epsilon^2 < \epsilon_{\max}^2, \tag{91}$$

$$\epsilon^2_{\max,\,\min} = \omega_c^2+\frac{1}{2}\left(\frac{W}{4}\right)^2 \pm \left(\frac{W}{8}\right)\sqrt{\left(\frac{W}{4}\right)^2+8\omega_c^2}. \tag{92}$$

For small y, we can simplify (87) and (88) as

$$\begin{aligned} \kappa \approx \frac{2}{W|\epsilon|}\Bigg\{&\epsilon^2\left[3-\epsilon^2\left(\frac{4}{W}\right)^2\right] \\ &+ y\left[\frac{W}{2}-\frac{4\epsilon^2}{W}-\epsilon^4\left(\frac{4}{W}\right)^3\right]\Bigg\}, \end{aligned} \tag{93}$$

with the following restrictions for ϵ:

$$y>0\text{: } y^2 < \epsilon^2 < 3\left(\frac{W}{4}\right)^2 - \frac{5}{6}Wy + \frac{29}{27}y^2, \tag{94}$$

$$y<0\text{: } \frac{W}{6}|y| + \frac{25}{27}y^2 < \epsilon^2 < 3\left(\frac{W}{4}\right)^2 + \frac{5}{6}W|y| + \frac{29}{27}y^2. \tag{95}$$

Large values of T_c can be achieved for $\kappa \not\ll 1$ and in the vicinity of half-filling ($\epsilon \neq 0$),

$$\begin{aligned} k_B T_c &\approx \frac{W\delta}{12(\delta-1)}, \quad \delta = \frac{6\epsilon^2}{W|y|} > 1, \\ y &= \frac{W}{4}-\omega_c, \end{aligned} \tag{96}$$

but only for $y < 0$, and hence, $\omega_c > W/4$.

7. CONCLUSIONS

We have discussed the occurrence of superconductivity in the strong-coupling limit ($\bar{g} \gg 1$) of the Hub-

bard–Holstein model. Strong coupling leads to a renormalization of the one-polaron Green's function already in the local approximation. For $\bar{g} \gg 1$, we found a collective mode for the phonon clouds estimated by evaluating integrals in the Laplace approximation. Because of the absorption and emission of this mode by polarons, the on-site energies of polarons are renormalized. Similarly the irreducible two-particle Green's functions are renormalized. Allowance for the exchange of polarons, including their phonon clouds, leads to a new irreducible Green's function that has been used to study spin–singlet pairing of polarons. Analytic results for the superconducting phase have been obtained in the limiting case where the local repulsion of polarons is exactly canceled by their attractive interaction. The resulting equation for the critical temperature has been obtained by assuming a large collective-mode frequency and a nearly half-filled band case. The parameters that determine T_c are ω_c ($\omega_c \geq W/4$), ϵ (with $\epsilon = 0$ corresponding to half-filling), and the bandwidth W. In the strong-coupling limit, we obtain a critical temperature on the order of $\omega_c/3$.

It is interesting to note that a similar result for the value of T_c has been established in [34, 35] for such anomalous low-temperature superconductors as Pb, Hg, and Nb realized within the framework of Eliashberg's theory [36].

In the Eliashberg theory, the retarded nature of the photon-induced interaction and the pseudopotential treatment of the screened Coulomb interaction are taken into account. For example, in [35], the maximum value of the critical temperature $T_c^{\max}$ is equal to $\langle\omega\rangle/\exp(3/2)$, where $\langle\omega\rangle$ is the average phonon frequency taken over the phonon density of states $\langle\omega\rangle \approx 0.5\omega_0$. Our equations involve only the collective frequency $\omega_c = \alpha\omega_0$, $\alpha > 1$.

