JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Effects related to spacetime foam in particle physics

A. A. Kirillov*)

Institute for Applied Mathematics and Cybernetics, 603005 Nizhniĭ Novgorod, Russia

(Submitted 24 November 1998)
Zh. Éksp. Teor. Fiz. **115**, 1921–1934 (June 1999)

It is found that the existence of spacetime foam leads to a situation in which the number of fundamental quantum bosonic fields is a variable quantity. The general aspects of an exact theory that allows for a variable number of fields are discussed, and the simplest observable effects generated by the foam are estimated. It is shown that in the absence of processes related to variations in the topology of space, the concept of an effective field can be reintroduced and standard field theory can be restored. However, in the complete theory the ground state is characterized by a nonvanishing particle number density. From the effective-field standpoint, such particles are ''dark.'' It is assumed that they comprise dark matter of the universe. The properties of this dark matter are discussed, and so is the possibility of measuring the quantum fluctuation in the field potentials.
[S1063-7761(99)00106-7]

1. INTRODUCTION

In gravitation theory it is assumed that spacetime is a flat manifold at scales much greater than the Planck length, while at the Planck scale all geometric properties disappear and spacetime itself acquires a foamlike structure.[1] There are two basic indications of such behavior of spacetime. The first is related to the fact that at the Planck scale the vacuum fluctuations of the metric and curvature are of the same order as the corresponding average quantities. Not only does this follow from simple estimates—rigorous calculations also support this idea. In particular, the fact that such fluctuations exist leads to the absence of a classical background space in the Planck stage of evolution of the early universe.[2] The second indication is the fact that at small scales the topology of space also experiences quantum fluctuations.[1] The study of possible observable effects related to changes in the topology of space is attracting ever more attention. In particular, to describe such effects, Hawking[3] used wormholes and virtual black holes. Another work worth noting is that of Garay,[4] who proposed a phenomenological method to account for spacetime foam.

The absence of a background space at small scales is a serious problem in quantum field theory. The possibility of resolving this problem is usually related to the development of nonperturbative methods,[5] in which the concept of background fields is not used. However, these theories also rely on the presence of a coordinate basis space, whose topology is fixed by the statement of the problem and therefore is not a dynamic characteristic.

This paper elaborates on a possible way to set up a quantum field theory in the case in which the topology and structure of physical space may vary. The main idea of this method was set forth in Ref. 6 in order to describe the quantum birth of the early universe.

The observation forms the basis of the proposed method. On the one hand, as noted earlier, variations in the topology of space can occur at scales where the very concept of a flat manifold breaks down, at least due to the presence of vacuum fluctuations. On the other hand, it is believed that there is no other way to describe the given region but to extrapolate the spatial relationships existing at larger scales to it. In other words, all possible topologies of physical space should be described in terms of a consistent coordinate basis space. We call this space simply a basis. Since measurement instruments, which are classical objects, play a fundamental role in quantum theory,[7] it is expected that the properties of the basis are determined entirely by the measuring device.

If we specify the quantum state corresponding to a fixed topology of physical space and if the topology differs substantially from that of the basis, its image in terms of the basis coordinate cannot be one-to-one. In the same way, when functions defined in physical space and corresponding to different physical observables are mapped to the basis space, they cease to be single-valued and become multivalued functions of the physical coordinates. Furthermore, the number of images of an arbitrary physical observable is an additional variable quantity, which, generally speaking, depends on the position in the basis space.

Thus, we arrive at a situation in which the number of fields corresponding to a physical observable is a variable quantity. In quantum theory this variable is an operator whose eigenvalues characterize the topological structure of space. The possible dependence of this quantity on spatial coordinates means that the given quantity is a characteristic or measure of the number density of the degrees of freedom of the field.

A natural way to describe systems with a variable number of degrees of freedom is to use second quantization. Before we begin to describe the method as applied to the problem in question, we make the following remark. In the standard second-quantization method, the number of degrees of freedom characterizes the number of particles or elementary excitations (quanta) in the system. Here it is assumed that the particles obey the identity principle or, as it is said, the indistinguishability principle. It can be expected that in measurements at small scales the different images of the same physical observable also obey the identity principle. Indeed, the possibility of distinguishing between the different images of observables would mean that physical space in itself has certain topology and structure, which by assumption is impossible (at least in view of the presence of quantum fluctuation in the topology).

Two types of statistics, Bose and Fermi, exist for particles, depending on the symmetry of the wave function under particle permutations. Accordingly, we must also select the type of statistics when performing second quantization of the degrees of freedom of the fields. Since second quantization reflects the properties and topology of physical space, this selection must be unique for all types of fields and physical quantities. Here it turns out that the only acceptable choice is Fermi–Dirac statistics, since otherwise in dealing with fermions we immediately confront a violation of the Pauli principle.

At a fundamental level, the composition of matter is determined by a set of fields and their sources. The sources are point particles, which in quantum theory behave like fermions. The need to perform second quantization of the sources arises already in relativistic theory and hence no changes in the description of fermions emerge. A new interpretation is added, however. For instance, pair production corresponds to a change in the structure of physical space (it can be said that processes related to changes in the properties of space proceed much more easily at isolated points than they do in entire regions).

When fields are quantized, the idea of particles, the quanta of a field, also emerges. Such particles, however, obey the Bose–Einstein statistics. Here, generally speaking, particle production is not associated with variations in the topology of space. There is a certain similarity between this aspect and the situation in solid state physics, where excitation of vibrations in a crystal lattice (phonon production) is not associated with variations in the true number of degrees of freedom.

Thus, the variation mentioned above primarily involves bosonic fields.

2. GENERAL SCHEME OF SECOND QUANTIZATION OF FIELDS

We consider a set M, which in the future acts like a basis manifold, and specify an arbitrary field φ on it. We also assume that there is a device that can do complete measurements of the quantum states of the field. A complete measurement can always be expanded in a set of elementary measurements. For instance, to make a complete measurement of a field state we must measure the field amplitude at every point $x \in M$, or equivalently measure the number of particles (or amplitude) in each Fourier mode. Thus, the device can be viewed as a set of elementary detectors.

Let A be the set of possible readings of an elementary detector.[1)] The structure of A can be described in the following way. In A we select an arbitrary system of coordinates ξ. Generally, there is a natural projection operator P ($P^2=P$) that partitions the coordinates into two groups: $\xi=((I-P)\xi,P\xi)=(\eta,\zeta)$, where I denotes the identity operator. The first group, ζ, refers to the manifold M and describes the position in space M^* at which the elementary measurement takes place (here and in what follows M^* denotes either space M or the mode space). The second group, η, refers to the field φ and describes the position in space V. The coordinate $\eta \in V$ denotes either the field amplitude or the number of particles corresponding to the field. Thus, the set A acquires the features of a fiber space with basis $P(A) \sim M^*$ and fiber $P^{-1}(\zeta)=V$. The result of a complete measurement of field φ is a fibration section, which is the map $\varphi: M^* \rightarrow A$. What is important is that in the usual picture an arbitrary section intersects each fiber only once, i.e., the projection of the section coincides with the space $M^*(P(\varphi)\equiv M^*)$, which implies that such sections can be represented by functions $\eta(\zeta)$ on M^* with values in V.

As noted above, the topology and geometric structure of the set A (and thus of M^*) reflects the macroscopic properties of the measurement process. On the other hand, the real physical space M_{ph} is assumed to have arbitrary topology and structure.[2)] Furthermore, in a general quantum state, the properties of space M^*_{ph} are, generally speaking, not fixed. Thus, a physical field must be defined as an extended section of the form $\tilde{\varphi}: M^*_{\text{ph}} \rightarrow A$. Here an arbitrary section can intersect each fiber an arbitrary number of times. Furthermore, if the topology of space M^*_{ph} changes, so does the number of intersections. Thus, the number of images of field $\tilde{\varphi}$ in space M^* is variable. An image of space M^*_{ph} is a subset in $M^*(P(\tilde{\varphi})=M'_{\text{ph}} \subset M^*)$ that can be represented as a union of distinct pieces, $M'_{\text{ph}}=\cup \sigma_j$, so that on each piece σ_j the field is described by a given number of functions $\eta_i(\zeta)$, $\zeta \in \sigma_j$ ($i=1,2,\dots,m$, where m is an integer characterizing the number of images of space M^*_{ph} in σ_j). Note that in general the dimensionality of the pieces σ_j can differ from the dimensionality of M^*.

Thus, if the topology of physical space in an additional degree of freedom, the result of a complete measurement of the state of the field will be represented by a definite set of functions $\{n_J(\zeta)\}(J=(i,\sigma)$ and $\zeta \in \sigma)$. Formally, such states can be classified in the following way.

We introduce a set of operators $C^\dagger(\xi)$ and $C(\xi)$, the creation and annihilation operators for an individual elements of the set A. For the sake of simplicity we assume that the measure of each individual point $\xi \in A$ is finite (as in the case in which the coordinates ξ take discrete values). We require that these operators satisfy the anticommutation relations

$$\{C(\xi)C^\dagger(\xi')\}=C(\xi)C^\dagger(\xi')+C^\dagger(\xi')C(\xi)=\delta_{\xi\xi'}. \quad (1)$$

We define the vacuum state $|0\rangle$ by the relationship $C(\xi)|0\rangle=0$ and build a Fock space F in which the basis consists of the vectors ($n=1,2,\dots$)

$$|\xi_1,\xi_2,\dots,\xi_n\rangle=\prod_{i=1}^{n} C^\dagger(\xi_i)|0\rangle. \quad (2)$$

The vacuum state corresponds to complete absence of a field and hence of the observables associated with the field. The state $|\xi\rangle$ describes the field φ with only one degree of freedom. This can be either a field concentrated at a single point or a field containing only one mode, and the quantity $\xi\in A$ describes the intensity (the number of quanta) and the position of the field in M^*. States described by single-valued functions are constructed in the following way:

$$|\eta(\zeta)\rangle=\prod_{\zeta\in M^*} C^\dagger(\eta(\zeta),\zeta)|0\rangle, \tag{3}$$

where the direct product is taken over the entire space M^*, and where we have partitioned the coordinates ξ into two groups: $\xi=(\eta,\zeta)$. Generally, such states do not belong to a Fock space. Furthermore, when the coordinates $\zeta\in M^*$ run through continuous values, this expression requires an extension of its definition and hence can be interpreted only formally. However, when the variations of the physical quantities in real processes involve only a finite part of the set M^*, we can stay within a Fock space.

We now examine an arbitrary domain $\sigma\in M^*$ and define a set of operators

$$D^\dagger(\eta(\zeta),\sigma)=\prod_{\zeta\in\sigma} C^\dagger(\eta(\zeta),\zeta), \tag{4}$$

where the domain of the function $\eta(\zeta)$ is limited to the set σ. Then the states with an arbitrary number of fields can be written

$$|\eta_1,\eta_2,\ldots,\eta_n\rangle=\prod_{i=1}^{n} D^\dagger(\eta_i(\zeta),\sigma_i)|0\rangle. \tag{5}$$

The interpretation of these states is obvious. Suppose that all functions $\eta_i(\zeta)$ are specified on a single set σ. Then in the given domain a complete measurement will show the presence of a set consisting of n different fields $\eta_1(\zeta)$, $\eta_2(\zeta)$, $\ldots$, $\eta_n(\zeta)$. It is convenient to introduce the number density operator of the fields:

$$N(\zeta)=\sum_{\eta\in V} C^\dagger(\eta,\zeta)C(\eta,\zeta). \tag{6}$$

Then for $\zeta\in\sigma$ the states (5) represent the eigenstates of the operator $N(\zeta)$ with eigenvalues

$$N(\zeta)|\eta_1,\eta_2,\ldots,\eta_n\rangle=n|\eta_1,\eta_2,\ldots,\eta_n\rangle. \tag{7}$$

Clearly, the states with a fixed number of fields correspond to a fixed topology of the space M^*_{ph}. Then under certain conditions (the requirement that the functions $\eta_i(\zeta)$ be smooth at cuts), instead of the set of functions $\eta_i(\zeta)$ we can introduce a single-valued function $\eta(\zeta)$ and thus restore the structure of the set M^*_{ph}. Conversely, each space M^*_{ph} can be projected on the basis of M^* by performing the necessary paste-up, so that the state vector of the field has the form (5).

The space H formed by the vectors (5) and their superposition lays the basis for building the Hilbert space of the theory. An arbitrary operator $\hat{O}(\xi)$ related to the field (and symmetrized in the number of the fields) can be expressed in the standard way in the terms of the set of basis operators C and $C^\dagger$:

$$\hat{O}=\sum D_I^\dagger O_{IJ} D_J$$

(where $I,J=(\eta_i(\zeta),\sigma)$, and σ is an arbitrary domain in M^*), thus defining the action of this operator in H. The specific way in which this Hilbert space is built is determined by the physical problem at hand.

3. SCALAR FIELD IN THE SECOND-QUANTIZATION REPRESENTATION

In Sec. 2 we discussed the general scheme of second quantization, irrespective of the dynamics of the field. We now turn to the example of a real scalar field φ (the generalization to the case of arbitrary fields is obvious). For the basis space we take ordinary flat Minkowski space.

One idea that is central to particle physics is the representation in which quantum states of a field are classified in terms of physical particles. Since quantum states of a field can in general contain an arbitrary number of the identical modes, the definition of particles and their relation to field operators require certain modifications. We find it more convenient to operate with discrete indices. To this end we require that the field in question be located in a cube with edge length L, and we introduce periodic boundary conditions. As necessary, we can replace sums with integrals (as $L\to\infty$) via the usual prescription: $\Sigma\to\int(L/2\pi)^3\,d^3k$.

We now examine the expansion of the field operator φ in plane waves,

$$\varphi(x)=\sum_k (2\omega_k L^3)^{-1/2}(a_k\exp\{ikx\}+a_k^\dagger\exp\{-ikx\}), \tag{8}$$

where $\omega_k=\sqrt{k^2+m^2}$, and $k=2\pi n/L$, with $n=(n_x,n_y,n_z)$. The general expression for the Hamiltonian is

$$H=H_0+V, \tag{9}$$

where H_0 describes free particles,

$$H_0=\sum_k \omega_k a_k^\dagger a_k+e_k, \tag{10}$$

and the potential term V is responsible for the interaction, and can be represented in the normal form:

$$V=\sum_{n,\{m\},\{m'\}} V^n_{\{m\},\{m'\}}, \tag{11}$$

$$V^n_{\{m\},\{m'\}}={\sum_{k_1,\ldots,k_n}}' V^n_{\{m\},\{m'\}}(k_1,\ldots,k_n) \times\prod_{i=1}^{n}(a_{k_i}^\dagger)^{m_i}(a_{k_i})^{m'_i}. \tag{12}$$

Here we assume that the sum with respect to the wave vectors k_i contains no terms with equal indices, i.e., $k_i\neq k_j$ for any pair of indices i and j (the sum is taken over distinct modes), and allow for the fact that for different wave numbers the operators a_{k_i} and $a_{k_j}^\dagger$ commute.

The quantity e_k in (10) is the energy of the ground state of the kth mode. In a flat space without particles, the energy must be zero, so we assume that $e_k=0$ throughout the present paper. However, as we show in the sections that follow, the nontrivial nature of the topology of the spacegenerally leads to a value of e_k that is finite. Note that the dependence of the zero energy on the topology of space is known as the Casimir effect[8] and is assumed to be an experimentally established fact.[9]

When the number of modes is variable, the set of field operators $\{a_k, a_k^\dagger\}$ is replaced by the somewhat expanded set $\{a_k(j), a_k^\dagger(j)\}$, where $j\in[1,\ldots,N_k]$, and N_k is the number of modes for a given wave number k. For a free field the energy is an additive quantity, which can be written

$$H_0=\sum_k \sum_{j=1}^{N_k} \omega_k a_k^\dagger(j)a_k(j). \tag{13}$$

Since the modes are indistinguishable, the interaction operator has the obvious generalization

$$V^n_{\{m\},\{m'\}}={\sum_{k_1,\ldots,k_n}}' \sum_{j_1,\ldots,j_n} V^n_{\{m\},\{m'\}}(k_1,\ldots,k_n) \times\prod_{i=1}^n (a^\dagger_{k_i}(j_i))^{m_i}(a_{k_i}(j_i))^{m'_i}, \tag{14}$$

where the indices j_i run through the corresponding intervals $j_i\in[1,\ldots,N(k_i)]$. It is convenient to introduce the notation

$$A_{m,n}(k)=\sum_{j=1}^{N(k)} (a_k^\dagger(j))^m(a_k(j))^n. \tag{15}$$

Then the expression for the field Hamiltonian takes the form

$$H=\sum_k \omega_k A_{1,1}(k) + \sum_{n,\{m\},\{m'\}} {\sum_{k_1,\ldots,k_n}}' V_{\{m\},\{m\}'}(k_1,\ldots,k_n) \times\prod_{i=1}^n A_{m_i,m'_i}(k_i). \tag{16}$$

We can now express the main quantities in terms of the fundamental operators $C^\dagger(\xi)$ and $C(\xi)$. For the operators a and $a^\dagger$ it is convenient to use the Fock–Bargmann representation, in which operators act in the space of entire analytic functions with a scalar product of the type

$$(f,g)=\int f^*(a)g(a^*)\exp\{-a^*a\}\frac{da^*\,da}{2\pi i}; \tag{17}$$

the action of these operators is defined as

$$a^\dagger f(a^*)a^*f(a^*), \quad af(a^*)=\frac{d}{da^*}f(a^*). \tag{18}$$

Then for the normal field coordinates we can take the complex-valued quantities a^*; thus, the set A consists of the pairs $\xi=(a^*,k)$. For the fundamental operators $C^\dagger(\xi)$ and $C(\xi)$ it is convenient to use the representation

$$C(a^*,k)=\sum_{n=0}^\infty C(n,k)\frac{(a^*)^n}{\sqrt{n!}},$$

$$C^\dagger(a,k)=\sum_{n=0}^\infty C^\dagger(n,k)\frac{a^n}{\sqrt{n!}}. \tag{19}$$

Then the anticommutation relations (1) become

$$\{C(n,k)C^\dagger(m,k')\}=\delta_{n,m}\delta_{k,k'}\,. \tag{20}$$

The physical meaning of the operators $C(n,k)$ and $C^\dagger(n,k)$ is that they create and annihilate modes with a given number of particles.

Now, to express the Hamiltonian (16) in terms of $C(n,k)$ and $C^\dagger(n,k)$ it suffices to derive the corresponding expressions for the operators (15). In the second-quantization representation, the expressions for the given operators are defined to be

$$\hat{A}_{m,n}(k)=\int \exp\{-a^*a\}\frac{da^*\,da}{2\pi i} \times C^\dagger(a,k)(a^*)^m\left(\frac{d}{da^*}\right)^n C(a^*,k) \tag{21}$$

or, with allowance for (19),

$$\hat{A}_{m_1,m_2}(k)=\sum_{n=0}^\infty \frac{\sqrt{(n+m_1)!\,(n+m_2)!}}{n!} \times C^\dagger(n+m_1,k)C(n+m_2,k). \tag{22}$$

An expression for the Hamiltonian in terms of the operators $C^\dagger(\xi)$ and $C(\xi)$ can be obtained by simply substituting (22) into (16). For a free field, the eigenvalues of the Hamiltonian take the form

$$\hat{H}_0=\sum_k \omega_k\hat{A}_{1,1}(k)=\sum_{n,k} n\omega_k N_{n,k}\,, \tag{23}$$

where $N_{n,k}$ is the number of modes for fixed values of the wave number k and the number of particles n $(N_{n,k}=C^\dagger(n,k)C(n,k))$.

Thus, the field state vector Φ is a function of the occupation numbers $\Phi(N_{k,n},t)$, and its evolution is described by the Schrödinger equation

$$i\partial_t\Phi=H\Phi. \tag{24}$$

Consider the operator

$$N_k=\sum_{n=0}^\infty C^\dagger(n,k)C(n,k). \tag{25}$$

Physically, this operator characterizes the total number of modes for a fixed wave number k. One can easily verify that for the Hamiltonian (16), N_k is a constant of the motion,

$$[N_k,H]=0, \tag{26}$$

and in this way Hamiltonians like (16) preserve the topological structure of the field. In the course of evolution, the number of modes for each k does not change.

We now turn to the problem of representing the particle creation and annihilation operators in this formalism. Since the individuality of the modes is limited, operators of type (22) act like the set of operators $\{a_k(j), a_k^\dagger(j)\}$. Among the operators (22) are some that change the number of particles by one:

$$b_m(k)=\hat{A}_{m,m+1}(k), \quad b_m^\dagger(k)=\hat{A}_{m+1,m}(k), \tag{27}$$

$$[\hat{n}, b_m^\dagger(k)]=\pm b_m^\dagger(k), \quad [H_0, b_m^\dagger(k)]=\pm\omega_k b_m^\dagger(k), \tag{28}$$

where

$$\hat{n}=\sum_k \hat{n}_k=\sum_{n,k} nN_{n,k}. \tag{29}$$

Then the ground state Φ_0 of the field can be defined as a vector satisfying the relationships $(m=0,1,\dots)$

$$b_m(k)\Phi_0=0 \tag{30}$$

and corresponding to the minimum energy for a fixed mode distribution N_k. Note that in contrast to standard theory, the ground state is generally characterized by a nonvanishing particle number density $\hat{n}\Phi_0=n_0\Phi_0$. Using the vector Φ_0, we can build a Fock space F whose basis consists of vectors obtained by cyclic application of the operators $b_m^\dagger(k)$ to Φ_0.

4. EFFECTIVE FIELD

In the absence of processes related to changes in the topology of space and for a mode distribution of the form $N_k=1$ (there is only one mode for each wave number k), the standard field theory is restored. Furthermore, there is a fairly general case in which the concept of an effective field can be introduced to restore the standard picture.

Indeed, consider the case in which the interaction operator in (16) is expressed solely in terms of the set of operators $b_0(k)$ and $b_0^\dagger(k)$. Then instead of the complete Fock space F we can limit ourselves to its subspace $F'\subset F$ formed by the cyclic application of the operators $b_0^\dagger(k)$ to the field ground state Φ_0. If the initial state vector Φ belongs to F', then as the system evolves, $\Phi(t)\in F'$ for all t (at least as long as the number of particles created remains finite).

We define the operators

$$a_k=N_k^{-1/2}b_0(k), \quad a_k^\dagger=N_k^{-1/2}b_0^\dagger(k), \tag{31}$$

where N_k is the operator defined in (25), which, when restricted to the Fock space F', is an ordinary number function. For (20) and (22) we find that the commutation relations for a_k and $a_k^\dagger$ have the standard form

$$[a_k, a_{k'}^\dagger]=\delta_{kk'}. \tag{32}$$

Thus, if the basic observable objects are particles, it is possible to revert to the usual picture in which the particles are quanta of an effective field $\tilde{\varphi}$ of type (8). Note that if the field potentials $\varphi(x)$ are measurable quantities, then the true expression for the field operators has the same form (8), where instead of the operators a_k and $a_k^\dagger$ we must put $b_0(k)$ and $b_0^\dagger(k)$. The expression for the effective-field energy operator has the form (9), but the ground-state energy in the kth mode, e_k, must be assumed not to vanish. The value of this energy can be found in the complete theory.

Since the operators a_k and $a_k^\dagger$ reflect only some of the information about the state of the system, the auxiliary nature of the effective field becomes manifest. Indeed, the only observables related to a given field are particles, and not just any particles, but only those that outnumber the particles in the ground state. For the particle number operator in F' we have

$$a_k^\dagger a_k=\delta\hat{n}_k=\hat{n}_k-\overline{n_k}, \tag{33}$$

where $\hat{n}_k$ is the operator defined in (29), and $\overline{n_k}$ can be found by solving $\hat{n}_k\Phi_0=\overline{n_k}\Phi_0$. Thus, the properties of the ground state Φ_0 remain beyond the scope of the effective field.

5. PROPERTIES OF THE FIELD GROUND STATE

Equations (20) and (22) imply that a true vacuum state has the property that all field modes (and hence all observables related to the field) are absent. A true vacuum state is one in which there are no particles and no zero-point vibrations related to particles. This situation is similar to the situation in solid state physics, where in the absence of a crystal there can be no phonons and no zero-point lattice vibrations. Since the properties of physical space are determined by the properties of material fields, we conclude that in a true vacuum state there can be no physical space. Obviously, in reality such a state cannot be achieved.

At first glance the most common situation in particle theory is the one in which physical space is ordinary flat Minkowski space, and nontrivial topology is manifest at the Planck scale (this is the conventional view; see Refs. 1 and 4). But since operating at the Planck scale requires using energies unattainable with present-day accelerators, and also requires serious consideration of quantum gravity effects, it would appear to be impossible to make any sort of directly measurable predictions with this theory.

In reality, the situation is somewhat different. First, the stability of the Minkowski space means that probably even at the Planck scale the topology of the space can be assumed to be simple (i.e., $N_k=1$ and $k\geqslant k_{\mathrm{Pl}}$), at least as long as we do not consider processes in which real particles with Planck energies are produced (naturally, virtual processes cannot lead to real changes in the topology of space).

Second, recall that the universe has already passed the quantum stage, in which real processes involving changes in spatial topology might occur. After the quantum stage, processes with topology variations are suppressed, and we can say that the topological structure of space has been ''tempered,'' so that the structure of the space is preserved as the universe expands. Thus, we expect that at the present time the nontrivial topology of space is most likely manifested on a cosmological scale.

In the foregoing theory, the structure of space is determined by the number density of the field modes. These modes are in turn governed by Fermi statistics, i.e., they act like a Fermi gas. To simplify matters, we examine free fields, since consistent allowance for the interaction of field warrants a separate investigation. We assume that the field-mode distribution was thermal in the Planck period of the evolution of the universe. As the universe expands, the temperature drops and the gas becomes degenerate, with the field winding up in the ground state. Thus, the field ground state Φ_0 can be characterized by occupation numbers of the type

$$N_{k,n}=\theta(\mu_k-n\omega_k), \tag{34}$$

where $\theta(x)$ is the Heaviside step function and μ_k is the chemical potential. Note that when the expansion is adiabatic, we must put $\mu_k=\mu$. When the evolution of the universe includes an inflationary period,[10,11] the adiabaticity condition can be violated, which generally leads to additional dependence of the chemical potential on the wave number. For the mode spectral density we have

$$N_k=\sum_{n=0}^{\infty}\theta(\mu_k-n\omega_k)=1+\left[\frac{\mu_k}{\omega_k}\right], \tag{35}$$

where $[x]$ denotes the integer part of the number x. Equation (35) shows, in particular, that at $\omega_k>\mu_k$ we have $N_k=1$, i.e., the field structure corresponds to a flat Minkowski space, with the result that $\omega_k<\mu_k$ is the range of wave vectors in which nontrivial field properties are expected to show up.

It can easily be verified that from the effective-field standpoint, the ground state Φ_0 is a vacuum state, i.e., $a_k\Phi_0=0$. On the other hand, the given state can be characterized by a nonvanishing particle number density. Indeed, for any wave number we have

$$\overline{n}_k=\sum_{n=0}^{\infty}n\,\theta(\mu_k-n\omega_k)=\frac{1}{2}\left(1+\left[\frac{\mu_k}{\omega_k}\right]\right)\left[\frac{\mu_k}{\omega_k}\right], \tag{36}$$

with the result that the spectral density of the ground-state energy is

$$e_k=\omega_k\overline{n}_k=\frac{\omega_k}{2}\left(1+\left[\frac{\mu_k}{\omega_k}\right]\right)\left[\frac{\mu_k}{\omega_k}\right]. \tag{37}$$

Since the given particles correspond to the ground state of the field, in ordinary processes (which do not change the topology of space) the particles in question are not manifested explicitly (but they enter into the renormalization of the parameters of the observed particles indirectly; here, in contrast to vacuum fluctuations, the contribution of the particles is naturally finite). We also note that although the particles are bosons, in the ground state they behave like fermions.

One possible explicit manifestation of a residual particle number density in the ground state is dark matter. Observations have shown that dark matter accounts for about 90% of visible matter in our universe, and the matter is clearly not of baryonic origin (see, e.g., Ref. 12). Its existence is usually related to the presence of various hypothetical particles (Higgs particles, axions, etc.), which for various reasons cease to interact with ordinary matter. But if this mass is ascribed to the ground state, then first it becomes obvious that the matter is truly dark, and second that the minimum set incorporating only the particles known at present is sufficient.

To describe the properties of dark matter, we begin with massive bosons ($m\neq 0$). For the sake of approximation, we ignore the possible dependence of the parameter μ on the wave number k. In this case, to avoid obtaining too large a value for dark matter, we require that

$$\mu^2-m^2=z^2\ll m^2. \tag{38}$$

Then the ground state contains only one particle per mode in the wave-number range $k^2\leqslant z^2$, where $N_k=2$. In other words, massive bosons in the ground state behave like an ordinary degenerate Fermi gas, and we obtain for the energy density and particle number density

$$\begin{aligned}\varepsilon&=\frac{1}{L^3}\sum_{n,k}n\omega_kN(k,n)\\&=\frac{g}{2\pi^2}\left(\frac{z^3\mu}{4}+\frac{m^2}{8}\left(z\mu-m^2\ln\frac{z+\mu}{m}\right)\right),\end{aligned} \tag{39}$$

$$n=\frac{1}{L^3}\sum_{n,k}nN(k,n)=\frac{g}{6\pi^2}z^3, \tag{40}$$

where g is the number of polarization states. In the limit $z\ll m$, this expression leads to the well-known nonrelativistic relationship

$$\varepsilon=nm+\frac{3}{2}p,\quad p=\frac{g}{30\pi^2}\frac{z^5}{m}, \tag{41}$$

where p is the gas pressure. The principal contribution to the ground-state energy density is provided by the rest mass of the particle, i.e., in leading order this contribution comes from dust. Note, however, that the particle pressure is nonzero, and it yields a small correction of order $p/\varepsilon\sim z^2/m^2\sim n^{2/3}/m^2$.

We now study particles with zero rest mass (such as photons and gravitons). For the ground-state energy density we have

$$\varepsilon=\frac{g}{2\pi^2}\frac{\mu^4}{4}\xi(3). \tag{42}$$

The number density of vacuum particles is

$$n=\frac{g}{2\pi^2}\frac{\mu^3}{3}\xi(2), \tag{43}$$

where

$$\xi(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}.$$

The equation of state in this case is ultrarelativistic ($\varepsilon=3p$).

Massless particles are especially interesting, since one can also measure the intensity of quantum fluctuations of the field potentials, which for the ground state (34) are

$$\langle\varphi(x),\varphi(x+r)\rangle=\frac{1}{(2\pi)^2}\int_0^{\infty}\frac{dk}{k}\frac{\sin kr}{kr}\Phi^2(k), \tag{44}$$

where

$$\Phi^2(k)=k^2N_k=k^2\left(1+\left[\frac{\mu}{k}\right]\right).$$

At long wavelengths $k \ll \mu$, a substantial increase in the level of quantum fluctuations should be observable in comparison with pure vacuum noise $(\mu=0)$.

6. CONCLUDING REMARKS

We see then that the concept of spacetime foam introduced by Wheeler should lead to a number of observable effects in particle theory. The simplest are the emergence of dark matter and an increase in the intensity of quantum noise in the field potentials. In Sec. 5 we calculated such effects under the assumption that the field is in the ground state. However, the results can easily be generalized to a situation in which the state of the fields is characterized by nonzero temperature T^*. Since processes associated with changes in the topology of space are the first to stop in the early stages of the evolution of the universe, we expect that $T^* \ll T_\gamma$ (T_γ is temperature of the microwave background radiation). On the other hand, given the value of μ in Eq. (42), we can obtain an upper bound $\mu^* \sim 60T_\gamma$. Thus, we expect T^* to be much less than μ, and the temperature corrections to the ground state (34) to be small. Note, however, that the nature of the fluctuations of the field potentials in (44) can change substantially if the temperature is nonzero.[6]

In addition to the effects studied in this paper, there clearly remain many phenomena that require additional investigation. For example, given the existence of self-action, the ground state (34) can be transformed, which can lead to the emergence of scalar Higgs fields (by analogy with the well-known Cooper effect in superconductivity). Such fields are needed, in turn, to generate particle masses in grand unification theories. Note that in fields with self-action, a nonvanishing particle number density in the ground state, in and of itself, leads to the emergence of massive excitations, although the upper bound on masses that can be derived from cosmological constraints on the value of μ is many orders of magnitude less than the values observed in particle theory. Another possibility is that when measuring the Casimir force,[8,9] one must expect an anomalous dependence on distance at scales exceeding the value of μ.

The author is grateful to D. Turaev for useful discussions at all stages of the research, and to M. Reiner for an invitation to Potsdam University, where a substantial portion of this research was done. This work was supported by grants from the Russian Fund for Fundamental Research (Grant No. 98-02-16273) and DFG (Grant No. 436 RUS 113/236/0(R)).

---

[*] E-mail: kirillov@focus.nnov.ru, kirillov@unn.ac.ru

[1] The set A can be called an elementary system of quantum numbers.

[2] When speaking of the topology of the physical space, we mean either the topology of the space M_{ph} itself, or the topology of the related space M^*_{ph}, depending on the quantities being measured. Note, however, that the relationship between these two spaces is nontrivial.

1. J. A. Wheeler, in *Relativity, Groups, and Topology*, B. S. and C. M. DeWitt (eds.), Gordon & Breach, New York (1964); S. W. Hawking, Nuclear Phys. **114**, 349 (1978).
2. A. A. Kirillov, Phys. Lett. B **399**, 201 (1997); A. A. Kirillov and G. Montani, JETP Lett. **66**, 475 (1997).
3. S. W. Hawking, Phys. Rev. D **37**, 904 (1988); Nucl. Phys. B **335**, 155 (1990); S. Giddings and A. Strominger, Nucl. Phys. B **307**, 854 (1988); G. V. Lavrelashvili, V. A. Rubakov, and P. G. Tinyakov, Nucl. Phys. B **299**, 757 (1988); S. Coleman, Nucl. Phys. B **310**, 643 (1988); T. Banks, Nucl. Phys. B **309**, 493 (1988).
4. L. J. Garay, Phys. Rev. Lett. **80**, 2508 (1998).
5. K. Wilson, Phys. Rev. B **10**, 2445 (1974); V. Jones, Ann. Math. **126**, 59 (1987); E. Witten, Commun. Math. Phys. **121**, 351 (1980); C. Rovelli and L. Smolin, Nucl. Phys. B **331**, 80 (1990).
6. A. A. Kirillov, Gravit. Cosmol. **2**(5), 123 (1996); Astron. Astrophys. Trans. **10**, 95 (1996).
7. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Nonrelativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).
8. H. B. G. Casimir, Proc. Kon. Nederl. Akad. Wet. **51**, 793 (1948).
9. M. J. Sparnaay, Physica (Amsterdam) **24**, 751 (1958).
10. A. A. Starobinsky, Phys. Lett. B **91**, 100 (1980); A. H. Guth, Phys. Rev. D **23**, 347 (1981); A. D. Linde, Phys. Lett. B **108**, 389 (1982).
11. A. D. Linde, *Particle Physics and Inflationary Cosmology*, Harwood Academic, New York (1990).
12. G. Borner, *The Early Universe: Facts and Fiction*, Springer-Verlag, New York (1992).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Initial mass spectrum of black holes in galactic nuclei

Yu. N. Eroshenko*) and M. I. Zel'nikov†)

P. N. Lebedev Physical Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 17 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 1935–1949 (June 1999)

Within the framework of the cosmological model with cold dark matter we have calculated the initial mass function of supermassive black holes formed in galactic nuclei. The collapsing region is modeled by a homogeneous ellipsoid. It is assumed that the accumulation of angular momentum by the proto-object takes place under the action of external tidal forces, and that surmounting of the centrifugal barrier with subsequent gravitational collapse occurs as a result of turbulent viscosity. To determine the mass function, we first find the angular momentum distribution function of the nascent objects for an arbitrary spectrum of initial density perturbations. The initial mass function is compared with available observations, and some processes leading to its transformation are indicated. © *1999 American Institute of Physics.* [S1063-7761(99)00206-1]

1. INTRODUCTION

Data obtained with the aid of the Hubble telescope and a number of ground-based telescopes[1] indicate that the nuclei of ten galaxies contain supermassive black holes (BH) with masses $M_h \sim (10^6-10^{11})M_\odot$. There are also less definite indications of the presence of black holes in several doze other galaxies. The technique for determining the masses of black holes is based on a study of the dynamics of stars and gas near the centers of galaxies and on a number of other approaches. There are reasons to believe that the fraction of the galaxies containing black holes in their nuclei is not less than 20% (Refs. 1 and 2), so that the formation of a black hole in the nucleus of a galaxy is not an exceptional event but rather a regular phenomenon in the history of almost every galaxy.

Several models of black-hole formation have been proposed.[3–5] These models are not mutually exclusive; therefore black holes in different mass ranges and even of the same mass have possibly a different origin. According to the model that we will use in the present work, a black hole arises as the result of collapse of a gas cloud long before the formation of a stellar galaxy.[5] The possibility of such a collapse depends on the angular momentum collected under the action of external tidal forces. At the center of a virialized halo of dark matter the baryonic matter forms a self-gravitating rotating disk. Compression of the central region of this disk before the relativistic state is reached occurs thanks to an outward transfer of angular momentum due to turbulent viscosity. A black hole is formed if the disk is able to compress before the moment at which it would otherwise fragment into stars and before its disruption by supernova flareups. Thus, only objects with sufficiently small angular momentum collapse.

The accumulation of statistical material in recent years has made it possible to detect definite correlations between the masses of black holes and the characteristics of the galaxies containing them. The authors of Ref. 6 obtained an approximate relation $M_h \simeq 0.003 M_b$ between the mass of a black hole M_h and the mass M_b of the stellar component of a galaxy. A similar relation $M_h \simeq 0.006 M_b$ was obtained in Ref. 2 on the basis of a study of 32 galaxies. Correlations of a similar sort (albeit approximate) together with the known luminosity function of galaxies have made it possible to obtain an estimate for the mass distribution of black holes.[7] Any theory of galaxy formation must first address the problem of determining the mass function of the black holes. Efforts to calculate the mass function have been undertaken, e.g., in Refs. 8 and 9; however, the approaches proposed in these works do not take sufficient account of the statistical aspects of black-hole formation. Reference 8 does not take account of the angular-momentum distribution of the objects involved. In Ref. 9, within the framework of a linear theory, found only the asymptotic limit of the angular momentum (L) distribution at large L and assumed it to be valid in the limit $L \rightarrow 0$. Besides, Ref. 9 utilized an inadequate criterion for the formation of a black hole (black holes are formed when L does not exceed the maximum angular momentum of a Kerr black hole). The goal of the present paper is a more accurate calculation of the black-hole mass function by first finding the angular-momentum distribution for small L taking account of the nonlinear stage of evolution of a density perturbation. In contrast to Ref. 9, we employ what is in our opinion a more likely criterion of black-hole formation.

The model of accretion of matter to a black hole successfully explains many observed properties of quasars (their high luminosity, variability, and emission spectrum). Accretion in the Eddington regime, as is well known, leads to exponential growth of the black-hole mass. Another source of transformation of the mass spectrum is collisions and coalescence of galaxies with large red shifts.[10,11] Interacting galaxies, and also the remnants of such interactions have been directly observed. Black holes existing at the centers of galaxies before coalescence should, as a result of dynamic friction, settle rapidly into the center of the new galaxy and also coalesce. Both processes—growth of mass during accretion, and coalescence of black holes—lead to a redistribution of the black-hole mass function toward larger masses. Therefore, the presently observed black-hole mass spectrum should differ from the mass spectrum of nascent black holes. A comparison of the initial black-hole mass function calculated by us with the observed mass function, subjected to transformation, can thus be useful in studies of the coalescence of galaxies and accretion mechanisms in quasars.