It is possible to estimate the values of T_c not only analytically, but also via calculations using Eq. (83), or mere precisely, Eq. (73). Indeed, such numerical results have been obtained by assuming certain values of the theory's parameters and of the interval of interesting values of T_c. From the three parameters in our theory (ω_c, *W*, and *e*), we first choose the value of the collective frequency ω_c. Assuming that ω_0 is equal to 0.07 eV for cuprates and that the dimensionless interaction constant $\bar{q}$ is equal to 3, we obtain $\alpha = 4.5$ and $\omega_c = 0.315$ eV. Next we fix the value of T_c, e.g., as equal to 100 K. With these values of ω_c and T_c, the following pairs of the other two parameters are compatible:

$e = 0.10515$ eV and $W = 1.68057$ eV,

$e = 0.20348$ eV and $W = 2.07383$ eV,

$e = 0.30149$ eV and $W = 2.46594$ eV.

ACKNOWLEDGMENTS

This work was supported by the Heisenberg–Landau Program and by the Program of the High Council for Science and Technical Development of Moldova. We gratefully acknowledge discussions with Prof. Plakida and Prof. Crisan. V.A.M. would like to thank the universities of Duisburg and Salerno for financial support. P.E. would like to thank the Bogoliubov Laboratory of Theoretical Physics, JINR, for the hospitality that he received during his stay in Dubna.

APPENDIX

The Method of Moments

Using the Heisenberg representation of the one-polaron Green's function

$$\mathcal{G}_c(\mathbf{x}-\mathbf{x}'|\tau-\tau') = \langle T\hat{c}_{\mathbf{x}\sigma}(\tau)\hat{\bar{c}}_{\mathbf{x}'\sigma'}(\tau')\rangle_H, \quad \text{(A.1)}$$

where

$$\hat{c}_{\mathbf{x}\sigma}(\tau) = e^{\mathcal{H}\tau}c_{\mathbf{x}\sigma}e^{-\mathcal{H}\tau}, \quad \text{(A.2a)}$$

$$\hat{\bar{c}}_{\mathbf{x}\sigma}(\tau) = e^{\mathcal{H}\tau}c^{\dagger}_{\mathbf{x}\sigma}e^{-\mathcal{H}\tau}, \quad \text{(A.2b)}$$

we can write the asymptotic expansion of the Fourier representation in (25) for $|\omega_n| \longrightarrow \infty$ as

$$\mathcal{G}_\sigma(\mathbf{x}=0|i\omega_n) = g_\sigma(i\omega_n) = \frac{1}{i\omega_n} - \frac{M_1}{(i\omega_n)^2} + \frac{M_2}{(i\omega_n)^3} - \frac{M_3}{(i\omega_n)^4} + \ldots, \quad \text{(A.3)}$$

$$M_n = \langle\{c^{\dagger}_{\mathbf{x}\sigma}, \underbrace{[\mathcal{H}[\mathcal{H}\ldots[\mathcal{H}, c_{\mathbf{x}\sigma}]\ldots]\ldots]}_{n}\}\rangle_H, \quad \text{(A.4)}$$

where the statistical average $\langle\ldots\rangle_H$ is defined with respect to the full density matrix of the grand canonical ensemble. In the simplest approximation, we obtain the first three moments of the Green's functions as

$$M_1 = -(\epsilon + \omega_c\tanh(\beta\epsilon/2)), \quad \text{(A.5a)}$$

$$M_1 = \epsilon^2 + \omega_c^2 + (W/4)^2 + 2\epsilon\omega_c\tanh(\beta\epsilon/2) + \omega_0\omega_c\coth(\beta\omega_c/2), \quad \text{(A.5b)}$$

$$M_3 = -\{\epsilon^3 + 3\epsilon[\omega_c^2 + \omega_c\omega_0\cosh(\beta\omega_c/2) + (W/4)^2] + \omega_c\tanh(\beta\epsilon/2)[3\epsilon^2 + 3(W/4)^2 + \omega_c^2 + \omega_0^2 + 3\omega_0\omega_c\coth(\beta\omega_0/2)]\}. \quad \text{(A.5c)}$$