This paper is organized as follows. Section 2 examines the time dependence of the moment of the tidal forces and obtains an approximate formula for the integral of this moment over time. Section 3 finds the joint probability distribution for quantities characterizing the inner and outer density distributions. Section 4 carries out a statistical average and obtains the angular-momentum distribution of the objects of interest. Section 5 briefly describes some physical processes leading to formation of black holes, and estimates the critical value of the angular momentum. Section 6 calculates the initial mass function and provides an interpretation of the results obtained. Throughout this article we consider a flat cosmological model with cold dark matter without a Λ term.

2. MODEL OF THE PROTO-OBJECT

Let us consider an individual peak in the density distribution of the matter from which a gravitationally bound object is formed after collapse. We divide the region of space near the peak into inner and outer regions relative to a sphere of radius R such that the larger part of the mass that later falls into the object is found in the inner region. The potential of the outer part, representing the field of tidal forces, can be expanded in spherical harmonics[12]

$$\phi_{\mathrm{sh}}=\sum_{l,m}\frac{4\pi G}{2l+1}a_{lm}Y_{lm}|\mathbf{r}|^{l}, \tag{1}$$

where

$$a_{lm}=-\rho_b\int_{|\mathbf{r}|>R}d^3rY^{*}_{lm}\delta(\mathbf{r})|\mathbf{r}|^{-l-1}, \tag{2}$$

G is the gravitational constant, $\rho(\mathbf{r})$ is the density at the point $\mathbf{r}$, ρ_b is the mean cosmological density, and $\delta(\mathbf{r})\equiv[\rho(\mathbf{r})-\rho_b]/\rho_b$. The term with $l=0$ does not affect the compression dynamics, and the term with $l=1$ is responsible only for motion of the center of mass and does not affect the accumulation of angular momentum. According to Ref. 13, terms with $l\geqslant 3$ can be neglected in a treatment of tidal forces.

We model the inner collapsing region by a homogeneous ellipsoid. For the accumulation of angular momentum by the proto-object, of decisive importance is the presence of a quadrupole moment interacting with the tidal forces. We may nominally distinguish two sources of the quadrupole moment: nonsphericity of the proto-object and inhomogeneity of the inner distribution of matter.[14] The model of a homogeneous ellipsoid to a significant extent takes account only of the first of these factors. Despite the indicated shortcoming, the model of a homogeneous ellipsoid is the only model currently available that allows a detailed analysis of the dynamics of nonlinear compression. We will use it under the assumption that it gives a reasonable quantitative estimate at least for the time dependence of the moment of the tidal forces.

Equations for the evolution of a homogeneous ellipsoid were derived, for example, in Refs. 12 and 15. During the evolution of the ellipsoid the total mass inside the ellipsoid is conserved: $M=\rho_iV_i=M_e+\rho_bV$, where V is the volume of the ellipsoid and M_e is the excess mass in the ellipsoid above the background mass. The subscript i here and in subsequent formulas means that the quantity is taken at some initial time t_i that can be chosen close to the recombination time. We introduce the following parametrization of the coordinate r^α of the mass element of the ellipsoid:

$$r^{\alpha}=S^{\alpha\beta}(t)x^{\beta}, \tag{3}$$

where $|\mathbf{x}|\leqslant 1$. The quadratic potential

$$\phi=\frac{1}{2}\Phi_{\alpha\beta}(t)r^{\alpha}r^{\beta} \tag{4}$$

consists of the potential of the homogeneous ellipsoid, the potential of the homogeneous background, and the potential of the tidal forces:

$$\Phi=\Phi_{\mathrm{el}}+\Phi_b+\Phi_{\mathrm{sh}}, \tag{5}$$

$$\Phi_b=4\pi G\rho_b(t)I, \tag{6}$$

I is the unit matrix. Let the semi-axes of the ellipsoid be equal to a, b, and c. In the Cartesian coordinate system associated with the principal axes of the ellipsoid, we have

$$S=\left\|\begin{matrix}a&&\\&b&\\&&c\end{matrix}\right\|,\quad \Phi_{\mathrm{el}}=2\pi G\rho_e\left\|\begin{matrix}A_1&&\\&A_2&\\&&A_3\end{matrix}\right\|, \tag{7}$$

where

$$A_1=abc\int_0^{\infty}\frac{d\lambda}{(a^2+\lambda)[(a^2+\lambda)(b^2+\lambda)(c^2+\lambda)]^{1/2}}, \tag{8}$$

A_2 and A_3 are written analogously, and $\rho_e\equiv M_e/V$.

The equations of evolution of the ellipsoid are

$$\frac{d^2S^{\alpha\beta}}{dt^2}=-\Phi^{\alpha\gamma}S^{\gamma\beta}. \tag{9}$$

Formation of a black hole corresponds to a small angular momentum and correspondingly a small moment of the tidal forces. We assume that they are so weak that their influence on the dynamics of the halo of dark matter can be neglected. In this case the angular momentum due to the tidal forces becomes a significant factor only in the further compression of the baryon component. The initial velocities are written as

$$\left.\frac{dS^{\alpha\beta}}{dt}\right|_i=\left[HS^{\alpha\beta}-\frac{2}{3H}(\Phi^{\alpha\gamma}-\Phi_b^{\alpha\gamma})S^{\gamma\beta}\right]_i, \tag{10}$$

where H is the Hubble constant.

Black holes emerge from relatively high density peaks. In this situation the deviation of the shape of the object from spherically symmetrical is not large.[16] Therefore we seek the solution of Eqs. (9) and (10) in the limit of small deviations of the shape of the ellipsoid from spherically symmetrical, i.e., we represent $S^{\alpha\beta}$ in the form

$$S^{\alpha\beta}=IR+q, \tag{11}$$

where

$$q=\begin{Vmatrix} q_1 & & \\ & q_2 & \\ & & q_3 \end{Vmatrix}, \tag{12}$$

and $q_1^2+q_2^2+q_3^2 \ll R^2$. We introduce the notation $p\equiv q_1+q_2+q_3$. Then the matrix Φ on the right-hand side of Eq. (9) takes the form

$$\Phi=2\pi G\left\{\rho_i \frac{R_i^3}{R^3}\left(1+\frac{p_i}{R_i}\right)-\rho_b(t)\left(1+\frac{p}{R}\right)\right\} \times\left\{\left(\frac{2}{3}-\frac{2}{5}\frac{p}{R}\right)I-\frac{4}{5}\frac{q}{R}\right\}+\frac{4\pi}{3}G\rho_b(t)I. \tag{13}$$

In the zeroth approximation we neglect nonsphericity, setting $q=0$ and $p=p_i=0$. The solution in this case is well known[15] and is expressed by the parametric formulas

$$R=R_i\frac{3}{5\delta_i}\cos^2\vartheta,$$

$$\vartheta+\frac{1}{2}\sin\vartheta=\frac{2}{3}\left(\frac{5\delta_i}{3}\right)^{3/2}\frac{t-t_s}{t_i}, \tag{14}$$

where t_s is the time of maximum expansion:

$$t_s=t_i\left[1+\frac{3\pi}{4}\left(\frac{5\delta_i}{3}\right)^{-3/2}\right], \tag{15}$$

δ_i is the relative density fluctuation at the time t_i.

We will consider nonsphericity effects in the following approximation which is linear in q. Combining the equations of system (9) with the initial conditions (10), we obtain

$$\ddot{p}=a_1p+a_2p_i, \quad \dot{p}(t_i)=a_3p_i, \quad p(t_i)=p_i, \tag{16}$$

where the quantities a_1, a_2, and a_3 depend on the components of q. If at the initial time t_i we choose a sphere of radius R such that $p_i=0$, then it follows from Eqs. (16) that $p(t)=0$ for any time t. Then q satisfies the system of equations

$$\ddot{q}=q\frac{4\pi G}{5}\left(\frac{1}{3}\frac{R_i^3}{R^3}\rho_i-2\rho_b(t)\right),$$

$$\dot{q}(t_i)=\frac{2}{3t_i}\left(1+\frac{\delta_i}{15}\right)q_i,$$

$$q(t_i)=q_i. \tag{17}$$

The relation between R and t is given by Eqs. (14). The given system of differential equations can be integrated numerically, e.g., by the Runge–Kutta method. We require the solution in order to calculate the angular momentum. We will now clarify how the momentum of the tidal forces depends on q.

The moment of forces acting on the ellipsoid is expressed by an integral over its volume:

$$K_\alpha=-\int d^3r\rho_b\delta(\mathbf{r})[\mathbf{r}\nabla\phi_{\rm sh}(\mathbf{r})]_\alpha =-\frac{M_e}{5}e_{\alpha\beta\gamma}\Phi_{\gamma\sigma}(t)S_{\sigma\varepsilon}(t)S_{\beta\varepsilon}(t). \tag{18}$$

Substituting the expressions for Φ_{sh} and S, we obtain

$$K_\alpha=-\frac{2\sqrt{6\pi}GM_e}{5\sqrt{5}}\begin{Vmatrix}(c^2-b^2)\,\mathrm{Im}\,a_{21}\\(c^2-a^2)\,\mathrm{Re}\,a_{21}\\(a^2-b^2)\,\mathrm{Im}\,a_{22}\end{Vmatrix}. \tag{19}$$

The time dependence of a_{2m} in the linear approximation has the form[14]

$$a_{2m}(t)=a_{2m,i}\left(\frac{t}{t_i}\right)^{-4/3}. \tag{20}$$

Formula (20) is applicable for sufficiently isolated density peaks, as should hold for the majority of cases of black-hole formation.

We have the approximate equality $a^2-b^2\approx 2R(q_1-q_2)$, and analogously for c^2-b^2 and c^2-a^2. Equations (17) are linear in q; therefore the time dependence of q_i-q_j is the same as that of q_1. We assume that the virialization of the object sets up with its compression to roughly half its radius, starting at the time of maximum expansion.[15] The virialization time t_v depends only on δ_i and corresponds to $\vartheta=\pi/4$. Thus,

$$L_\alpha=K_{\alpha,i}t_i\int_1^{t_v/t_i}\frac{R}{R_i}\frac{q_1(t)}{q_{1,i}}\left(\frac{t}{t_i}\right)^{-4/3} \times\left[1+\delta_i-\left(\frac{t_i}{t}\right)^2\left(\frac{R}{R_i}\right)^3\right]d\left(\frac{t}{t_i}\right)=K_{\alpha,i}t_if(\delta_i). \tag{21}$$

The expression in brackets is the time dependence of M_e in formula (19). By numerical integration with accuracy acceptable for the subsequent calculations, we obtain

$$f(\delta_i)\approx 1.9\times 10^{-2}\delta_i^{-1}. \tag{22}$$

3. JOINT PROBABILITY DISTRIBUTION

Black holes are formed near peaks of the density field

$$\delta(\mathbf{r})=\frac{1}{(2\pi)^3}\int\delta_{\mathbf{k}}e^{i\mathbf{kr}}d^3k. \tag{23}$$

The density perturbations are assumed to be Gaussian, and their statistics are determined by assigning the power spectrum $P(k)$:

$$\langle\delta_{\mathbf{k}}^*\delta_{\mathbf{k}'}\rangle=(2\pi)^3P(k)\delta_D^{(3)}(\mathbf{k}-\mathbf{k}'), \tag{24}$$

where $\delta_D^{(3)}(\mathbf{k}-\mathbf{k}')$ is the Dirac delta function, and the angular brackets denote ensemble averaging. The spectral moments are defined as

$$\sigma_j^2 \equiv \int \frac{k^2 dk}{2\pi^2} P(k) k^{2j}. \tag{25}$$

We expand the density field (23) in a power series about the point $\mathbf{r}_p$:

$$\delta(\mathbf{r}_p + \mathbf{r}) = \delta(\mathbf{r}_p) + \eta_i r_i + \frac{1}{2} \zeta_{ij} r_i r_j + \ \ldots\ . \tag{26}$$

If $\mathbf{r}_p$ is the position vector of the density extremum, then $\boldsymbol{\eta} = 0$. The expression for the concentration distribution of the maxima of the density field, obtained in Ref. 16 and generalized with the angular momentum taken into account, has the form

$$n(\mathbf{r}, L_{z0}, \nu_0) dL_z d\nu = |\det(\zeta_{ij})| \, \delta_D^{(3)}(\boldsymbol{\eta}) \, \theta(\lambda_1) \, \theta(\lambda_2) \, \theta(\lambda_3) \times \delta_D(\nu - \nu_0) \, \delta_D(L_z - L_{z0}) dL_z d\nu, \tag{27}$$

where $\nu \equiv \delta(\mathbf{r}_p)/\sigma_0$, λ_i are the eigenvalues of the matrix ζ_{ij}, and θ is the Heaviside step function.

Our immediate goal is to construct the joint probability distribution for the quantities characterizing the density field inside the sphere of radius R centered at the point $\mathbf{r}_p$ and the quantities characterizing the tidal forces, i.e., the variables $\delta(\mathbf{r}_p)$, η_i, ζ_{ij}, and a_{2m} or others that are expressed in terms of them. We find the correlator

$$\langle \zeta_{ij} a_{2m} \rangle = C \int d\Omega_{\mathbf{k}} n_i n_j Y_{2m}^*(\Omega_{\mathbf{k}}), \tag{28}$$

where $n_i = k_i/k$ and $\Omega_{\mathbf{k}}$ is the solid angle in momentum space,

$$C \equiv \frac{\rho_b}{16\pi^5} \int_0^\infty dk k^4 P(k) \, \frac{kR \cos(kR) - \sin(kR)}{(kR)^3}. \tag{29}$$

In the course of the calculation we took the Fourier transform, used formula (24) to calculate the average, and applied the relation

$$\int e^{i\mathbf{kr}} Y_{lm}(\theta, \varphi) d\Omega = 4\pi Y_{lm}(\theta_{\mathbf{k}}, \varphi_{\mathbf{k}}) i^l j_l(kr), \tag{30}$$

where $j_l(x)$ are the spherical Bessel functions and $d\Omega$ is the element of solid angle.

If we take some definite linear combinations of the correlators (28) with fixed m, then the corresponding linear combinations of the expressions $n_i n_j$ give the spherical functions Y_{2n}. By virtue of the orthogonality of the spherical functions it is thus possible to reduce a large subset of the correlators to zero. The form of the necessary linear combinations can be seen by writing down Y_{2n} in Cartesian coordinates r_1, r_2, and r_3, substituting $r_1 r_1$ for ζ_{11}, $r_1 r_2$ for ζ_{12}, etc. Thus we introduce the new variables $\vec{\xi}$:

$$\xi_i = \begin{pmatrix} \xi_3 \\ \xi_2 \\ \xi_1 \\ \xi_0 \\ \xi_{-1} \\ \xi_{-2} \end{pmatrix} = \frac{1}{\sigma_2} \begin{pmatrix} -(\zeta_1 + \zeta_2 + \zeta_3) \\ -(\zeta_1 - \zeta_2)/2 + i\zeta_6 \\ \zeta_5 + i\zeta_4 \\ -(\zeta_1 - 2\zeta_3 + \zeta_2)/2 \\ \zeta_5 - i\zeta_4 \\ (\zeta_1 - \zeta_2)/2 - i\zeta_6 \end{pmatrix}. \tag{31}$$

Here we have changed over to the new notation $\zeta_{ij} \to \zeta_A$, $A = \overline{1,6}$: $\zeta_1 \equiv \zeta_{11}$, $\zeta_2 \equiv \zeta_{22}$, $\zeta_3 \equiv \zeta_{33}$, $\zeta_4 \equiv \zeta_{23}$, $\zeta_5 \equiv \zeta_{13}$, $\zeta_6 \equiv \zeta_{12}$. The choice of the normalized coefficients is made in line with the notation of Ref. 16, from which we will use some results in what follows. We write the inverse transformation in the form

$$\zeta_A = \sigma_2 \sum_{n=-2}^{3} U_A^{(n)} \xi_n, \tag{32}$$

where $U_A^{(n)}$ is a constant matrix. With the help of $U_A^{(n)}$ we express $n_i n_j$ in terms of Y_{2n}:

$$n_i n_j = \left(\sum_{n=-2}^{2} U_{ij}^{(n)} D^{(n)} Y_{2n} \right) + U_{ij}^{(3)} D^{(3)} Y_{00}, \tag{33}$$

where the vector

$$D^{(n)} = \sqrt{\frac{8\pi}{15}} \left(-\sqrt{\frac{15}{2}}; 1; -1; \frac{\sqrt{6}}{2}; 1; 1 \right). \tag{34}$$

We find the correlators

$$\langle \xi_m a_{2n} \rangle = \frac{C D^{(n)}}{\sigma_2} \delta_{mn}, \quad \langle \xi_m \xi_n^* \rangle = \frac{D^{(n)2}}{4\pi} \delta_{mn},$$
$$\langle a_{2m} a_{2n}^* \rangle = \Psi \delta_{mn}, \tag{35}$$

where

$$\Psi \equiv \frac{\rho_b^2}{4\pi^4} \int_0^\infty dk k^2 P(k) \left[\frac{kR \cos(kR) - \sin(kR)}{(kR)^3} \right]^2. \tag{36}$$

To diagonalize the correlator matrix, we introduce in place of a_{2m} the new variables

$$v_n = \alpha (a_{2n}^* - f_n \xi_n), \quad n = \overline{-2,2}. \tag{37}$$

It follows from the requirement $\langle v_n \xi_m^* \rangle = 0$ that $f_n = \alpha \langle a_{2n} \xi_n \rangle / \langle \xi_n \xi_n^* \rangle$. For such a choice of f_n we have

$$\langle v_m v_n^* \rangle = \alpha^2 (\Psi - 4\pi C^2/\sigma_2^2) \delta_{mn} = \delta_{mn}, \tag{38}$$

if we take $\alpha = (\Psi - 4\pi C^2/\sigma_2^2)^{-1/2}$. We normalize ξ_n: $w_n \equiv \sqrt{4\pi} \xi_n / D^{(n)}$; then $\langle w_m w_n \rangle = \delta_{mn}$. The joint probability distribution of the quantities v_n and w_n (for the time being without w_3) has the form

$$P_v P_w d^5 v \, d^5 w = (2\pi)^{-5} \times \exp\left(-\frac{1}{2} \vec{v}^{\,+} \vec{v} - \frac{1}{2} \vec{w}^{\,+} \vec{w} \right) d^5 v \, d^5 w, \tag{39}$$

the cross denotes the Hermitian conjugate, and $\vec{v}$ and $\vec{w}$ are vectors with components v_n and w_n, respectively.

4. ANGULAR MOMENTUM

By virtue of the isotropicity of the density field distribution, the distribution function F over the components of the angular momentum depends only on the magnitude of the angular momentum $L^2=L_x^2+L_y^2+L_z^2$. We denote

$$\tilde{F}(L_z^2)\equiv\int F(L_x^2+L_y^2+L_z^2)dL_x dL_y$$

$$=2\pi\int_0^\infty F(L_*^2+L_z^2)L_* dL_*, \tag{40}$$

where $L_*^2=L_x^2+L_y^2$. Differentiating with respect to L_z^2, we obtain

$$F(L^2)=-\frac{1}{\pi}\left.\frac{d\tilde{F}(L_z^2)}{dL_z^2}\right|_{L_z^2=L^2}. \tag{41}$$

Thus, to find the total distribution function $F(L^2)$ it suffices to know the distribution in L_z.

We introduce the angular momentum operator $\hat{l}_z=-i[\mathbf{r}\nabla]_z$. The spherical functions are the eigenfunctions of this operator. Thus, the z component of the moment of the tidal forces takes the form

$$K_z=-\int_{|\mathbf{r}|<R}d^3r\rho_b\delta(\mathbf{r})[\mathbf{r}\nabla\phi_{\mathrm{sh}}(\mathbf{r})]_z$$

$$=-i\rho_b\int_{|\mathbf{r}|<R}d^3r\,\delta(\mathbf{r})\hat{l}_z\phi_{\mathrm{sh}}(\mathbf{r}). \tag{42}$$

As was indicated in the previous section, it is sufficient to restrict ourselves to the terms with $l=2$. Making the necessary substitutions of variables, we obtain

$$K_z=-i\frac{2\pi G\rho_b}{5}\sum_{m=-2}^{2}a_{2m}\sum_{ij}\zeta_{ij}\int_{|\mathbf{r}|<R}d^3r r^2 r_i r_j\hat{l}_z Y_{2m}(\Omega_{\mathbf{r}})$$

$$=i\frac{2\pi G\rho_b R^7}{35}\sum_{m=-2}^{2}\sum_{A=1}^{6}\sum_{k=-2}^{2}ma_{2m}\sigma_2\xi_k U_A^{(k)}U_A^{(m)*}D^{(m)}$$

$$=-i\frac{2\pi G\rho_b R^7}{35\sqrt{4\pi}}\sum_{m=-2}^{2}\sum_{A=1}^{6}\sum_{k=-2}^{2}\left[\frac{m}{\alpha}w_k D^{(k)}U_A^{(k)}U_A^{(m)*}\right.$$

$$\left.\times D^{(m)}v_m^*+\frac{\sqrt{4\pi}\,mC}{\sigma_2}w_k D^{(k)}U_A^{(k)}U_A^{(m)*}D^{(m)}w_m^*\right]. \tag{43}$$

To calculate the integral over time of this expression, we make use of the results of Sec. 2. Note that in such an approach, to determine the initial value of the moment of the tidal forces we employ the quadratic approximation of the exact peak profile, and to find the time dependence we employ the approximation of the homogeneous ellipsoid. Thus we find that the angular momentum is equal to $L_z=K_{zi}t_i f(\delta_i)$, where $f(\delta_i)$ is given by formula (22). We write it conditionally in the form

$$L_z=w_k^*E_{km}v_m+w_k^*E'_{km}w_m, \tag{44}$$

where E_{km} and E'_{km} are matrices which are independent of $\vec{v}$ and $\vec{w}$.

In the averaging of expression (27) the following mean of the delta function arises:

$$\langle\delta_D(L_z-L_z(\vec{v},\vec{w}))\rangle_{P_v}$$

$$=\int\frac{d^5v}{(2\pi)^{5/2}}\exp\left(-\frac{1}{2}\vec{v}^{\,+}\vec{v}\right)\int\frac{ds}{2\pi}$$

$$\times\exp[is(L_z-w_k^*E_{km}v_m-w_k^*E'_{km}w_m)]$$

$$=\int\frac{ds}{2\pi}\exp\left[is(L_z-w_k^*E'_{km}w_m)-\frac{1}{2}s^2w_kE_{km}^*E_{lm}w_l^*\right]$$

$$=\frac{1}{(2\pi)^{1/2}2^{3/2}Xy}\exp\left(-\frac{L_z^2}{16X^2y^2}\right), \tag{45}$$

where

$$X\equiv\frac{\pi^{1/2}G\rho_b R^7}{35}\left(\frac{8\pi}{15}\right)^2(\sigma_2^2\Psi-4\pi C^2)^{1/2}t_i f(\delta_i). \tag{46}$$

After integrating over d^5v, the coordinate axes can be aligned with the eigenvectors of the matrix ζ_{ij}, which significantly simplifies the process of integrating over s. In the last equality in Eqs. (45) we changed over to the following notation (see also Ref. 16):

$$\sigma_2x=-(\zeta_1+\zeta_2+\zeta_3),\quad \sigma_2y=-(\zeta_1-\zeta_2)/2,$$

$$\sigma_2z=-(\zeta_1-2\zeta_3+\zeta_2)/2, \tag{47}$$

$$\gamma\equiv\sigma_1^2/\sigma_2\sigma_0,\quad x_*\equiv\gamma\nu,\quad R_*\equiv\sqrt{3}\sigma_1/\sigma_2. \tag{48}$$

Further statistical averaging of expression (27)

$$N(L_z,\nu)=\langle n(\mathbf{r},L_z,\nu)\rangle \tag{49}$$

is realized by the following integral over x, y, and z (a way of obtaining the normalization factor can be found in Ref. 16):

$$N(L_z,\nu,x,y,z)dL_z d\nu$$

$$=\int\frac{5^{5/2}3^{1/2}}{(2\pi)^{7/2}2^{3/2}}\left(\frac{\sigma_2}{\sigma_1}\right)^3\frac{1}{Xy(1-\gamma^2)^{1/2}}$$

$$\times e^{-Q}F(x,y,z)\chi dL_z d\nu dx dy dz, \tag{50}$$

where

$$Q\equiv\frac{\nu^2}{2}+\frac{(x-x_*)^2}{2(1-\gamma^2)}+\frac{5}{2}(3y^2+z^2)+\frac{L_z^2}{16X^2y^2}, \tag{51}$$

$$F(x,y,z)=y(x-2z)[(x+z)-(3y)^2](y^2-z^2). \tag{52}$$

The characteristic function χ assigns the region of integration:

$$\chi:\quad y\geqslant 0,\quad y\geqslant z\geqslant -y,\quad x+z\geqslant 3y. \tag{53}$$

The integral over y can be done analytically. We will not write out the result in light of its cumbersomeness.

We are interested in objects with angular momentum much less than the mean value. We introduce the smallness parameter

$$\varepsilon \equiv \frac{L_z^2}{16X^2} \ll 1 \tag{54}$$

and expand expression (50) integrated over y in a series in ε. The validity of such an expansion is verified by expanding expression (50) in powers of x and z at the point $x = z = 0$. By virtue of the additional differentiation in expression (41) the zeroth term of the expansion (independent of ε) does not contribute. Let us limit ourselves to the term linear in ε. We integrate over z and x numerically and find an approximate function for the dependence of the result on the variables γ and x_* within the ranges of variation of these latter quantities of interest. Finally, employing expression (41), we obtain the expression

$$N(L,\nu)dL\,d\nu = 2.4\times 10^{-2}(1-\gamma^2)^{0.17}\frac{x_*^{2.84}}{X^3 R_*^3} e^{-\nu^2/2} L^2 dL\,d\nu. \tag{55}$$

The accuracy of approximation formula (55) for the values of the quantities entering into it used in our calculations is not worse than 20%. In the following sections on the basis of this distribution we investigate the statistics of black holes in galactic nuclei.

5. FORMATION OF BLACK HOLES

The main processes leading to formation of black holes at the centers of galaxies have been discussed, in particular, in Refs. 5, 17, and 18. Our treatment is similar to that given in Ref. 5 except that we have used equations for a self-gravitating disk without predominance of the central mass. Let us consider an object with total mass M (dark matter + baryons). Let the baryonic matter make up the fraction f_g of the mass: $M_g = f_g M$. The gas cloud cools rapidly thanks to inverse Compton scattering by photons of the microwave background (fossil radiation) and to emission from free–free and free–bound transitions. During cooling the baryonic matter falls toward the center of the object and reaches the centrifugal barrier at the radius

$$r_d \simeq \frac{L^2}{GM^3 f_g}, \tag{56}$$

where L is the total angular momentum of the object with mass M. It is assumed that the angular-momentum L density per unit mass is identical for baryons and for dark matter.

At the center of the object the baryonic matter forms a self-gravitating rotating disk. Further compression of the central region of the disk takes place thanks to outward transfer of angular momentum. We assume that the main factor in the transfer of angular momentum is turbulent viscosity. The turbulence is usually characterized by the parameter $\alpha_t = v_t/v_s \sim 0.1$ (Ref. 19), where v_t is the velocity of the turbulent pulsations, $v_s = (5k_B T/3m_p)^{1/2}$ is the velocity of sound, k_B is the Boltzmann constant, and m_p is the mass of the proton.

The thickness of the self-gravitating disk is

$$H_d \simeq r_d \frac{v_s^2}{v_\phi^2}, \tag{57}$$

where v_ϕ^2 is the angular velocity of rotation of the disk. The coefficient of kinematic viscosity and the characteristic compression time are equal respectively to

$$\mu \simeq \frac{\alpha_t r_d^2 v_s^3}{GM_g}, \quad t_{\text{vis}} \simeq \frac{GM_g}{\alpha_t v_s^3}. \tag{58}$$

In order to find v_s it is necessary to consider the process of radiative transfer of the gravitational energy being liberated outward during compression. The rate of liberation of energy is

$$\frac{dE_{\text{grav}}}{dt} \simeq E_{\text{grav}} t_{\text{vis}}^{-1} \simeq 2\pi r_d^2 \sigma T_s^4, \tag{59}$$

where $E_{\text{grav}} \simeq GM_g^2/r_d$, σ is the Stefan–Boltzmann constant, T_s is the temperature on the surface of the disk. The relation between the effective temperature at the center T_c and the effective temperature on the surface T_s has the form $T_c^4 \sim \tau T_s^4$ (Ref. 20), where $\tau = \sigma_T M_g / 2\pi r_b^2 m_p \gg 1$ is the optical thickness of the plasma disk and σ_T is the Thomson cross section.

From the above relations it is possible to obtain the compression time t_{vis} of the disk due to viscous transfer of angular momentum. We assume that a black hole is formed if this time does not exceed the characteristic time of star formation.[21]

$$t_* = \tilde{t}\left(\frac{M_c}{M_\odot}\right)^{-\beta}, \tag{60}$$

where M_c is the characteristic mass of a star nucleus. In the simplest model $\tilde{t} \approx 5\times 10^7$ years and $\beta \approx 2.8$ (Ref. 21). If the stars have the standard Salpeter mass function, then we can expect $M_c \sim 0.5 M_\odot$. However, the conditions in a compact disk are substantially different from those in ordinary star-forming regions in galaxies; therefore M_c at this stage of development of the theory should be treated as a free parameter.

From the condition $t_{\text{vis}} < t_*$ we obtain $L < J$, where

$$J\left[\frac{\text{g}\cdot\text{cm}^2}{\text{s}}\right] = 6.1\times 10^{68}\left(\frac{\alpha_t}{0.1}\right)^{4/15}\left(\frac{f_g}{0.03}\right)^{2/3}\left(\frac{\tilde{t}}{5\times 10^7 \ \text{years}}\right)^{1/6} \times \left(\frac{M_c}{M_\odot}\right)^{-\beta/6}\left(\frac{M}{10^8 M_\odot}\right)^{23/15}. \tag{61}$$

If the reverse inequality holds, then the disk fragments into stars and instead of a black hole a star cluster is formed.

A second limitation arises from the requirement that the gas be ionized after the virialization time of the dark halo. In the opposite case cooling and compression of baryons will not take place. Ionization can be a consequence of secondary ionization of the Universe by the first pregalactic generation of stars.[17,18] In this case, however, the concentration of nascent black holes of small mass significantly exceeds the limits of observation. A second source of ionization is heating of the gas during virialization of the object under the action of shock waves and vigorous mixing. We take this ionization source to be defining. We find the temperature from the relation $m_p v^2/2 = 3k_B T/2$, where v is the virial velocity. The condition $T > T_{\mathrm{ion}} \sim (1-5)\times 10^4$ K yields

$$\delta_l \equiv \delta_i(1+z_i) > \frac{27}{5\sqrt{2}} \frac{k_B T_{\mathrm{ion}}}{m_p} \left(\frac{t_0}{GM}\right)^{2/3}, \tag{62}$$

where T_{ion} is the ionization temperature, t_0 is the current time, and $1+z_i \equiv (t_0/t_i)^{2/3}$.

We have investigated formation of black holes within the framework of a simple model without a detailed treatment of the internal structure or evolution of the disk. The choice of such a model is based on the absence of detailed studies of real models of self-gravitating disks. For example, the authors of Ref. 22 found a self-similar solution for the evolution of an isothermal disk in the weak-accretion limit and under the assumption that $\mu(r)H_d(r)/r = \mathrm{const}$. The solution obtained correctly reflects some qualitative aspects of the evolution of a disk; however, use of the quantitative results of Ref. 22 is hardly justified in a real situation. Accurate solution of the problem may require complicated, three-dimensional hydrodynamic calculations allowing for thermal processes. In any additional calculations what is important for us is the presence of a maximum value of the angular momentum. In quantitative estimates we will use results obtained within the framework of the above model.

6. BLACK-HOLE MASS SPECTRUM

In Secs. 3 and 4 the power spectrum $P(k)$ is taken to be arbitrary. For concrete calculations we will use the spectrum arising in the model with cold dark matter with the Hubble constant $H = 75\,\mathrm{km}\cdot\mathrm{s}^{-1}\cdot\mathrm{Mpc}^{-1}$ (Ref. 16):

$$P(k) = \frac{a[\ln(1+4.164k)]^2}{k(1+6.947k+828.9k^2+925.4k^3+20710k^4)^{1/2}}, \tag{63}$$

where k is measured in the co-moving coordinates and in units of Mpc^{-1}. The normalization constant $a \approx 2.13\times 10^5$ is determined, as usual, from the requirement that the relative fluctuation of the mass on a scale of 8 Mpc be unity.

We smooth the density field with the help of two filters

$$\delta_R(\mathbf{r}) = \int \delta(\mathbf{r}_1) W_1(\mathbf{r}_1-\mathbf{r}_2) W_2(\mathbf{r}_2-\mathbf{r}) d^3r_1 d^3r_2, \tag{64}$$

where

$$W_1(\mathbf{r}_1-\mathbf{r}_2) = \frac{3}{4\pi R^3}\theta\left(1-\frac{|\mathbf{r}_1-\mathbf{r}_2|}{R}\right),$$

$$W_2(\mathbf{r}_2-\mathbf{r}) = \frac{1}{(2\pi\tilde{R}^2)^{3/2}}\exp\left(-\frac{|\mathbf{r}_2-\mathbf{r}|^2}{2\tilde{R}^2},\right). \tag{65}$$

Use of the filter W_1 allows us to interpret $\nu\sigma_0$ directly as the density averaged over a sphere of radius R. It is specifically this quantity that defines the moment of collapse and is denoted in Sec. 2 [formula (22)] as δ_i. In the final calculations we transform to co-moving coordinates and express all quantities in terms of quantities at the current time t_0. The Gaussian filter W_2 with $\tilde{R} = R/10^{1/3}$ has only practical significance and allows us to avoid problems with rapid oscillations in the integrand during numerical calculation of the quantities σ_j. By virtue of the fact that the filter W_2 smooths out masses $M/10 \ll M$, the presence of the filter W_2 has only a weak influence on the result.

The smallness parameter ε (54) used in the calculations does not exceed 5×10^{-4} for masses varying within the limits $10^7 M_\odot < M_h < 5\times 10^{11} M_\odot$, which justifies the assumption of inequality (54). The angular momentum collected by the object depends on the orientation of its quadrupole moment relative to the outer density distribution. However, as calculation shows, these quantities correlate only weakly. Indeed, their mutual correlation is characterized by the expression $4\pi C^2$ in relation (46). For the masses under consideration here, $4\pi C^2/\sigma_2^2\Psi < 1/700$. Thus, the outer density distribution can be considered to be statistically independent of the inner distribution. The explanation of this fact probably lies in the statistical independence of fluctuations on different scales. Evidence in favor of this is provided by the results of Ref. 13, where it was found that tidal forces generated by spherical shells of radii r and $2r$ are anticorrelated, and the main contribution comes from density fluctuations a distance $5R$ away from the object of radius R and corresponding to a mass of 5^3M.

In studies of a hierarchical mechanism of formation of the large-scale structure of the Universe the problem of "cloud within a cloud" is well known.[23] The statement of this problem is that an object of mass M_1 can become incorporated in an object of mass $M_2 > M_1$ at a later time. In connection with this the problem arises of calculating the concentration of independent objects. It is natural to interpret formula (55) as the concentration of objects with mass $> M$. To calculate the mass (M) distribution it is necessary to differentiate formula (55) with respect to M. The dependence on ν in formula (55) has the form $\nu^{5.84}\exp(-\nu^2/2)$; this function attains its maximum at $\nu \approx 2.4$ in line with the fact that black holes are formed from relatively high density peaks. The final distribution is given by the formula

$$\phi(M)dM = dM \int_0^J dL \frac{d}{dM} \int_{\tilde{\delta}/\sigma_0}^{\infty} d\nu N(L,\nu), \tag{66}$$

where $\tilde{\delta} \equiv \max\{\delta_c, \delta_l\}$. The value $\delta_c = 1.69$ corresponds to collapse at the current time. To transform to the mass distribution of the black holes $M_h = f_g M$ it is necessary to make the corresponding change of variables:

$$\phi_1(M_h)d(\ln M_h) = \phi(M_h/f_g)d(M_h/f_g).$$

The result of the calculation is shown in Fig. 1. For comparison, the dashed curve *3* depicts the assumed distribution obtained in Ref. 7 based on observational data. The difference in the curves can be attributed to inaccuracies in the observations or in the subsequent transformation of the mass function. The figure demonstrates the possible direction and magnitude of such a transformation. It is interesting to observe that for masses $M_h > 10^8 M_\odot$ according to formula (66) the initial black-hole spectrum is close to a power-law: $\phi \propto M^{-1.84}$.

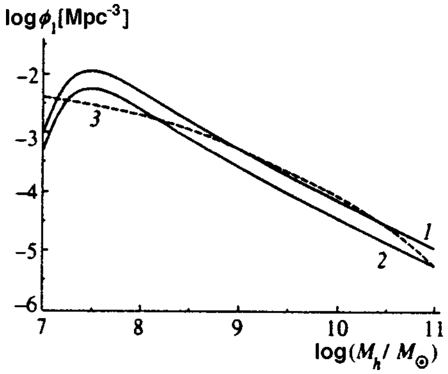

FIG. 1. Mass distribution ϕ_1 of black holes in galactic nuclei. Curves *1* and *2* plot the results obtained using formula (66) for $M_c=0.3M_\odot$ $M_c=0.5M_\odot$, respectively. For comparison, the dashed curve depicts the distribution obtained in Ref. 7.

Quantity (66) is proportional to the following combination of parameters of our model:

$$\left(\frac{\alpha_t}{0.1}\right)^{4/5}\left(\frac{f_g}{0.03}\right)^{2}\left(\frac{\tilde{t}}{5\times 10^{7}\ \text{years}}\right)^{1/2}\left(\frac{M_c}{M_\odot}\right)^{-\beta/2}. \qquad (67)$$

For definiteness we set all parameters except for M_c to be equal to their normalized values in expression (67). Thus, depending on the mass, the relationship between the initial mass function and the distribution in Ref. 7 can be understood as follows. It is clear from the figure that if $M_c \approx 0.3M_\odot$ then the calculated mass function reproduces the distribution from Ref. 7 quite well for $5\times 10^8 M_\odot < M_h < 5\times 10^{10} M_\odot$. Note that it is precisely in this mass range that the largest number of black holes have been recorded. The statistics outside the indicated range are incomplete. The difference in the distributions may also be due to observational selection at small black-hole masses. If $M_c > 0.3M_\odot$, then for small masses the initial mass function is also substantially greater than its values obtained from the observations,[7] but for large masses it is substantially less. Processes that can lead to such a transformation were indicated in the Introduction, these being growth of the mass of black holes as a result of their coalescence and as a result of accretion. In the hierarchical picture of the formation of the large-scale structure of the Universe every galaxy during its existence can experience up to ten coalescence events with other galaxies.

We obtained a similar picture of the transformation of the mass function on the basis of a comparison of an observational estimate of the black-hole mass distribution and the luminosity and red-shift distributions of quasars. Of fundamental significance is the question of the duration t_q of the Eddington accretion stage. For $t_q \ll t_{\mathrm{Ed}}$ (t_{Ed} is the Eddington time) the masses of black holes remain essentially unchanged after their formation. On the contrary, for $t_q \geqslant t_{\mathrm{Ed}}$ the masses grow significantly and the black-hole mass spectrum observed at the current time differs strongly from the mass spectrum of nascent black holes. Calculations show that almost all black holes with masses $M > 3\times 10^8 M_\odot$ arose from less massive black holes as a result of accretion. The concentration of nascent black holes originally formed with masses $M > 3\times 10^8 M_\odot$ can be neglected in comparison with the concentration of black holes that have entered this range from below.

The authors express their gratitude to the Russian Fund for Fundamental Research for its financial support in the form of Grants No. 96-15-96614 and 96-02-16670.