The expressions for the moments can be used to determine the unknown coefficients $A_n(\epsilon)$ and $B_n(\epsilon)$ in the relation for $\Lambda_\sigma(i\omega)$, Eq. (78), by also considering the

asymptotic behavior of $g_\sigma(i\omega)$ in (66) for small values of $\lambda_\sigma(i\omega)$,

$$g_\sigma(i\omega) = (2/W)\lambda_\sigma(i\omega) \times [1 + (\lambda^2/4) + 2(\lambda^2/4)^2 + \ldots]. \quad \text{(A.6)}$$

We then insert the asymptotic form of $\lambda_\sigma(i\omega)$ from (78). Comparison of the corresponding equations fixes the coefficients $A_n(\epsilon)$ and $B_n(\epsilon)$ as

$$A_1(\epsilon) = 1, \quad \text{(A.7a)}$$

$$B_1(\epsilon) = -\frac{1}{\omega_c}[M_1 + \epsilon] \approx \tanh(\beta\epsilon/2), \quad \text{(A.7b)}$$

$$A_2(\epsilon) = \frac{1}{\omega_c^2}[M_2 + 2\epsilon M_1 + \epsilon^2 - \omega_c^2 - (W/4)^2] \approx \frac{\omega_0}{\omega_c}\coth(\beta\omega_0/2), \quad \text{(A.7c)}$$

$$B_2(\epsilon) = \frac{1}{\omega_c^3}[-M_3 - 3\epsilon M_2 + M_1(\omega_c^2 - 3\epsilon^2 + 3(W/4)^2) + \epsilon\omega_c^2 - \epsilon^3 + 3\epsilon(W/4)^2] \approx \frac{\omega_0}{\omega_c}\tanh(\beta\epsilon/2)\left[\frac{\omega_0}{\omega_c} + 3\coth(\beta\omega_0/2)\right]. \quad \text{(A.7d)}$$