*)E-mail: erosh@td.lpi.ac.ru
†)E-mail: zelnikov@td.lpi.ac.ru

---

[1] R. P. Van Der Marel, in *Proceedings of the IAU Symposium 186*, Kyoto, August 1997, edited by D. B. Sanders and J. Barnes, Kluwer Academic Publishers, (in press); E-prints archive, astro-ph/9712076.
[2] J. Magorrian, S. Tremaine, D. Richstone, *et al.*, Astron. J. **115**, 2285 (1998).
[3] M. J. Rees, Annu. Rev. Astron. Astrophys. **22**, 471 (1984).
[4] A. V. Gurevich and K. P. Zybin, Zh. Eksp. Teor. Fiz. **97**, 20 (1990) [Sov. Phys. JETP **70**, 10 (1990)].
[5] D. J. Eisenstein and A. Loeb, Astrophys. J. **443**, 11 (1995).
[6] J. Kormendy and D. Richstone, Annu. Rev. Astron. Astrophys. **33**, 581 (1995).
[7] P. Salucci, E. Szuszkiewicz, P. Monaco, and L. Danese, E-prints archive astro-ph/9811102
[8] M. G. Haehnelt and M. J. Rees, Mon. Not. R. Astron. Soc. **263**, 168 (1993); E-prints archive astro-ph/9712259.
[9] H. Susa, M. Sasaki, and T. Tanaka, Prog. Theor. Phys. **92**, 961 (1994).
[10] D. S. Krivitsky and V. M. Kontorovich, E-prints archive astro-ph/9801195.
[11] M. G. Haehnelt, E-prints archive astro-ph/9809328.
[12] D. J. Eisenstein and A. Loeb, Astrophys. J. **439**, 520 (1995).
[13] T. Quinn and J. Binney, Mon. Not. R. Astron. Soc. **255**, 729 (1992).
[14] B. S. Ryden, Astrophys. J. **329**, 589 (1988).
[15] P. J. E. Peebles, *The Large-Scale Structure of the Universe* (Princeton University Press, Princeton, N.J., 1980).
[16] J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay, Astrophys. J. **304**, 15 (1986).
[17] A. Loeb, Astrophys. J. **403**, 542 (1993).
[18] A. Loeb and F. A. Rasio, Astrophys. J. **432**, 52 (1994).
[19] N. I. Shakura and R. A. Sunyaev, Astron. Astrophys. **24**, 337 (1973).
[20] Ya. B. Zel'dovich and Yu. P. Raĭzer, *Physics of Shock Waves and High-Temperature Hydrodynamic Phenomena* (Fizmatgiz, Moskva, 1963).
[21] V. G. Surdin, *The Birth of Stars*, URSS, Moskva (1997).
[22] S. Mineshige and M. Umemura, Astrophys. J. **480**, 167 (1997).
[23] J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser, Astrophys. J. **379**, 440 (1991).

Translated by Paul F. Schippnick

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Dynamics of Mandel'shtam–Brillouin induced scattering during self-focusing of a laser beam

N. E. Andreev and M. V. Chegotov

Scientific-Research Center for the Thermal Physics of Pulsed Actions, Unified Institute of High-Temperature Physics, Russian Academy of Sciences, 127412 Moscow, Russia

L. M. Gorbunov*)

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 17 August 1998)
Zh. Éksp. Teor. Fiz. **115**, 1950–1960 (June 1999)

It is shown that the size of the focal spot has a substantial influence on the dynamics of Mandel'shtam–Brillouin induced scattering (MBIS) for the laser beam power near critical for striction self-focusing. For small focal spots MBIS suppresses self-focusing. An increase in the size of the focal spot leads to growth of the MBIS pulsations and the steady-state setup time. For large enough focal spots MBIS arises in the form of regular intense spikes. Physical processes shaping the dynamics of MBIS are discussed. © *1999 American Institute of Physics.* [S1063-7761(99)00306-6]

1. INTRODUCTION

In nonlinear optics one of the most widely known processes is Mandel'shtam–Brillouin induced scattering (MBIS), in which an incident electromagnetic wave, by interacting with sound waves, creates scattered electromagnetic waves with shifted frequencies (see, e.g., Refs. 1–3). This process is observed in many material media and is of great significance for a number of applications. In particular, it is used for phase conjugation (see, e.g., Refs. 4–6). It is also of great significance for laser nuclear fusion, lowering the fraction of radiation absorbed in the target (see, e.g., Ref. 7).

If the laser beam power P exceeds a certain critical value P_{cr} (Refs. 8–10), then self-focusing can also take place along with MBIS. It is well known that in steady state either self-focusing (like filamentation of the laser beam) leads to a growth of MBIS (see, e.g., Refs. 11–13) or MBIS suppresses self-focusing (see, e.g., Refs. 14–16). However, the interaction of MBIS and self-focusing can be most uniquely manifested in their simultaneous development in time during a transitory process. Thus, in Ref. 17 it was shown that because of self-focusing MBIS acquires the form of periodic spikes during which the intensity of the scattered radiation can exceed the intensity of the incident radiation. A nonmonotonic time dependence of MBIS is also indicated by numerical calculations.[18]

The present paper discusses the possibility of modifying the dynamics of MBIS when the laser beam power exceeds the critical power of striction self-focusing by varying the size of the focal spot. It follows from a numerical solution of the system of nonlinear equations describing the incident beam and scattered beams and also large-scale perturbations of the density of the medium associated with self-focusing that for relatively large focal spots the development of self-focusing leads to the result that MBIS has the form of periodic intense spikes.[17] As the size of the focal spot is reduced, self-focusing is suppressed and the spiking character of MBIS gives way to oscillations about some mean value, where the amplitude of these oscillations decays with time. For a fixed focal spot size self-focusing is suppressed that much more effectively, the higher the initial level of scattering. In this case, the setting up of steady state takes place faster. Under conditions in which the initial MBIS reflection coefficient, calculated in the one-dimensional theory neglecting self-focusing, amounts to several percent, the dynamics of MBIS approaches the dynamics described by the one-dimensional nonlinear theory.[19]

Numerical calculations of the spatiotemporal variation of the intensity of the incident and scattered radiation, and also the density of the medium, have made it possible to interpret the physical processes responsible for the above-described MBIS dynamics. These questions, and also experiments in which MBIS pulsations have been observed, are discussed in the Conclusion.

2. STATEMENT OF THE PROBLEM, AND BASIC EQUATIONS

Let us consider a planar layer of a nonlinear, transparent medium, onto which, starting at the time $t=0$, a beam of electromagnetic radiation having characteristic width $2a$ (diameter of the focal spot) is incident. We assume that the radiation power exceeds its critical value for striction (ponderomotive) self-focusing and the thickness of the layer exceeds the diffraction length. Together with self-focusing, we consider MBIS in directions close to directly backward.[1)]

To describe the incident and scattered beams, and also perturbations of the medium density, we use, respectively, Maxwell's equations and the equations of acoustics, in which we allow for the action of the averaged ponderomotive force (see, e.g., Refs. 2 and 3). We represent the electric field strength in the medium in the form

$$\mathbf{E}(\mathbf{r},t)=\frac{1}{2}\sum_{\sigma=\pm1}\{\mathbf{E}_\sigma(\mathbf{r},t)\exp[-i\omega_\sigma t+i\sigma k_\sigma z]+\text{c.c.}\},\tag{1}$$

where $\mathbf{E}_\sigma$ is the complex amplitude of the spatiotemporally slowly varying (on scales of k_σ^{-1} and ω_σ^{-1}, respectively) field of the incident ($\sigma=1$) and scattered ($\sigma=-1$) beams; ω_σ and k_σ are respectively the frequencies and wave numbers, which are interrelated by the dispersion relation $k_\sigma^2c^2=\omega_\sigma^2\varepsilon(\omega_\sigma)$; $\varepsilon(\omega)$ is the linear dielectric constant of the medium, which is assumed to be isotropic.

For simultaneous development of self-focusing and MBIS two types of density perturbations develop. Large-scale variations in the density of the medium $\delta\rho$ are responsible for self-focusing while small-scale sound waves are responsible for MBIS in the directions close to directly backward, where the density perturbations in these sound waves can be represented in the form

$$\delta\rho_s=\frac{1}{2}\{\delta\rho_{s0}(\mathbf{r},t)\exp[-i\omega_s t+ik_s z]+\text{c.c.}\},\tag{2}$$

where the frequencies of the sound waves $\omega_s=\omega_1-\omega_{-1}$ and their wave numbers $k_s=k_1+k_{-1}$ are interrelated by the dispersion relation $\omega_s=k_sV_s$ ($V_s=\sqrt{\partial p/\partial\rho}$ is the speed of sound), and $\delta\rho_{s0}$ is the amplitude of the sound wave.

If the mean free path of sound is small in comparison with those scales on which the amplitude of the sound wave $\delta\rho_{s0}$ varies, then the latter is expressed in terms of the amplitude $\mathbf{E}_\sigma$ as follows[1–3] 2):

$$\delta\rho_{s0}=\frac{ik_s}{16\pi\gamma_sV_s}\left(\rho\frac{\partial\varepsilon}{\partial\rho}\right)\mathbf{E}_1\mathbf{E}_{-1}^*,\tag{3}$$

where γ_s is the decay decrement of the sound waves, which is assumed to be small in comparison with ω_s; ρ is the density of the medium; and $\rho(\partial\varepsilon/\partial\rho)$ is the electrostriction coefficient.[2]

Substituting relations (1)–(3) in the Maxwell equations and discarding small terms proportional to the second derivative of the slowly varying amplitude $\mathbf{E}_\sigma$, we obtain

$$\left\{i\left(\frac{1}{V_g}\frac{\partial}{\partial t}+\sigma\frac{\partial}{\partial z}\right)+\frac{1}{2k}\Delta_\perp+\frac{\omega}{2V_g}\left(\frac{\partial\varepsilon}{\partial\rho}\right)\delta\rho+i\sigma\frac{\omega^2}{32\pi c^2\gamma_sV_s\rho}\left(\rho\frac{\partial\varepsilon}{\partial\rho}\right)^2|E_{-\sigma}|^2\right\}E_\sigma=0,\tag{4}$$

where, taking the inequality $\omega_s\ll\omega_1,\omega_{-1}$ into account, we assume that $\omega_1\approx\omega_{-1}=\omega$ and $-k_{-1}\approx k_1=k$; the group velocity is equal to

$$V_g=c[d(\omega\sqrt{\varepsilon(\omega)})/d\omega]^{-1},$$

and $\Delta_\perp$ is the transverse part of the Laplace operator

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right).$$

In Eqs. (4) self-focusing is taken into account by the third term inside the braces. The equation for the corresponding perturbation of the medium density has the form

$$\left(\frac{\partial^2}{\partial t^2}+2\Gamma\frac{\partial}{\partial t}-V_s^2\Delta_\perp\right)\delta\rho=-\frac{1}{16\pi}\left(\rho\frac{\partial\varepsilon}{\partial\rho}\right)\Delta_\perp\sum_{\sigma=\pm1}|E_\sigma|^2,\tag{5}$$

where Γ is the attenuation coefficient for the small-scale density perturbations.

Treatments of MBIS do not usually take account of self-focusing, and the term proportional to $\delta\rho$ in system of equations (4) is usually neglected. On the other hand, treatments of self-focusing neglect MBIS. This corresponds to discarding the fourth term in Eq. (4) for $\sigma=1$ and taking account only of the term with $\sigma=1$ on the right-hand side of Eq. (5).

For the numerical calculations, we represent Eqs. (4) and (5) in dimensionless form:

$$\left\{i\left(\frac{\partial}{\partial\tau}+\sigma\frac{\partial}{\partial\xi}\right)+\overline{\Delta}_\perp-A+i\sigma G|e_{-\sigma}|^2\right\}e_\sigma=0,\tag{6}$$

$$\left(\frac{1}{\beta^2}\frac{\partial^2}{\partial\tau^2}+\frac{2\overline{\Gamma}}{\beta}\frac{\partial}{\partial\tau}-\overline{\Delta}_\perp\right)A=\alpha\overline{\Delta}_\perp\sum_{\sigma=\pm1}|e_\sigma|^2,\tag{7}$$

where

$$\tau=tc/2ka^2,\quad \xi=z/2ka^2,$$

$$\overline{\Gamma}=\Gamma a/V_s,\quad e_\sigma=E_\sigma/E_m,$$

$$\overline{\Delta}_\perp=\frac{1}{\eta}\frac{\partial}{\partial\eta}\left(\eta\frac{\partial}{\partial\eta}\right),\quad \eta=r/a,$$

$$\beta=2ka\frac{V_s}{V_g},\quad G=\frac{kV_s}{\gamma_s}\alpha,$$

$$\alpha=\frac{1}{2}\left(\frac{a\omega}{c}\right)^2\frac{E_m^2}{8\pi\rho V_s^2}\left(\rho\frac{\partial\varepsilon}{\partial\rho}\right)^2,\quad A=-\left(\frac{\omega a}{c}\right)^2\delta\rho\frac{\partial\varepsilon}{\partial\rho},$$

E_m is the maximum amplitude of the electric field of the incident beam on the axis $\eta=0$, and a is the characteristic initial size of the focal spot at the boundary $\xi=0$.

In the solution of Eqs. (6) and (7) we assume that the incident beam at the left-hand boundary of the layer $\xi=0$ and the scattered "primer" beam at the right-hand boundary $\xi=\xi_0\gg1$ have planar wavefronts and a Gaussian intensity distribution in the radius:

$$e_1(\eta,\xi=0,\tau)=f(\tau)\exp(-\eta^2),$$

$$e_{-1}(\xi,\xi=\xi_0)=e_b\exp(-\eta^2/b^2).\tag{8}$$

Here the function $f(\tau)\le1$ describes the variation with time of the amplitude of the incident beam, b is the dimensionless width of the scattered "primer" beam. For $b\gg1$ the amplitude e_{-1} is essentially constant over the width of the incident beam and it is possible to speak of spatially homogeneous "priming." The boundary conditions in the radial variable corresponded to the symmetry of the fields and the density perturbations relative to the $\eta=0$ axis and to their falloff at infinity (in reality at some $\eta_{\max}\gg1$). Initial density perturbations $\delta\rho$ were assumed to be absent.

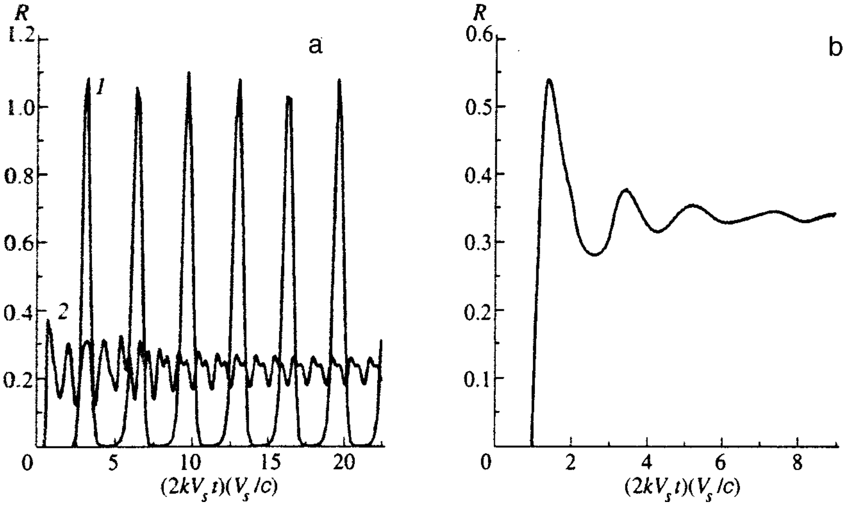

FIG. 1. Variation of the MBIS reflection coefficient R as a function of the dimensionless time $\tau = 2kV_s^2t/c$ for various widths of the focal spot: a — $G=15$, $|e_b|^2=10^{-7}$, $\beta=1.5$ (*1*), $\beta=0.5$ (*2*); b — $G=20$, $|e_b|^2=10^{-6}$, $\beta=1$.

3. RESULTS OF NUMERICAL CALCULATIONS

The critical power of self-focusing $P_{\rm cr}$ corresponds to a definite value of the parameter $\alpha=\alpha_{\rm cr}$. This value can be estimated if we neglect MBIS ($e_{-1}=0$), assuming that the beam is stationary [$\partial/\partial\tau=0$ in Eqs. (6) and (7)] and homogeneous in the longitudinal direction [$\partial/\partial\xi=0$ in Eq. (6)]. Employing expression (8), we find in the limit $\eta\rightarrow 0$ that $\alpha_{\rm cr}=4$. A more consistent analysis, correctly taking into account the radial dependence of the amplitude e_1 inside the medium, gives the value $\alpha_{\rm cr}=7.54$ (Ref. 10). In the calculations below we used the value of the parameter $\alpha=8.4 > \alpha_{\rm cr}$.

The quantity $\overline{\Gamma}$, characterizing the attenuation of the large-scale perturbations of the medium density, was taken to be equal to 0.25. Earlier it was shown[20] that for $\overline{\Gamma}<1$ attenuation has a weak effect on the dynamics of self-focusing.

In the results obtained below, growth of the amplitude of the incident beam with time was assumed to be instantaneous [in formula (8) the function $f(\tau)$ is equal to zero for $\tau\leqslant 0$ and unity for $\tau>0$]. Calculations with a linearly increasing function $f(\tau)$ on the interval $0\leqslant\tau\leqslant 50$ do not differ substantially from results obtained for a step function, in line with the results obtained in Refs. 20 and 21.

The system of equations (6), (7) was solved for various values of the three parameters: the parameter G, which for a fixed value of α is determined by the dissipation of the short-wavelength sound waves γ_s; the parameter $|e_b|^2$ characterizing the "primer" level of the scattered radiation; and the parameter β proportional to the width of the focal spot. In the absence of self-focusing ($A=0$) the parameters G and $|e_b|^2$ in the steady-state linear theory of MBIS [when $\partial/\partial\tau=0$ and $G=0$ for $\sigma=1$ in Eq. (6)] uniquely determine the level of the backscattered radiation for a given plasma length ξ_0:

$$R_0=|e_b|^2\exp[(G/2)\arctan(4\xi_0)],$$

where diffraction lowering of the pump field intensity has been taken into account. In the calculations we set $\xi_0=1$, which corresponds to the plasma length exceeding the diffraction length of the laser beam $ka^2/2$ by a factor of four. Note that for $\xi_0\geqslant 1$ the scattering level does not depend on the plasma length.

Figure 1 plots the dependence on the dimensionless time $(2kV_st)(V_s/c)$ of the total (integrated over the radial variable) MBIS backscattering coefficient R:

$$R=\frac{1}{S_0}\int_0^\infty d\eta\,\eta|e_{-1}(\xi=0,\tau,\eta)|^2, \tag{9}$$

where the quantity

$$S_0=\int_0^\infty d\eta\,\eta e^{-2\eta^2}$$

defines the total energy flux of the incident radiation. In Fig. 1a we used the parameter values $|e_b|^2=10^{-7}$ and $G=15$, which corresponds to $\gamma_s/2kV_s\approx 0.25$ and small initial scattering $R_0\approx 2\times 10^{-3}$. As can be seen from the figure, for a wider beam ($\beta=1.5$) the scattering coefficient pulsates in time (the so-called spiking regime of MBIS) and the maximal values of R exceed unity. As the width of the focal spot is decreased ($\beta=0.5$) the range of variation of R decreases and the spiking character of the scattering gives way to oscillations about a mean value. Here the amplitude of the oscillations of the function R falls off with time.

Similar properties of the MBIS coefficient were also obtained over a wide range of variation of the parameters G and $|e_b|^2$, where G varied from 10 to 20, and $|e_b|^2$ varied from 10^{-8} to 10^{-6}. In all of these cases the initial scattering level was less than or on the order of 0.1%.

A different picture arises when the MBIS level is sufficiently high even in the absence of self-focusing. Figure 1b shows the variation of R with time for $R_0=0.573$ ($\beta=1$, $G=20$, $|e|^2=10^{-6}$). Clearly, the process has a relaxational character and after a few oscillations the function R relaxes to a steady-state level.

In order to understand the physical processes that determine the properties of the reflection coefficient R, we have constructed isocontours of the amplitude of the radiation in the incident and scattered beams in the $(\beta^2\tau,\xi)$ plane for

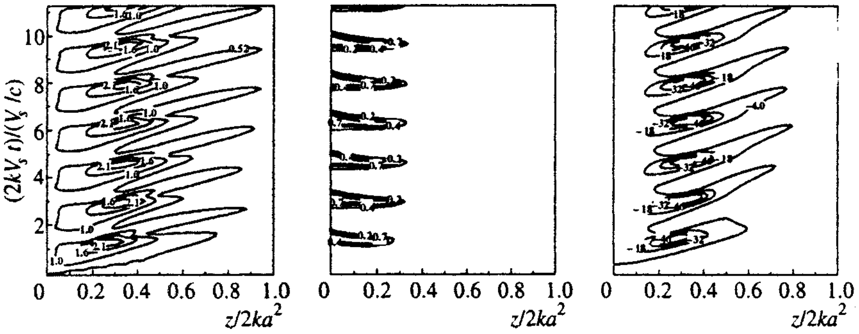

FIG. 2. Lines of constant values of the modulus of the amplitude of the incident (*a*) and scattered (*b*) radiation, density of the medium (*c*) in the $(2kV_s^2t/c,\ z/2ka^2)$ plane for $\eta=0$ (beam axis), $G=15$, $|e_b|^2=10^{-7}$ for a wide focal spot $\beta=1.5$.

$\eta=0$ (beam axis), and also for the large-scale density perturbations associated with self-focusing. Figure 2 plots such results for a relatively wide laser beam ($\beta=1.5$) for a low initial level of MBIS. The parameters G and $|e_b|^2$ were taken to be the same as in Fig. 1a. It can be seen that deep within the medium a region is periodically created in which the intensity of the incident radiation reaches a maximum [Fig. 2a] and the density of the medium, a minimum [Fig. 2c]. The location of this region varies slightly with time, and it is there mainly that growth of the intensity of the scattered radiation takes place [Fig. 2b].

For a narrower beam ($\beta=0.5$, Fig. 3) the process at first resembles what happened for a wide beam (Fig. 2) although it develops more slowly. Then, however, a state arises in which MBIS occurs mainly near the boundary [Fig. 3b], where the intensity of the incident radiation is the highest [Fig. 3a]. The dimensions of the region of amplification of the scattered radiation pulsate with time, and this is reflected in its intensity.

Note that the time-averaged MBIS coefficient depends only weakly on the parameter β. It is equal to 0.277 for $\beta=1.5$ and 0.226 for $\beta=0.5$.

Figure 4 plots the same functions as in Fig. 2, but for the parameters G and $|e_b|^2$ corresponding to Fig. 1c, where the scattering level is quite high even in the absence of self-focusing. It can be seen that although at the outset of its development the scattering dynamics are similar to what took place for a low initial MBIS level, the scattering region then localizes near the boundary of the nonlinear medium [Fig. 4b] and the variation of all the quantities in time becomes weaker and weaker and steady state is established comparatively rapidly.

As was already noted, switching on the incident beam gradually at the boundary, with characteristic time up to $\tau_0=50$, does not alter the character of the MBIS dynamics.

4. DISCUSSION

The results of numerical calculations can be understood by comparing the times characterizing the development of various nonlinear processes. Striction self-focusing is associated with a redistribution of the medium density in the radial direction and develops during a characteristic time $t_{sf}\approx a/V_s$ (Refs. 20–22). The transit time of radiation through the interaction region equal in length to the diffusion length $ka^2/2$ coincides in order of magnitude with the time of energy exchange, t_{NL}, between the incident and scattered waves due to MBIS. According to Ref. 19, a more accurate expression for it has the form

$$t_{NL}=\frac{4\pi ka^2}{V_gGp''}, \tag{10}$$

where p'' is the imaginary part of the complex solution of the equation

$$|e_b|^2\exp G-p(1-p+|e_b|^2)\exp(pG)=0. \tag{11}$$

The steady-state setup time, according to Ref. 19, is equal to

$$t_r=\frac{2ka^2}{V_gG|p_0-p'|}, \tag{12}$$

where p_0 and p' are the real root and real part of the complex root of Eq. (11). For the parameter $|e_b|^2$ varying within the interval from 10^{-6} to 10^{-8} and the parameter G varying within the interval from 10 to 20 the quantities Gp'' and $G|p_0-p'|$ vary respectively from 5.3 to 7.8 and from 2 to 24.5.

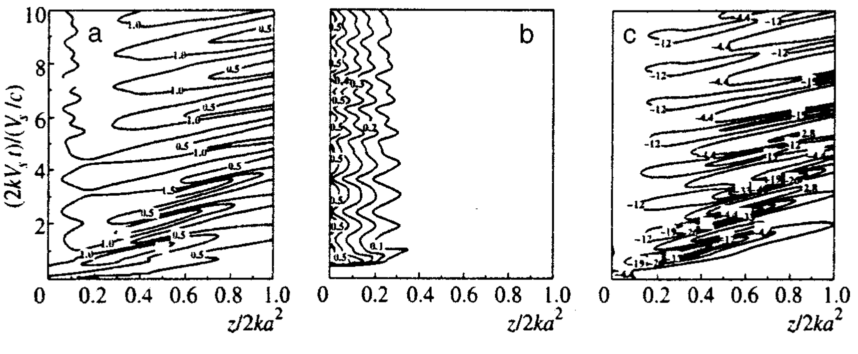

FIG. 3. The same as in Fig. 2, but for a narrow focal spot $\beta=0.5$.

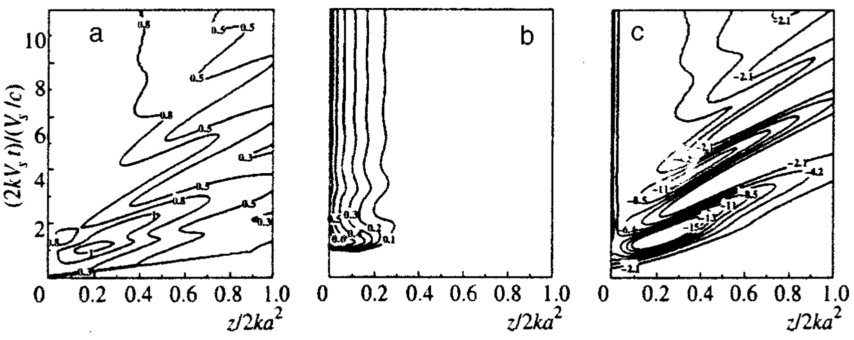

FIG. 4. The same as in Fig. 2, but for a relative high initial MBIS level. Parameters the same as in Fig. 1(b).

We may begin our discussion with a consideration of the case in which the initial MBIS level is quite low ($R_0 \leqslant 10^{-3}$) and the steady-state setup time (12) exceeds the times t_{sf} and t_{NL}. If $t_{sf}<t_{NL}$ and correspondingly, according to relation (10),

$$\beta=\frac{2kaV_s}{V_g}>\beta_{01}=\frac{Gp''}{2\pi}$$

(wide beam), then self-focusing develops faster than energy exchange takes place between the incident and scattered waves due to MBIS. This is the case to which Fig. 2 corresponds. In the initial stage of the process, MBIS has almost no effect on the development of self-focusing, and by the time t_{sf} a region is formed in which the density of the medium is lowered while the intensity of the incident radiation is increased.[20–22] However, thanks to the exponential dependence on $|e_1|^2$, starting at some time in this region, there is an abrupt growth in MB induced backscattering. The scattered radiation, propagating toward the boundary $\xi=0$ is amplified and depletes the energy of the incident beam. MBIS arises in the form of spikes [Fig. 1a]. During the time it takes a new portion of laser radiation to reach the self-focusing region the perturbation of the medium density has time to relax. The entire process begins practically anew. As a result, MBIS has the form of almost periodic short spikes of duration t_{sf}. This regime was discussed in Ref. 17.

Note that increasing the size of the focal spot leads, on the one hand, to an increase in the self-focusing time ($t_{sf} \propto a$) and, on the other, to an increase in the length of the region and the characteristic time of the interaction between the incident and scattered waves ($t_{NL} \propto a^2$). Since the latter time increases more rapidly with growth of a for sufficiently wide laser beams, the time t_{sf} is smaller than t_{NL} and the above-described region of pulsating MBIS should arise.

Let us turn now to the case $t_{sf} \geqslant t_{NL}$ and $\beta<\beta_{01}$ (narrow beam) depicted in Fig. 3. Self-focusing does not have time to develop completely during the time it takes MBIS to amplify to a level sufficient to substantially lower the intensity of the incident beam in the focal region. It can be said that for sufficiently narrow laser beams MBIS suspends and suppresses the development of self-focusing [Fig. 3c]. Only near the boundary of the medium [Fig. 3b] does a quasiperiodic process of energy exchange between the incident and scattered beams arise. The oscillations in the MBIS intensity in Fig. 1a for $\beta=0.5$ is quite well described by theory.[19] In particular, the mean level of the MBIS reflection $R=0.226$ is found to be in reasonable agreement with the nonlinear stationary theory,[23] according to which $R=0.18$ for the parameters corresponding to Fig. 1a.

The effect of suppression of self-focusing by MBIS is even more vividly manifested at larger gain coefficients G or at a higher primer scattering level $|e_b|^2$. Figure 4, which corresponds to Fig. 1b, depicts the dynamics of the process for those parameters for which the initial MBIS level is several percent. It can be seen that self-focusing is suppressed [Fig. 4c] and the function R has a relaxational character similar to that considered in Ref. 19. However, also in this case an increase in the size of the focal spot to a value such that $\beta=7$ leads to a spiking regime of MBIS.

Let us discuss the possibility of satisfying the condition

$$\beta=\beta_{01} \quad \text{or} \quad \frac{2kaV_s}{V_g}=\frac{Gp''}{2\pi}, \tag{14}$$

for which the oscillatory regime gives way to the spiking regime (see above) in the experiments. For liquids, solids, and gases the ratio of the velocity of light to the sound velocity is on the order of 10^5. For $k=1.26\times10^5\,\mathrm{cm}^{-1}$; (the second harmonic of a neodymium-glass laser) it follows from condition (14) that the diameter of the focal spot should exceed one centimeter, and the diffraction length, one kilometer. Under these conditions our assumption that the thickness of the layer of nonlinear medium exceeds the diffraction length is difficult to fulfill under laboratory conditions. In a hot, rarefied plasma the situation is different. Thus, at an electron temperature $T_e=1$ keV in a hydrogen plasma the ratio $V_g/V_s\approx3\times10^3$ and for the second harmonic of a neodymium laser, condition (14) is fulfilled for the diameter of the focal spot on the order of 50 μm, for which the diffraction length is around 8 mm.

Pulsations of the MBIS intensity have been experimentally observed more than once in a hot plasma (see, e.g., Refs. 24–28). However, the conditions in which these experiments were performed did not always correspond to the formulation discussed in this work. In Refs. 24 and 25 the plasma was inhomogeneous and the incident beam was partially reflected from the region with critical density. This created additional possibilities for the MBIS pulsations, which are discussed, for example, in Refs. 26 and 27. In other experiments the plasma was transparent to radiation, but its dimensions were less than the diffraction length, which suppressed self-focusing and increased the role of MBIS over a wide range of angles.[14,16] We know of only one experiment[28] in which the conditions corresponded to the formulation under discussion. The main characteristics of the experiment were: $\omega \approx 2\times 10^{14}\,\mathrm{s}^{-1}$ (CO_2 laser), ratio of the electron density n to its critical value $n_c = m\omega^2/4\pi e^2$ equal to 0.16; $a\approx 10^{-2}$ cm, $I=2\times 10^{11}-10^{12}\,\mathrm{W/cm}^2$, $T_e = 100\,\mathrm{eV}$; the length of the plasma was 1.2 cm and exceeded the diffraction length. In the experiment pulsations of MB induced backscattering were observed at a level of 10–20% with characteristic period on the order of 150 ps.

According to the definition given above, the quantity α can be written for a plasma in the form[29]

$$\alpha = \left(\frac{\omega_p a V_E}{2cV_T}\right)^2, \tag{15}$$

where ω_p is the plasma frequency, V_T and V_E are respectively the thermal velocity of the electrons and the velocity of their oscillations in the laser radiation field ($V_E = eE_m/m\omega$). Substituting the above values into formula (15), we find that for the conditions of the experiment in Ref. 28 $\alpha\approx 40 > \alpha_{\mathrm{cr}}$ for $I = 6\times 10^{11}\,\mathrm{W/cm}^2$. In addition, we find that for the given experiment $\beta \approx 0.04$ while the quantity β_{01} is on the order of 2. Since β is much less than β_{01}, the spiking regime of MBIS cannot be realized. An estimate of the period of the oscillations according to formula (10) gives a value on the order of 50 ps for $Gp''=5$. Thus, the results of the experiment in Ref. 28 correspond more closely to the MBIS regime considered in Ref. 19.

Note that Ref. 30 investigated experimentally the dynamics of MBIS in a rarefied plasma $[(\omega_p/\omega)^2 \approx 0.045]$ as a function of the means of focusing the radiation ($k \approx 10^5\,\mathrm{cm}^{-1}$, $a\approx 50-70\,\mu$m, $I\sim 10^{14}\,\mathrm{W/cm}^2$, $T_e \approx 1$ keV). The parameter β was equal approximately to 0.2, which visibly corresponds to the oscillatory regime of MBIS. However, the short duration of the laser pulse (~ 1.6 ns) did not allow the authors of Ref. 30 to observe variations in MBIS with time.

To realize the spiking regime of MBIS in a hot plasma with $V_s/c \approx 3\times 10^{-4}$ and $P/P_{\mathrm{cr}} \geqslant 1$ the ratio of the length of the plasma L to the radius of the focal spot a should exceed 1.5×10^3. If we assume that $a\approx 20\,\mu$m, we obtain a lower bound on the length of the plasma: $L \geqslant 3$ cm. An experiment with this goal, employing the scheme under discussion would make it possible, on the one hand, to check the validity of our picture of the dynamics of MBIS under conditions of self-focusing and, on the other, in the case of the spiked MBIS regime, to obtain short, intense pulses of backscattered radiation.

This work was carried out with partial financial support of the Russian Fund for Fundamental Research (Projects No. 97-02-16537 and 96-02-00021c).

*) E-mail: gorbun@lpi.ac.ru

1) If the layer thickness is much less than the diffraction length and is commensurate with the diameter of the focal spot, then it is necessary to allow for MBIS over a wide interval of angles as was done in Refs. 14 and 16.

2) For weakly attenuating sound waves the combined dynamics of MBIS and self-focusing in a plasma were investigated in Ref. 15.

---

[1] N. Bloembergen, *Nonlinear Optics* (Benjamin, New York, 1965) p. 150.
[2] V. S. Starunov and I. L. Fabelinskiĭ, Usp. Fiz. Nauk **98**, 441 (1969) [Sov. Phys. Usp. **12**, 463 (1970)].
[3] Y. R. Shen, *The Principles of Nonlinear Optics* (John Wiley & Sons, New York, 1984) p. 190.
[4] B. Ya. Zel'dovich, N. F. Pilipetskiĭ, V. V. Shkunov, *Principles of Phase Conjugation* (Springer-Verlag, New York, 1985).
[5] *Wavefront Reversal in Nonlinear Media*, edited by V. I. Bespalov [in Russian] (IPF AN SSSR, Gor'kiĭ, 1982).
[6] V. I. Bespalov and G. A. Pasmanik, *Nonlinear Optics and Adaptive Laser Systems* [in Russian] (Nauka, Moscow, 1986).
[7] K. Brakner and S. Dzhorna, *Controllable Laser Synthesis* [in Russian] (Atomizdat, Moskva, 1977) p. 43.
[8] G. A. Askar'yan, Zh. Éksp. Teor. Fiz. **42**, 1567 (1962) [Sov. Phys. JETP **15**, 1088 (1962)].
[9] V. I. Talanov, Pis'ma Zh. Éksp. Teor. Fiz. **2**, 222 (1965) [*sic*].
[10] V. N. Lugovoĭ and A. M. Prokhorov, Usp. Fiz. Nauk **111**, 203 (1973) [Sov. Phys. Usp. **16**, 658 (1973)].
[11] C. J. Randall, Laser Program Annual Report **2**, 3 (1979).
[12] M. S. Sodha, G. Umesh, and R. R. Sharma, J. Appl. Phys. **50**, 4678 (1979).
[13] H. A. Rose, Phys. Plasmas **2**, 2216 (1995).
[14] M. R. Amin, C. E. Capjack, P. Frycz *et al.*, Phys. Fluids B **5**, 3748 (1993).
[15] V. V. Eliseev, W. Rozmus, V. T. Tikhonchuk *et al.*, Phys. Plasmas **2**, 1712 (1995).
[16] S. Huller, Ph. Mounaix, and D. Pesme, Phys. Scr. **63**, 151 (1996).
[17] N. E. Andreev, L. M. Gorbunov, JETP Lett. **56**, 139 (1992).
[18] V. T. Tikhonchuk, S. Huller, and Ph. Mounaix, Phys. Plasmas **4**, 4369 (1997).
[19] M. V. Chegotov, Fizika plazmy **23**, 41 (1997) [Plasma Phys. Rep. **23**, 36 (1997)].
[20] N. E. Andreev, L. M. Gorbunov, S. V. Tarakanov *et al.*, Phys. Fluids B **5**, 1986 (1993).
[21] N. E. Andreev, L. M. Gorbunov, A. I. Zykov *et al.*, Zh, Éksp. Teor. Fiz. **106**, 1676 (1994) [JETP **79**, 905 (1994)].
[22] L. M. Gorbunov and S. V. Tarakanov, Zh. Éksp. Teor. Fiz. **99**, 58 (1991) [Sov. Phys. JETP **72**, 30 (1991)].
[23] C. L. Tang, J. Appl. Phys. **37**, 2945 (1966).
[24] L. M. Gorbunov, Yu. S. Kas'yanov, V. V. Korobkin *et al.*, JETP Lett. **27**, 226 (1978).
[25] V. L. Artsimovich, L. M. Gorbunov, Yu. S. Kas'yanov *et al.* Zh. Éksp. Teor. Fiz. **80**, 1859 (1981) [Sov. Phys. JETP **53**, 963 (1981)].
[26] S. Huller, P. Mulser, and A. M. Rubenchik, Phys. Fluids B **3**, 3339 (1991).
[27] A. V. Chirokikh, S. M. Kozochkin, A. P. Streltsov, B. D. Ochirov, and A. Rubenchik, Phys. Rev. Lett. **71**, 723 (1993).
[28] K. Henkel, Dissertation, Bochum (1992).
[29] A. J. Schmitt, Phys. Fluids **31**, 3079 (1988).
[30] O. Willi *et al.*, Phys. Fluids B **2**, 1318 (1990).

Translated by Paul F. Schippnick

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Field splitting of a nonabsorbing state

D. A. Shapiro*)

Institute of Automation and Electrometry, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia

(Submitted 14 July 1998)
Zh. Éksp. Teor. Fiz. **115**, 1961–1972 (June 1999)

This paper analyzes a four-level system interacting with four strong fields with frequencies combining into a cycle. The conditions for coherent population trapping, when a nonabsorbing superposition of states forms, are calculated. It is found that, in contrast to a three-level system, the nonabsorbing state splits, i.e., is realized at two values of the detuning. Such splitting manifests itself as narrow dips in the frequency dependence of the upper-level population. Similar dips are found in the spectrum of the nonlinear susceptibility, which is responsible for the conversion efficiency in the process of four-wave mixing. © *1999 American Institute of Physics.* [S1063-7761(99)00406-0]

1. INTRODUCTION

Nonlinear optics has three basic ways to change the frequency of coherent radiation: stimulated Raman scattering, parametric mixing, and higher harmonics generation. To achieve highly efficient conversion of radiation in the cw mode by any of these methods, the frequencies of the waves must be tuned in resonance with the atomic transitions so as to increase the nonlinear susceptibility. The gain may be substantial in gases, since the spectra of gases usually have narrow lines.