REFERENCES

1. J. G. Bednorz and K. A. Müller, Z. Phys. B **64**, 189 (1986).
2. E. Dagotto, Rev. Mod. Phys. **66**, 763 (1994).
3. A. P. Kampf, Phys. Rep. **249**, 219 (1994).
4. D. J. Scalapino, Phys. Rep. **250**, 329 (1995).
5. W. Brenig, Phys. Rep. **251**, 153 (1995).
6. Z.-X. Shen and D. S. Dessau, Phys. Rep. **253**, 1 (1995).
7. C. Thomson and M. Cardona, in *Physical Properties of High Temperature Superconductors*, Ed. by D. M. Ginsberg (World Sci., Singapore, 1989; Mir, Moscow, 1990), Vol. 1, p. 409.
8. C. Frank, in *Physical Properties of High Temperature Superconductors*, Ed. by D. M. Ginsberg (World Sci., Singapore, 1993), Vol. 4, p. 189.
9. S. I. Vedneev, P. Samuely, S. V. Meshkov, *et al.*, Physica C (Amsterdam) **198**, 47 (1992); N. Miyakawa, Y. Shiina, T. Taneko, and N. Tsuda, J. Phys. Soc. Jpn. **62**, 2445 (1993); Y. Ohyagi, D. Shimada, N. Miyakawa, *et al.*, J. Phys. Soc. Jpn. **64**, 3376 (1995).
10. M. I. Vladimir and V. A. Moskalenko, Theor. Math. Phys. **82**, 301 (1990).
11. S. I. Vakaru, M. I. Vladimir, and V. A. Moskalenko, Theor. Math. Phys. **85**, 1185 (1990).
12. V. A. Moskalenko, S. P. Cojocaru, and M. I. Vladimir, IC/94/182 of ICTP Trieste.
13. N. N. Bogoliubov and V. A. Moskalenko, Theor. Math. Phys. **86**, 10 (1991).
14. N. N. Bogoliubov and V. A. Moskalenko, Theor. Math. Phys. **92**, 820 (1992).
15. V. A. Moskalenko, Theor. Math. Phys. **111**, 744 (1997).
16. V. A. Moskalenko, L. A. Dogotaru, and I. G. Porcescu, Physica C (Amsterdam) **282–287**, 1719 (1997).
17. V. A. Moskalenko, Theor. Math. Phys. **113**, 432 (1997).
18. J. Hubbard, Proc. R. Soc. London, Ser. A **276**, 233 (1963).
19. T. Holstein, Ann. Phys. (N.Y.) **8**, 325 (1959).
20. J. K. Freericks and M. Jarrell, Phys. Rev. B **50**, 6939 (1994).
21. A. S. Alexandrov and N. F. Mott, *Polarons and Bipolarons* (World Sci., Singapore, 1996).
22. A. S. Alexandrov, in *Lectures on the Physics of Highly Correlated Electron Systems V*, Ed. by F. Mancini (Melville, New York, 2001), AIP Conf. Proc., Vol. 580, p. 1.
23. M. Mierzejewski, J. Zielinski, and P. Entel, Phys. Rev. B **57**, 590 (1998).
24. I. G. Lang and Yu. A. Firsov, Zh. Éksp. Teor. Fiz. **16**, 1301 (1963) [Sov. Phys. JETP **3**, 18431 (1962)].
25. V. A. Moskalenko, P. Entel, and D. F. Digor, Phys. Rev. B **59**, 619 (1999).
26. V. A. Moskalenko, P. Entel, and D. F. Digor, Physica B (Amsterdam) **259–261**, 781 (1999).
27. J. Hubbard, Proc. R. Soc. London, Ser. A **296**, 82 (1967).
28. P. M. Slobodyan and I. V. Stasyuk, Theor. Math. Phys. **19**, 423 (1974).
29. R. O. Zaĭtsev, Zh. Éksp. Teor. Fiz. **70**, 1100 (1976) [Sov. Phys. JETP **43**, 574 (1976)]; R. O. Zaĭtsev, Fiz. Tverd. Tela (Leningrad) **19**, 1097 (1977) [Sov. Phys. Solid State **19**, 1874 (1977)].
30. Yu. A. Izyumov and B. M. Letfulov, J. Phys.: Condens. Matter **2**, 8905 (1990); Yu. A. Izyumov, B. M. Letfulov, E. V. Shipitsyn, *et al.*, Phys. Rev. B **46**, 15697 (1992); Yu. A. Izyumov, B. M. Letfulov, and E. V. Shipitsyn, J. Phys.: Condens. Matter **6**, 5137 (1994).
31. W. Metzner, Phys. Rev. B **43**, 8549 (1991).
32. M. V. Fedoryuk, *Asymptotics, Integrals, and Series* (Nauka, Moscow, 1987).
33. A. A. Abrikosov, L. P. Gor'kov, and I. E. Dzyaloshinskiĭ, *Methods of Quantum Field Theory in Statistical Mechanics* (Fizmatgiz, Moscow, 1962; Pergamon, Oxford, 1962).
34. B. T. Geilikman and V. Z. Kresin, Fiz. Tverd. Tela (Leningrad) **7**, 3294 (1965) [Sov. Phys. Solid State **7**, 2659 (1965)]; Fiz. Tverd. Tela (Leningrad) **9**, 3111 (1967) [Sov. Phys. Solid State **9**, 2453 (1967)]; Pis'ma Zh. Éksp. Teor. Fiz. **5**, 271 (1967) [JETP Lett. **6**, 220 (1967)].
35. W. L. McMillan, Phys. Rev. **167**, 331 (1968).
36. G. M. Eliashberg, Zh. Éksp. Teor. Fiz. **38**, 966 (1960) [Sov. Phys. JETP **11**, 696 (1960)]; Zh. Éksp. Teor. Fiz. **39**, 1437 (1960) [Sov. Phys. JETP **12**, 1000 (1960)].