The most promising schemes are those of four-wave mixing with frequency subtraction, which occasionally make it possible to compensate for Doppler broadening.[1] A high efficiency of conversion of the light frequency has been achieved in pulsed experiments involving Pb vapor.[2] Quasi-cw generation has been demonstrated by Marangos *et al.*,[3] who used the rhombic scheme and atomic krypton. The cw mode of four-wave mixing with substantial upconversion in the optical range has been realized by Apolonsky *et al.*,[4] who used sodium dimers in the double Λ scheme. Both mixing schemes are depicted in Fig. 1. The cycle conditions for the double Λ scheme and the rhombic scheme are

$$\omega_a-\omega_b-\omega_c+\omega_d=0, \quad \omega_a-\omega_b+\omega_c-\omega_d=0, \tag{1}$$

where ω_a, ω_b, ω_c, and ω_d are the frequencies of the electromagnetic fields, and differ in the sequence of signs (either $\mp$ or $\pm$) in the third and fourth terms. Hence the schemes are sometimes called difference–sum and sum–difference, respectively. The conversion efficiency in the double Λ scheme proved to be high (25% in relation to the lowest intensity of the exciting fields[5]). However, efficient conversion is usually hindered by the resonant absorption of light by the medium.

To avoid absorption, Harris *et al.*[6] proposed using the effect of electromagnetically induced transparency, which emerges in the medium because of coherent population trapping (see also Refs. 2, 7 and 8). Coherent population trapping in three-level system has been thoroughly studied.[9,10] The physics of this effect amounts to interference of the quantum states mixed by a strong field. In a three-level Λ system, coherent population trapping manifests itself at equal detunings of the fields in relation to the corresponding transitions as a split-off state, a linear combination of the pair of lower states not interacting with the field. If the detuning of one field is fixed, the curve representing the fluorescence intensity as a function of the detuning of the second field exhibits a narrow deep dip, which manifests itself as a ‘‘dark’’ resonance. Generalization of the theory of coherent population trapping to *N*-levelsystems has also been discussed (see Ref. 10 and the literature cited therein). However, the role of coherent population trapping in resonant four-wave mixing remained unclear. In their work devoted to amplification without inversion in a double Λ scheme, Kocharovskaya and Mandel[11] took the criterion for coherent population trapping from the theory of three-level systems. On the other hand, in their studies of completely resonant four-wave mixing in a strong field, Coppeta *et al.*[12] ignored the criterion for coherent population trapping entirely.

In the present work we will calculate the nonlinear susceptibility in the simplest schemes of four-wave mixing with frequency subtraction. To this end, in Sec. 2 we will use the equations for the vector of the amplitudes without relaxation to derive, by a simple algebraic method, the criterion for coherent population trapping in the four-level system. We will find that the nonabsorbing state splits, i.e., a split-off state not interacting with the field is observed at two values of detuning. In Sec. 3 we will solve numerically the equations for the density matrix in the double Λ scheme with allowance for relaxation. We will find that the abrupt dip in the curve representing the frequency dependence of the upper-level splits into two components. There we will also calculate the manifestation of coherent population trapping in the spectral contours of nonlinear susceptibility. Section 4 will be devoted to a discussion of why splitting of a nonabsorbing state has not been observed in either measurements or computer simulations in four-wave mixing known from the literature. We will also establish the conditions for ex-

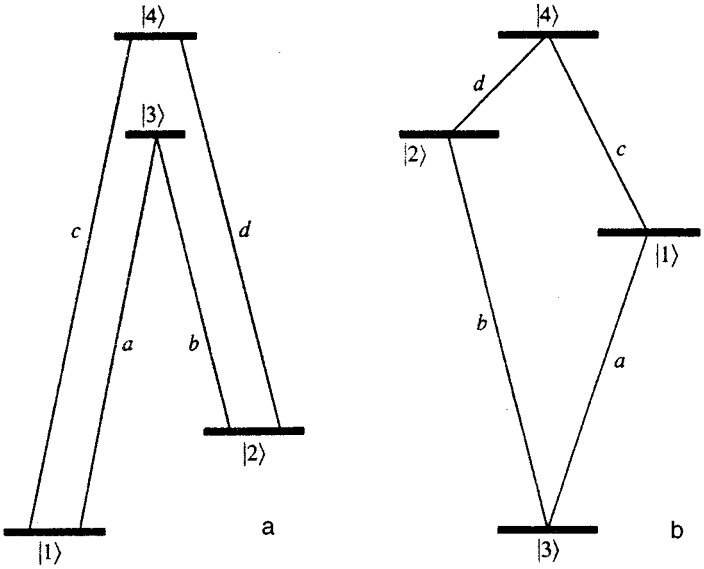

FIG. 1. Double Λ (a) and rhombic (b) schemes of four-wave mixing with frequency subtraction.

perimental verification of the effect. In Sec. 5 we will sum up our findings.

2. CONDITIONS FOR EMERGENCE OF A NONABSORBING STATE

We will examine a double Λ system (Fig. 1a) that resonantly interacts with four electromagnetic fields. The energy levels are labeled by the numbers $j=1$, 2, 3, and 4, and the fields by the letters $\nu=a$, b, c, and d:

$$\mathbf{E}(t)=\frac{1}{2}\sum_{\nu=a,b,c,d}\mathbf{E}_\nu\exp\{-i\omega_\nu t\}+\text{c.c.},\tag{2}$$

where $\mathbf{E}_\nu$ and ω_ν are the amplitudes and frequencies of the fields. At first we ignore relaxation and write the Schrödinger equation for the column vector $\mathbf{a}=(a_1,a_2,a_3,a_4)^T$ (T stands for ''transposed'') of the probability amplitude of the system being in states $|j\rangle$ $(j=1,\ldots,4)$ in the interaction picture:

$$i\dot{\mathbf{a}}=V\mathbf{a}.\tag{3}$$

The interaction matrix elements are nonzero for allowed transitions:

$$V_{13}=G_a\exp\{-i\Omega_a t\},\quad V_{14}=G_b\exp\{-i\Omega_b t\},$$
$$V_{23}=G_c\exp\{-i\Omega_c t\},\quad V_{24}=G_d\exp\{-i\Omega_d t\}.\tag{4}$$

Here the Ω_ν are the detunings of the field with respect to the corresponding transitions $\omega_{ij}\equiv(E_i-E_j)/\hbar,\ i,j=1,\ldots,4$:

$$\Omega_a=\omega_a-\omega_{31},\quad \Omega_b=\omega_b-\omega_{41},$$
$$\Omega_c=\omega_c-\omega_{32},\quad \Omega_d=\omega_d-\omega_{42},\tag{5}$$

and $G_\nu=-\mathbf{E}_\nu\cdot\mathbf{d}_\nu/2\hbar$ are the Rabi frequencies, i.e., the amplitudes of the matrix elements of the interaction of the fields $\mathbf{E}_\nu$ with the dipole moments of the transitions, $\mathbf{d}_\nu$.

Now we will proceed with the search for a nonabsorbing state. We transform the Schrödinger equation (3) into a system of equations with constant coefficients by transforming the Hermitian matrix $V=V^\dagger$ into a time-independent matrix via a unitary transformation U:

$$\tilde{\mathbf{a}}=U\mathbf{a},\quad \tilde{V}=UVU^\dagger-iUU_t^\dagger.\tag{6}$$

Such a ''stopping'' transformation can be selected if we look for its matrix in diagonal form:

$$U=\exp\{-iHt\},\quad H=\text{diag}(\Delta_1,\Delta_2,\Delta_3,\Delta_4).\tag{7}$$

For the undetermined coefficients Δ_i we have a degenerate system of homogeneous linear equations, which has a solution if

$$\Omega_a-\Omega_b-\Omega_c+\Omega_d=0.\tag{8}$$

This condition is a corollary of the cycle (1) in the frequencies of the fields. One parameter can be specified arbitrarily, say $\Delta_4=0$, and the other parameters are uniquely defined: $\Delta_1=\Omega_c$, $\Delta_2=\Omega_d$, and $\Delta_3=\Omega_c-\Omega_a$. Reasoning in a similar manner, we use the unitary transformation

$$U_1=\exp\{-i\Phi\},\quad \Phi=\text{diag}(\phi_1,\phi_2,\phi_3,\phi_4),$$

to transform the Rabi frequencies G_a, G_b, G_c, and G_d into real positive parameters when the synchronism condition is met, or $\varphi_a-\varphi_b-\varphi_c+\varphi_d=0$, with the φ_ν the phases of the fields. The synchronism condition is met if one of the fields,for example c, is generated as a result of four-wave mixing. But if there is no phase synchronism, only three amplitudes can be made real. We write the resulting matrix $\tilde{V}$ explicitly. For conciseness we denote a Rabi frequency by its index (i.e., we drop the letter G), or $G_\nu\to\nu$, with $\nu=a$, b, c, and d:

$$\tilde{V}=\begin{pmatrix}\Omega_c & 0 & a & c\\ 0 & \Omega_d & b & d\\ a & b & \Omega_c-\Omega_a & 0\\ c & d & 0 & 0\end{pmatrix}.\tag{9}$$

In a real four-level system there is relaxation. If the upper level $|4\rangle$ decays much faster than the states $|1\rangle$, $|2\rangle$, and $|3\rangle$ $(\Gamma_{1,2,3}\ll\Gamma_4)$, the problem of mixing these states by a field into a coherent superposition can be studied without allowing for relaxation when $\Gamma_{1,2,3}\ll G_\nu$. In particular, under a certain restriction imposed on the parameters such a superposition of long-lived states may split off and not interact with the field. We call such a superposition a split-off state.

Below we derive a criterion for the emergence of such a split-off state. Note that we began with the amplitude equations (3) without relaxation. If we account for spontaneous decay, the population is transferred to the split-off state through level $|4\rangle$. If this split-off state is long-lived, the fields cease to be absorbed. Such a nonabsorbing state of a four-level system manifests itself as a narrow dip in the frequency dependence of the population of level $|4\rangle$. A similar narrow resonance appears in the frequency dependence of the fluorescence signal from level $|4\rangle$.

We begin with the thoroughly studied simple case of a three-level Λ scheme, which we will attempt to generalize to a four-level scheme. If $a=b=0$, the diagrams in Figs. 1a and 1b reduce to the three-level diagram $|1\rangle$–$|4\rangle$–$|2\rangle$. To find the condition for emergence of a split-off state, we seek the eigenvector (a_1,a_2,a_4) of the operator

$$V_3=\begin{pmatrix} \Omega_c & 0 & c \\ 0 & \Omega_d & d \\ c & d & 0 \end{pmatrix}, \tag{10}$$

in which $a_4=0$ and a_1, $a_2\neq 0$. For the nonzero components we obtain an overdetermined system of linear equations:

$$\Omega_c a_1=\lambda a_1, \quad \Omega_d a_2=\lambda a_2, \quad c a_1+d a_2=0, \tag{11}$$

where λ is the eigenvalue. The system (11) has a solution if $\lambda=\Omega_c=\Omega_d$, i.e., only for equal detunings. The normalized eigenvector of the split-off state has the form

$$|\mathbf{a}_\lambda\rangle=-\frac{d}{\sqrt{c^2+d^2}}|1\rangle+\frac{c}{\sqrt{c^2+d^2}}|2\rangle. \tag{12}$$

Such a state is known from the theory of coherent population trapping in a three-level system.[10]

Now we will derive the conditions for split-off in a four-level system with fields $a,b\neq 0$. Bearing in mind the direction of the spontaneous decay of level $|4\rangle$, we must find the split-off coherent superposition of the states $|1\rangle$, $|2\rangle$, and $|3\rangle$. To do this, we must solve the eigenvalue problem

$$\tilde{V}\mathbf{a}_\lambda=\lambda\mathbf{a}_\lambda \tag{13}$$

and find an eigenvector $\mathbf{a}_\lambda=(a_1,a_2,a_3,0)^T$ in which the fourth component is zero. For the first three nonzero components, a_1, a_2, and a_3, we arrive at an overdetermined system of linear equations, a system that has a solution if there is a common root λ for the two equations

$$b(ad-bc)+c(\Omega_c-\Omega_a-\lambda)(\Omega_d-\lambda)=0,$$

$$bd(\Omega_c-\lambda)+ac(\Omega_d-\lambda)=0. \tag{14}$$

If we combine this with (8), we arrive at the condition

$$\Omega_a-\Omega_b=\frac{ac+bd}{bd}\left[-\frac{\Omega_b}{2}\pm\sqrt{\frac{\Omega_b^2}{4}+\frac{b}{c}(bc-ad)}\right], \tag{15}$$

which links two detunings and four Rabi frequencies.

The curve representing the Ω_a vs. Ω_b dependence (Fig. 2) is a hyperbola with its vertex at the origin and two asymptotes, $\Omega_a=\Omega_b$ and $\Omega_a=-\Omega_b ac/bd$. The first asymptote corresponds to a criterion for coherent population trapping in the three-level Λ scheme, while the second emerges in the four-level scheme. If $bc<ad$, a vertical band $\Omega_b^2\leqslant 4b(ad-bc)/c$, known as the forbidden band, appears in the (Ω_a,Ω_b) plane, and inside this band there can be no coherent population trapping, no matter what the value of Ω_a is. When bc is exactly equal to ad, the hyperbola degenerates into a pair of straight lines intersecting at the origin.

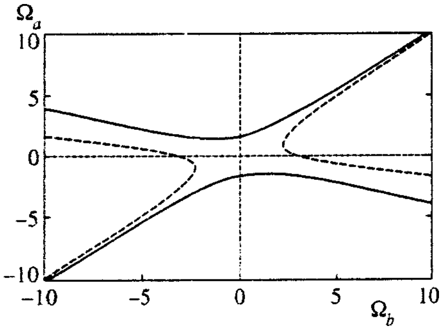

FIG. 2. Criterion for coherent population trapping in the (Ω_a,Ω_b) plane. The solid curves correspond to the case $bc>ad$ ($a=1$, $b=2$, $c=3$, and $d=4$); the dashed curves represent a resonant hyperbola in the opposite case $bc<ad$ (a, b, and c are the same and $d=8$).

The two branches of the hyperbola can be interpreted as a manifestation of the splitting of a nonabsorbing state in a four-level system into two components. To find the corresponding wave functions, we must solve the spectral problem (13). The eigenvalue can be found from (14),

$$\lambda=\frac{ac\Omega_d+bd\Omega_c}{ac+bd},$$

and the unnormalized eigenvector is

$$\mathbf{a}_\lambda=\left(-bd,\, bc,\, c\left[-\frac{\Omega_b}{2}\pm\sqrt{\frac{\Omega_b^2}{4}+\frac{b}{c}(bc-ad)}\right], 0\right)^T. \tag{16}$$

In a special case we can also find the nonabsorbing state in which only two states are mixed, $|1\rangle$ and $|2\rangle$ (as in a three-level system) rather than three. For such mixing to occur, additional conditions linking the field amplitudes and the detunings must be imposed. To obtain these conditions, we seek the eigenvector of the matrix $\tilde{V}$ in the subspace of states orthogonal to the vectors $|3\rangle$ and $|4\rangle$ $(a_3=a_4=0)$. The solvability condition amounts to $\Omega_c=\Omega_d$ (which also means that $\Omega_a=\Omega_b$) and an equation that links the amplitudes:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}=0. \tag{17}$$

In the given special case, the hyperbola degenerates into a pair of straight lines intersecting at the origin. Here, as Eq. (16) implies, a_3 vanishes, too. The normalized split-off state $\mathbf{a}_\lambda$ is the same as in the three-level system [Eq. (12)].

We sought a linear transformation that would leave the state $|4\rangle$ in place to account for the direction of relaxation processes. Coherent population trapping in which states $|1\rangle$, $|2\rangle$, and $|3\rangle$ are mixed can occur if the relaxation constants of these levels, $\Gamma_{1,2,3}$, are small compared to Γ_4. The special case (17) can be realized if the relaxation constants of levels $|1\rangle$ and $|2\rangle$ are relatively small.

In the symmetric case $a=c$, $b=d$, $\Omega_a=\Omega_c$, and $\Omega_b=\Omega_d$, the interaction operator (9) becomes invariant under the permutation $|3\rangle\leftrightarrow|4\rangle$. Then $\mathbf{a}_\lambda=(0,0,1,-1)^T$ becomes an eigenvector and the difference of amplitudes, $a_3(t)-a_4(t)$, a constant of the motion. In the subspace orthogonal to the vector $\mathbf{a}_\lambda$ the system reduces to a three-level scheme. At $a=b$, $c=d$, $\Omega_a=\Omega_b$, and $\Omega_c=\Omega_d$, the vector $(1,-1,0,0)^T$ becomes an eigenvector and the system becomes invariant under the permutation $|1\rangle\leftrightarrow|2\rangle$. The symmetric linear combination (12) of these two states is the split-off state.

Reasoning along similar lines, we arrive at an interaction matrix for the rhombic scheme (Fig. 1b). The matrix differs from (9) by the signs of two detunings: $\Omega_c\to-\Omega_c$ and $\Omega_d\to-\Omega_d$. Using the synchronism condition, which differs from (8) in the signs of Ω_c and Ω_d,

$$\Omega_a-\Omega_b+\Omega_c-\Omega_d=0, \tag{18}$$

we arrive at the same criterion (15) for the detunings.

3. DARK RESONANCE IN THE NONLINEAR SUSCEPTIBILITY

In this section we will use the quantum kinetic equation for the density matrix in the relaxation-constant model to allow for the decay of states. For an open four-level system there are fourteen such parameters: four population relaxation constants Γ_j $(j=1,2,3,4)$, four relaxation constants Γ_ν $(\nu=a,b,c,d)$ for the polarizations of the allowed transitions, two coherence decay constants $\Gamma_{12}\equiv\Gamma_e$ and $\Gamma_{34}\equiv\Gamma_f$ for the forbidden transitions, and four Einstein coefficients A_ν.

As in Sec. 2, the Hamiltonian in the dynamical equation $\dot{\sigma}=-i[V,\sigma]$ for the density matrix σ can be made time-independent via the unitary transformation (7):

$$\rho=\exp\{iHt\}\,\sigma\exp\{-iHt\},$$
$$\tilde{V}=\exp\{-iHt\}\,V\exp\{iHt\}+H,$$

and matrix (9) can also be obtained.