Journal of Experimental and Theoretical Physics, Vol. 97, No. 3, 2003, pp. 645–648.
From Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki, Vol. 124, No. 3, 2003, pp. 714–717.

SOLIDS
Electronic Properties

Magnetotransport in a Modulated Two-Dimensional Electron Gas¶

N. A. Zimbovskaya

The City College of the City University of New York, 138th St. and Convent Ave., New York, NY, 10031 USA
Ural Technical Academy of Mining and Geology, ul. Kuibesheva 30, Yekaterinburg, 620000 Russia
e-mail: nzimbov@physlab.sci.ccny.cuny.edu

Received January 29, 2003

Abstract—We propose a semiclassical theory of dc magnetotransport in a two-dimensional electron gas modulated along one direction with weak electrostatic modulations. We show that oscillations of the magnetoresistivity $\rho_\parallel$ corresponding to the current driven along the modulation lines observed at moderately low magnetic fields can be explained as commensurability oscillations. *© 2003 MAIK "Nauka/Interperiodica".*

The theory of dc magnetotransport in modulated a 2*D* electron gas is well developed at present, and most of the effects observed in such systems at low magnetic fields have been explained by both quantum mechanical (in a semiclassical limit) [1–7] and classical [8–12] transport calculations, giving consistent results. One of the few exceptions is the effect of oscillations of the resistivity component $\rho_\parallel$ that corresponds to the current driven parallel to the modulation lines. These oscillations were observed along with the commensurability oscillations of the other resistivity component $\rho_\perp$ corresponding to the current driven across the modulation lines. The $\rho_\parallel$ oscillations have the same period as the $\rho_\perp$ ones and the opposite phase. The oscillations of $\rho_\parallel$ have been explained as an effect that originates in quantum oscillations of the electron density of states in the applied magnetic field [1, 3].

On the other hand, the observed coincidence of periods of the low-field commensurability (Weiss) oscillations of the resistivity component $\rho_\perp$ and the weaker antiphase oscillations of $\rho_\parallel$ provides grounds for the assumption that these oscillations have the same nature and origin for both resistivity components. The purpose of the present paper is to demonstrate that the most important characteristic features of the low-field oscillations of the resistivity $\rho_\parallel$ can be qualitatively reproduced within the semiclassical transport theory. To simplify the following calculations, the anisotropy effects in electron scattering are neglected and the relaxation time approximation is used. It is also assumed that the external magnetic field is moderately weak, such that the electron cyclotron radius R is considerably smaller than the electron mean free path l but larger than the period of modulations λ, and that $R \gg \sqrt{l\lambda}$. This provides the preferred conditions for observation of commensurability oscillations of transport coefficients of the 2*D* electron gas.

We consider electrostatic modulation with a single harmonic of the period $\lambda = 2\pi/g$ along the y direction given by

$$\Delta E(y) = -\frac{dV(y)}{dy}.$$

The screened modulation potential $V(y)$ is parameterized as

$$eV(y) = \epsilon E_F \sin gy,$$

where E_F is the Fermi energy of the 2*D* electron gas. We examine weak modulations, and hence, $|\epsilon gl| \ll 1$. The electron current density in the 2*D* electron gas modulated along the y direction also depends on y and can be written as

$$\mathbf{j}(y) = Ne^2 \int_0^{2\pi} \frac{d\psi}{2\pi} \mathbf{v}(y, \psi)\Phi(y, \psi), \qquad (1)$$

where $N = m/\pi\hbar^2$ is the electron density of states on the Fermi surface, and m and e are the effective mass and charge of the electron. The electron velocity vector $\mathbf{v}(y)$ has a direction of $\mathbf{u}(\psi) = (\cos\psi, \sin\psi)$ and a magnitude of

$$v(y) = v_F\sqrt{1 + \epsilon \sin gy},$$

where v_F is the Fermi velocity in the unmodulated 2*D* electron gas. The distribution function $\Phi(y, \psi)$ satisfies

¶This article was submitted by the author in English.

the linearized Boltzmann transport equation

$$D[\Phi] + C[\Phi] = \mathbf{E} \cdot \mathbf{v}, \tag{2}$$

where $\mathbf{E}$ is the electric field. The collision term $C[\Phi]$ is written in the relaxation tune approximation with the relaxation towards the local equilibrium distribution,

$$C[\Phi] = \frac{1}{\tau}\left(\Phi(y, \psi) - \frac{1}{2\pi}\int_0^{2\pi} \Phi(y, \psi), d\psi\right), \tag{3}$$

and the drift term is given by

$$D[\Phi] = v(y)\sin\psi\frac{\partial\Phi}{\partial y} + (v'(y)\cos\psi + \Omega)\frac{\partial\Phi}{\partial\psi}, \tag{4}$$

where Ω is the electron cyclotron frequency. Linearized transport equation (2) with the collision and drift terms of form (3), (4) was used in [8] and agrees with the transport equations in [9–11].

Following the standard approach [10], we write $\Phi(y, \psi)$ as

$$\Phi(y, \psi) = \Phi_0(\psi)\frac{v(y)}{v_F} + \rho_0\tau\chi(y, \psi), \tag{5}$$

where ρ_0 is the Drude resistivity and τ is the relaxation time. The homogeneous distribution function

$$\Phi_0(\psi) = \rho_0\tau\mathbf{v}_0 \cdot \mathbf{j}_0$$

describes the linear response of the 2*D* electron gas to the field $\mathbf{E}$ in the absence of modulations, and the function $\chi(y, \psi)$ satisfies the transport equation

$$D[\chi] + C[\chi] = -v'(y)v(y)j_{0y}. \tag{6}$$

Here, as before, $\mathbf{j}_0$ is the current density for the unmodulated 2*D* electron gas.

To proceed, we expand $\chi(y, \psi)$ in a Fourier series in the spatial variable y, which leads to a system of differential equations for the Fourier components. Solving these equations and keeping the terms on the order of or larger than $(\epsilon gR)^2$, we obtain the approximation for the distribution function $\chi(y, \psi)$

$$\begin{aligned}\chi(y, \psi) = &-\frac{v_F}{2}\epsilon g l j_0^y Q\left(\cos\psi - \frac{\sin\psi}{\Omega\tau}\right)\\ &\times\left\{\sin(gR\cos\psi + gy)\right.\\ &-\frac{1}{2}\epsilon gR\left[\cos(gR\cos\psi)\cos^2 gy\right.\\ &\left.\left.-\frac{1}{2}\sin(gR\cos\psi)\sin 2gy\right]\right\},\end{aligned} \tag{7}$$

where

$$Q = \frac{J_0(gR)}{1 - J_0^2(gR)} \tag{8}$$

and $J_0(gR)$ is the Bessel function. Using the distribution function obtained, we can easily calculate the electron current density given by (1).

Keeping only the leading terms in the expansion of $\chi(y, \psi)$ in powers of the small parameter $(\Omega\tau)^{-1}$, we find that only the j_x component receives a correction due to the modulations along the *y* direction, whereas the component j_y remains equal to j_y^0 and does not depend on spatial coordinates. This agrees with the continuity equation

$$\nabla \cdot \mathbf{j} = 0,$$

which is necessary in order to obtain correct results for electron transport coefficients in a modulated 2*D* electron gas [10].

To proceed, we define the effective conductivity tensor σ_{eff} as

$$\mathbf{j} \equiv \langle \mathbf{j}(y)\rangle \equiv \frac{g}{2\pi}\int_0^{2\pi/g} \mathbf{j}(y), dy \equiv \sigma_{\text{eff}}\mathbf{E}. \tag{9}$$

To justify the definition adopted in (9), we note that the expressions for transport coefficients obtained with either quantum mechanical or classical calculations must be consistent at low magnetic fields. Quantum mechanical calculations of magnetoconductivity [3, 6, 7] yield an expression that passes to the classical conductivity tensor averaged over the period of modulations. The latter is therefore an accurate semiclassical analog of the conductivity calculated within the proper quantum mechanical approach, and our definition of σ_{eff} agrees with this.[1] The same definition was previously used in [11].

As a result, we find that only σ_{eff}^{xx} is affected by the modulations,

$$\sigma_{\text{eff}}^{xx} = \frac{\sigma_0}{1 + (\Omega\tau)^2} + \frac{1}{4}(\epsilon gR)^2\sigma_0\frac{J_0^2(gR)}{1 - J_0^2(gR)}, \tag{10}$$

where $\sigma_0 = 1/\rho_0$ is the Drude conductivity of the electron system. The second term in (10) represents the

[1] For electrostatic modulations, definition (9) actually gives the same results for magnetoresistivity components as the alternative definition $\rho_{\text{eff}}\langle\mathbf{j}\rangle \equiv \mathbf{E}$ used in [10]. However, there is a significant difference in results based on these definitions when magnetic modulations are considered.

electron diffusion along the x direction caused by the guiding center drift [8].

The effective magnetoresistivity tensor is here defined as the inverse of the effective conductivity,

$$\rho_{\text{eff}} = \sigma_{\text{eff}}^{-1}.$$

For the current driven across the modulation lines, the corresponding resistivity is

$$\rho_\perp = \rho_{yy} = \rho_0\left\{1 + \frac{1}{4}(\epsilon g l)^2 \frac{J_0^2(gR)}{1 - J_0^2(gR)}\right\}. \quad (11)$$

Assuming that the current flows along the modulation lines, we obtain

$$\rho_\parallel = \rho_{xx} = \rho_0\left\{1 - \frac{1}{4}(\epsilon g R)^2 \frac{J_0^2(gR)}{1 - J_0^2(gR)}\right\}. \quad (12)$$

For moderately weak magnetic fields ($gR \gg 1$), the results in (12) and (13) describe oscillations of both magnetoresistivity components periodic in the magnitude of the inverse magnetic field. The oscillations of $\rho_\perp$ and $\rho_\parallel$ have the same period in $1/B$ and opposite phases, which corroborates the experimental results in [1]. The amplitude of the oscillations of $\rho_\parallel$ is considerably smaller than that of $\rho_\perp$, and this also agrees with the experiments of [1] and with the results of numerical quantum mechanical calculations in the limit of a weak magnetic field [3]. The result for the resistivity $\rho_\perp$ also agrees with the corresponding results in [8–11] obtained within the classical magnetotransport theory.

However, expression (12) for $\rho_\parallel$ differs from the well-known result in the current semiclassical theory. To analyze this discrepancy, we now calculate the current density taking into account the next terms in the expansion of distribution function (7) in powers of $(\Omega\tau)^{-1}$. Keeping terms on the order of not less than $(\epsilon gR/\Omega\tau)^2$, we find that the grating-induced correction to the Drude conductivity tensor $\hat{\sigma}_0$ is given by

$$\delta\hat{\sigma}(y) = \frac{\sigma_0}{1 + (\Omega\tau)^2}\alpha(y)\begin{pmatrix} (\Omega\tau)^2 & \Omega\tau \\ -\Omega\tau & -1 \end{pmatrix}. \quad (13)$$

Here, correction $\alpha(y)$ is on the order $(\epsilon gR)^2$. With some formal transformations of transport equation (6), we can represent $\langle\alpha(y)\rangle$ in the form

$$\langle\alpha(y)\rangle = \frac{1}{2\pi}\int_0^{2\pi}\langle v(y)\sin\psi G(y,\psi)\rangle, d\psi, \quad (14)$$

where $G(y, \psi)$ satisfies the equation

$$D[G] + C[G] = -\frac{2v'(y)v(y)}{v_F^2}. \quad (15)$$

This gives expressions for σ_{eff} and ρ_{eff} that are totally consistent with the existing semiclassical theory [8–12]. However, expression (13), which is the starting point in the derivation of these results, is obviously incorrect because it violates the continuity equation for the current density. This gives grounds to seriously doubt the results of earlier works [8–12], especially those concerning $\rho_\parallel$.

A detailed analysis shows that simplified transport equation (2) can be successfully used in calculating the leading terms in the expansions of transport coefficients in powers of $(\Omega\tau)^{-1}$, and expressions (12)–(14) are therefore valid. To obtain the next terms in these expansions, we must modify transport equation (2) in both the drift and collision terms. For this, we must consistently and systematically consider effects of the internal electrochemical field arising due to grating-induced inhomogeneity of the electron density. This is important because redistribution of the electron density in the presence of modulations provides the local equilibrium of the system.[2] Taking into consideration magnetic modulations, we arrive at similar results [13].

Finally, the novel result in this paper is a qualitative explanation of the low-field oscillations of the magnetoresistivity component $\rho_\parallel$ in the $2D$ electron gas modulated along one direction within a semiclassical approach. We have shown that these oscillations of $\rho_\parallel$ at low magnetic fields are commensurability oscillations. At low temperatures, with the quantum oscillations of the electron density of states at the Fermi surface resolved, Shubnikov–de Haas oscillations can be superimposed on the geometric oscillations of the magnetoresistivity. However, this does not change the classical nature of the effect itself.

The author thanks G.M. Zimbovsky for help with the manuscript.

REFERENCES

1. R. R. Gerhardts, D. Weiss, and K. von Klitzing, Phys. Rev. Lett. **62**, 1173 (1989); R. W. Winkler, J. P. Kotthaus, and K. Ploog, Phys. Rev. Lett. **62**, 1177 (1989).

[2] The method of calculations adopted in the current theory is mostly based on averaging the transport equation multiplied by a velocity component with respect to both ψ and y (see, e.g., [10, 11]). This procedure ignores the contribution from the second term of collision integral (3), which provides relaxation of the system towards the local equilibrium. As a result, one arrives at the expressions for transport coefficients corresponding to the relaxation of the system towards the total equilibrium, which contradicts the continuity equation.

2. P. Vasilopoulos and F. M. Peeters, Phys. Rev. Lett. **63**, 2120 (1989).
3. C. Zhang and R. R. Gerhardts, Phys. Rev. B **41**, 12850 (1990).
4. F. M. Peeters and P. Vasilopoulos, Phys. Rev. B **46**, 4667 (1992); Phys. Rev. B **47**, 1466 (1993).
5. Q. W. Shi and K. Y. Szeto, Phys. Rev. B **53**, 12990 (1996).
6. A. Manolescu, R. R. Gerhardts, M. Suhrke, and U. Rossler, Phys. Rev. B **63**, 115322 (2001).
7. J. Grob and R. R. Gerhardts, Phys. Rev. B **66**, 155321 (2002).
8. C. W. J. Beenakker, Phys. Rev. Lett. **62**, 2020 (1989).
9. R. R. Gerhardts, Phys. Rev. B **53**, 11064 (1996).
10. R. Menne and R. R. Gerhardts, Phys. Rev. B **57**, 1707 (1998).
11. A. D. Mirlin and P. Wolfle, Phys. Rev. B **58**, 12986 (1998).
12. A. D. Mirlin, P. Wolfle, Y. Levinson, and O. Entin-Wohlman, Phys. Rev. Lett. **81**, 1070 (1998); F. von Oppen, A. Stern, and B. I. Halperin, Phys. Rev. Lett. **80**, 4494 (1998); S. D. Zwerscke and R. R. Gerhardts, Phys. Rev. Lett. **83**, 2616 (1999).
13. N. A. Zimbovskaya and J. L. Birman, Solid State Commun. **116**, 21 (2000); N. A. Zimbovskaya, J. L. Birman, and G. Gumbs, cond-mat/0204290.