When relaxation is included in the picture, we can write the kinetic equation as a system of equations

$$\dot{R}=-i\hat{L}R+Q, \tag{19}$$

where Q is the column of incoherent excitation rates ($Q_{5j-4}=\Gamma_jN_j$, while the other components of Q are zeros), N_j are the steady-state values of the populations of the levels unperturbed by the field $(N_1+N_2+N_3+N_4=1)$, R is the column of the elements of the density matrix ordered in rows,

$$R=(\rho_{11},\rho_{12},\rho_{13},\rho_{14},\rho_{21},\rho_{22},\rho_{23},\rho_{24},\rho_{31},$$
$$\times\rho_{32},\rho_{33},\rho_{34},\rho_{41},\rho_{42},\rho_{43},\rho_{44})^T,$$

and $\hat{L}$, known as the superoperator,[13] is the 16×16 matrix

$$\hat{L}=\begin{pmatrix}
\Gamma_1 & 0 & a & c & 0 & 0 & 0 & 0 & -a & 0 & -A_a & 0 & -c & 0 & 0 & -A_c\\
0 & \hat{\Gamma}_{12} & b & d & 0 & 0 & 0 & 0 & 0 & -a & 0 & 0 & 0 & -c & 0 & 0\\
a & b & \hat{\Gamma}_{13} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -a & 0 & 0 & 0 & -c & 0\\
c & d & 0 & \hat{\Gamma}_{14} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -a & 0 & 0 & 0 & -c\\
0 & 0 & 0 & 0 & \hat{\Gamma}_{21} & 0 & a & c & -b & 0 & 0 & 0 & -d & 0 & 0 & 0\\
0 & 0 & 0 & 0 & 0 & \Gamma_2 & b & d & 0 & -b & -A_b & 0 & 0 & -d & 0 & -A_d\\
0 & 0 & 0 & 0 & a & b & \hat{\Gamma}_{23} & 0 & 0 & 0 & -b & 0 & 0 & 0 & -d & 0\\
0 & 0 & 0 & 0 & c & d & 0 & \hat{\Gamma}_{24} & 0 & 0 & 0 & -b & 0 & 0 & 0 & -d\\
-a & 0 & 0 & 0 & -b & 0 & 0 & 0 & \hat{\Gamma}_{31} & 0 & a & c & 0 & 0 & 0 & 0\\
0 & -a & 0 & 0 & 0 & -b & 0 & 0 & 0 & \hat{\Gamma}_{32} & b & d & 0 & 0 & 0 & 0\\
0 & 0 & -a & 0 & 0 & 0 & -b & 0 & a & b & \Gamma_3 & 0 & 0 & 0 & 0 & 0\\
0 & 0 & 0 & -a & 0 & 0 & 0 & -b & c & d & 0 & \hat{\Gamma}_{34} & 0 & 0 & 0 & 0\\
-c & 0 & 0 & 0 & -d & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \hat{\Gamma}_{41} & 0 & a & c\\
0 & -c & 0 & 0 & 0 & -d & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \hat{\Gamma}_{42} & b & d\\
0 & 0 & -c & 0 & 0 & 0 & -d & 0 & 0 & 0 & 0 & 0 & a & b & \hat{\Gamma}_{43} & 0\\
0 & 0 & 0 & -c & 0 & 0 & 0 & -d & 0 & 0 & 0 & 0 & c & d & 0 & \Gamma_4
\end{pmatrix}, \tag{20}$$

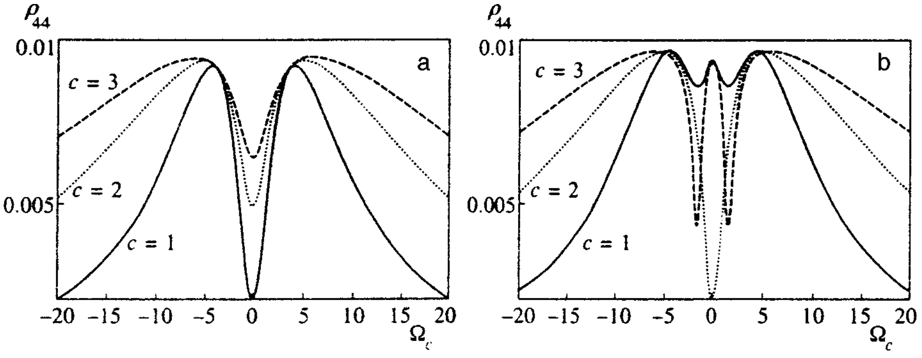

FIG. 3. Population ρ_{44} as a function of the detuning Ω_c for different values of c at $b=0$ (a), $b=2$ (b), $a=1$, $d=4$, $\Gamma_1=\Gamma_2=\Gamma_3=0.01$, and $\Gamma_4=1$. The off-diagonal relaxation constants are equal to the half-sums of the corresponding diagonal relaxation constants: $\Gamma_{ij}=(\Gamma_i+\Gamma_j)/2$; $N_1=1$, $N_2=N_3=N_4=0$, $\Omega_b=\Omega_d=0$, and $\Omega_c=\Omega_a$.

with

$$\hat{\Gamma}_{12}=i\Gamma_e-(\Omega_d-\Omega_c),\quad \hat{\Gamma}_{13}=i\Gamma_a+\Omega_a,\quad \hat{\Gamma}_{14}=i\Gamma_c+\Omega_c,$$

$$\hat{\Gamma}_{21}=i\Gamma_e+(\Omega_d-\Omega_c),\quad \hat{\Gamma}_{23}=i\Gamma_b+\Omega_b,\quad \hat{\Gamma}_{24}=i\Gamma_d+\Omega_d,$$

$$\hat{\Gamma}_{31}=i\Gamma_a-\Omega_a,\quad \hat{\Gamma}_{32}=i\Gamma_b-\Omega_b,\quad \hat{\Gamma}_{34}=i\Gamma_f+(\Omega_c-\Omega_a),$$

$$\hat{\Gamma}_{41}=i\Gamma_c-\Omega_c,\quad \hat{\Gamma}_{42}=i\Gamma_d-\Omega_d,\quad \hat{\Gamma}_{43}=i\Gamma_f-(\Omega_c-\Omega_a).$$

The matrix $\hat{L}$ becomes symmetric only if there is no spontaneous decay via the allowed transitions, $A_\nu=0$. When there is no relaxation, $\Gamma_j=\Gamma_\nu=0$, $j=1,\dots,4$ and $\nu=a,b,c,d,e,f$, the matrix becomes traceless.

Below we give the steady-state solutions of Eqs. (19) obtained by numerically inverting the matrix (20) by Gauss's method. Figure 3 depicts the dependence of the population of level $|4\rangle$ on the frequency detuning $\Omega_c=\Omega_a$ at $\Omega_d=\Omega_b=0$. To land in the range of parameters that ensures coherent population trapping, the relaxation constants of levels $|1\rangle$, $|2\rangle$, and $|3\rangle$ were chosen so as to be one hundred times smaller than the relaxation constant of $|4\rangle$. Figure 3a depicts the case where $b=0$, with the result that the field cycle is broken and the system reduces to a three-level scheme. The reader can clearly see a narrow dip near zero detuning, $\Omega_d=0$, corresponding to the condition for coherent population trapping. As the amplitude of field c increases, the dip becomes shallower due to the smoothing effect of saturation. Figure 3b corresponds to the case where $b\neq 0$, with the result that a "dark" resonance appears at $c=2$, when $bc=ad$. When $bc>ad$ $(c=3)$, the dip in the population of the upper level (and hence in the fluorescence signal) splits into two narrow components.

The absolute value of the nonlinear susceptibility $\beta \propto |\rho_{14}|$ at the frequency ω_c was also calculated. In the simplest model, which ignores the depletion of the pump fields a, b, and d, the propagation of field c is described by a truncated Maxwell equation with polarization ρ_{14} on the right-hand side. Hence for an optically thin medium we have $c^2\propto|\rho_{14}|^2L^2$, where L is the range of the medium. The dependence of the conversion coefficient β (in arbitrary units) on the detuning is depicted in Fig. 4a for three values of the amplitude $c=1,2,3$. The parameters are chosen in such a way that the curve for $c=1$ corresponds to $bc<ad$ and no coherent population trapping at $\Omega_a=0$ is present on it. Equation (17) is valid on the curve for $c=2$, with the result that coherent population trapping appears at the center of the line, $\Omega_a=0$. The opposite inequality $bc>ad$ holds for the curve with $c=3$; hence according to Fig. 2 this curve reflects the presence of two resonances, in accordance with two signs in (15). We see that the resonance, split into two broad symmetric humps, also acquires a wide central component, and a narrow dip grows with increasing field c at the center of the component. Then, as the field gets stronger, the dip splits into two components. The narrow dips in the curve for $c=3$ are positioned at approximately the same frequencies as the abrupt dips in the population of state $|4\rangle$ in Fig. 3b. At $\Omega_b\neq 0$ the "dark" resonance shifts. An example of an asymmetric frequency dependence is illustrated by Fig. 4b. When $\Omega_b\neq 0$, the radicand in (15) changes sign at a smaller value of c. Hence in Fig. 4b the splitting of the nonabsorbing state is clearly visible already in the curve for $c=2$.

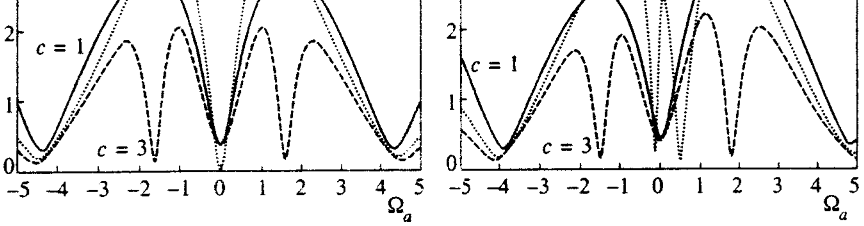

FIG. 4. Nonlinear susceptibility (in arbitrary units) as a function of the detuning Ω_a at $\Gamma_1=\Gamma_2=\Gamma_3=0.01$, $\Gamma_4=1$, $\Gamma_{ij}=(\Gamma_i+\Gamma_j)/2$, $A_\nu=0$, $a=1$, $b=2$, $d=4$, $c=1,2,3$, $\Omega_d=0$, $\Omega_c=\Omega_a-\Omega_b$, $N_1=1$, and $N_2=N_3=N_4=0$; $\Omega_b=0$ (a) and $\Omega_b=0.5$ (b).

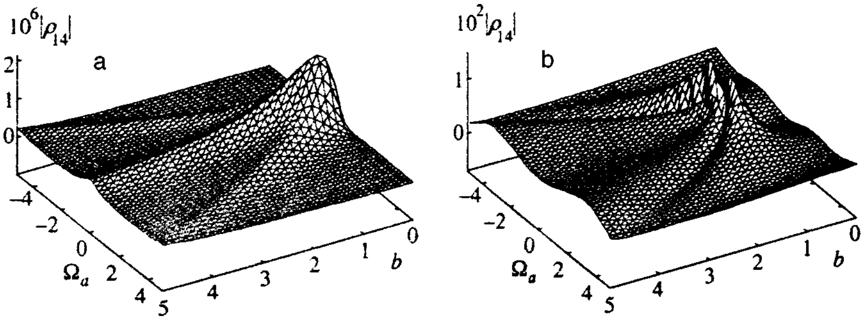

FIG. 5. Absolute value of polarization $\propto|\rho_{14}|$ as a function of the detuning Ω_a and the field b: (a) $a=d=0.001$, $c=0$, and $\Gamma_1=\Gamma_2=\Gamma_3=\Gamma_4=0.5$; and (b) $a=1$, $c=3$, $d=4$, $\Gamma_1=\Gamma_2=0.01$, $\Gamma_3=0.5$, and $\Gamma_4=3$. Everywhere $A_\nu=0$, $N_1=1$, and $N_2=N_3=N_4=0$.

4. DISCUSSION

Figure 5 depicts the quantity $\propto|\rho_{14}|$ as a function of the detuning of the field a and the Rabi frequency b if initially only level $|1\rangle$ was populated. In the case of weak fields a, c, and d, Fig. 5a shows the splitting of the resonance into three peaks: one peak is not shifted, and the two peaks that are shifted in relation to $\Omega_a=0$ represent the components of the Autler–Townes doublet, as they do in Ref. 12. The picture changes dramatically as the fields become strong. Figure 5b shows that when the field strength exceeds a certain critical value, the frequency dependence acquires a ''dark'' resonance, which splits into two components. The distance between the dips grows as field b becomes stronger. As b increases, the central component (which is split into two components) transforms into a triplet and then into a quartet. Without the side components (not shown in the figure), there are altogether six components in the nonlinear susceptibility spectrum.

Lin *et al.*[14] studied the degenerate four-wave mixing of the light of a titanium–sapphire laser at the D_2 line of ^{85}Rb. The level configuration corresponded to a two-level system, so that the spectrum of the probe field exhibited a doublet or triplet. In our setting a two-level system corresponds to the case of three weak (or far from resonance) fields, say a, c, and d and one strong field, b (see Fig. 5b). A two-level system does not exhibit coherent population trapping—there is only the splitting of the resonance into three components corresponding to the transitions between the different quasienergy levels.

To achieve frequency conversion in Pb vapor, Jain *et al.*[2] used the mixing of the second and third harmonics of the light of a titanium–sapphire laser ($\lambda_b=406\,\mathrm{nm}$, $\lambda_d=283\,\mathrm{nm}$, $\lambda_a=293\,\mathrm{nm}$, and $\lambda_c=425\,\mathrm{nm}$). However, in the double Λ scheme they selected for the intermediate state $|3\rangle$ (Fig. 1a) a virtual level detuned by $1112\,\mathrm{cm}^{-1}$ from $|4\rangle$. Hence the system was three-level and no splitting of coherent population trapping was observed either in the experiment or in computer simulations.

In their numerical calculations, Petch *et al.*[7] studied the mixing process in the rhombic krypton scheme. For the state $|2\rangle$ (Fig. 1b) they took a virtual level lying between $|3\rangle$ and $|4\rangle$ and corresponding to the two-photon transition $|3\rangle\to|4\rangle$. The system becomes three-level, the superoperator is a 9×9 matrix, and the nonabsorbing state does not split. The frequency dependence of the nonlinear susceptibility in Ref. 7 consists of two components.

In studying the linear regime of amplification of a bichromatic field without inversion in a double Λ scheme, Kocharovskaya and Mandel[11] examined, for the sake of simplicity, a special symmetric case. In the notation adopted in the present paper this case can be written as $\Omega_\nu=0$, $a=b$, and $c=d$. The truncated Maxwell equations retain this property even if we allow for depletion of pumping in the medium. Then $bd=ac$, i.e., condition (17) is met and the same nonabsorbing state as in the three-level system [Eq. (12)] is retained.

In their experiment, Babin *et al.*[5] fed the fields d, c, and b ($\lambda_d=488\,\mathrm{nm}$, $\lambda_c=599\,\mathrm{nm}$, and $\lambda_b=655\,\mathrm{nm}$) to an external cell from an argon and dimer Raman laser and a dye laser, respectively. The field a was generated in the process of resonant four-wave mixing, The fields a and c turned out to be weak, and the detunings Ω_d and Ω_c could not be varied independently, since they were rigidly coupled by the condition for Raman generation. Hence no splitting of coherent population trapping was observed in the experiment. Neither was there any splitting in the theory of a double Λ scheme with two strong fields in the opposite transitions,[15] while the appearance of two additional peaks in the nonlinear susceptibility spectrum (i.e., in addition to the four ordinary peaks) was a consequence of averaging over velocities.

If we want to observe the effects of coherent population trapping in four-wave mixing, we must make all four fields strong, $|G_\nu|\gg|\Gamma_\nu-i\Omega_\nu|$, and the three lower levels sufficiently narrow compared to the upper level. The effect is most appreciable at $ad=bc$ in exact resonance $\Omega_a=\Omega_b=\Omega_c=\Omega_d=0$. In this case, according to (17), there is no splitting, with the result that the dip in the dependence of the nonlinear susceptibility on the amplitude of wave c remains abrupt even if the relaxation constant of level $|3\rangle$ is large (Fig. 6).

The highest intensity of the field generated in the process of continuous resonant four-wave mixing with upconversion was attained by Hinze *et al.*[16] In exciting molecular sodium vapor by laser light of wavelengths $\lambda_a=665\,\mathrm{nm}$, $\lambda_b=756\,\mathrm{nm}$, and $\lambda_d=532\,\mathrm{nm}$ and input powers of about 100 mW, the output power of the generated light at $\lambda_c=480\,\mathrm{nm}$ was found to be about 0.1 mW, which is too low to observe the effect. However, the power of the generated wave can be raised by stabilizing the frequency of the exciting radiation. Much stronger fields can be generated in the pulsed mode. Recently, Dorman and Marangos[8] observed a five-fold pulsed increase in the efficiency of four-wave mix-

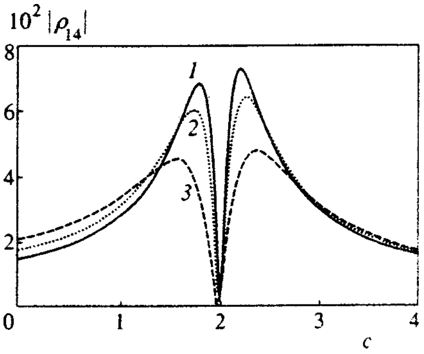

FIG. 6. Nonlinear susceptibility (in arbitrary units) as a function of the Rabi frequency c at $a=1$, $b=2$, $d=4$, $\Gamma_1=\Gamma_2=0.1$, and $\Gamma_4=1$. Curve *1* corresponds to $\Gamma_3=0.1$, curve *2* to $\Gamma_3=0.3$, and curve *3* to $\Gamma_3=1$ ($A_\nu=0$ and $\Omega_\nu=0$).

ing due to electromagnetically induced transparency in krypton.

5. CONCLUSION

In this paper we have focused on a new coherence effect: the splitting of a nonabsorbing state in a four-level system. The aim of the work was to demonstrate that the effect strongly influences the dependence of populations and nonlinear susceptibility of the frequencies and amplitudes of the waves. It was found (and corroborated by numerical calculations) that under certain conditions deep narrow dips form in the dependence. Experimental verification of the effect requires that all four fields be strong.

The author is grateful to S. A. Babin, A. I. Chernykh, E. V. Podivilov, and M. G. Stepanov for useful discussions. The work was supported by the Russian Fund for Fundamental Research (Grants Nos. 96-02-00069 and 96-15-96642) and the Ministry of Science and Technology programs in optics (laser physics, 1.53) and fundamental spectroscopy (08.02.32).

*)E-mail: shapiro@iae.nsk.su

---

[1] A. B. Budnitskiĭ and A. K. Popov, Opt. Spektrosk. **29**, 1032 (1970).

[2] M. Jain, H. Xia, G. Y. Yin, A. J. Merriam, and S. E. Harris, Phys. Rev. Lett. **77**, 4326 (1996).

[3] J. P. Marangos, N. Shen, H. Ma, M. H. R. Hutchinson, and J. P. Connerade, J. Opt. Soc. Am. B **7**, 1254 (1990).

[4] A. Apolonsky, S. Baluschev, U. Hinze, E. Tiemann, and B. Wellegehausen, Appl. Phys. B: Lasers Opt. **64**, 435 (1997).

[5] S. Babin, U. Hinze, E. Tiemann, and B. Wellegehausen, Opt. Lett. **21**, 1186 (1996).

[6] S. E. Harris, J. F. Field, and A. Imamoğlu, Phys. Rev. Lett. **64**, 1107 (1990).

[7] J. C. Petch, C. H. Keitel, P. L. Knight, and J. P. Marangos, Phys. Rev. A **53**, 543 (1996).

[8] C. Dorman and J. P. Marangos, Phys. Rev. A **58**, 4121 (1998).

[9] E. Arimondo, in *Progress in Optics*, Vol. 35, Elsevier, Amsterdam (1996), p. 257.

[10] B. D. Agap'ev, M. B. Gornyĭ, B. G. Matisov, and Yu. V. Rozhdestvenskiĭ, Usp. Fiz. Nauk **163**(9), 1 (1993) [Phys. Usp. **36**, 763, 1086 (1993)].

[11] O. Kocharovskaya and P. Mandel, Phys. Rev. A **42**, 523 (1990).

[12] D. A. Coppeta, P. L. Kelley, P. J. Harshman, and T. K. Gustavson, Phys. Rev. A **53**, 925 (1996).

[13] N. Chencinski, W. M. Schreiber, A. M. Levine, and Y. Prior, Phys. Rev. A **42**, 2939 (1990).

[14] J. Lin, A. I. Rubiera, and Y. Zhu, Phys. Rev. A **52**, 4882 (1995).

[15] S. A. Babin, E. V. Podivilov, and D. A. Shapiro, JETP Lett. **66**, 816 (1997).

[16] U. Hinze, L. Meyer, A. Apolonskii, S. Babin. E. Tiemann, and B. Wellegehausen, in *Proc. European Quantum Electronic Conference* (Glasgow, September 14–18, 1998) Technical Digest, QTuI3.

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Subthreshold resonances in three-particle molecular systems

F. M. Pen'kov*)

Joint Institute for Nuclear Research, 141980 Dubna, Moscow Region, Russia

(Submitted 20 October 1998)
Zh. Éksp. Teor. Fiz. **115**, 1973–1986 (June 1999)

This paper examines the resonant scattering of a light particle by a pair of identical particles in the Efimov limit. An analytic expression for the resonance widths is derived. The results of calculations are compared with the solution of Faddeev integral equations within a broad range of masses of the light particle. It is shown that the widths of the subthreshold resonances in the scattering amplitude obtained from the integral equations with Yamaguchi potentials are accurately described by the analytic expression, which makes it possible to use this expression in a range of masses inaccessible to numerical calculations. The conclusion is drawn that the lifetime of highly excited negative molecular ions is infinite. © *1999 American Institute of Physics.* [S1063-7761(99)00506-5]

1. INTRODUCTION

Recent publications[1–3] examined exotic states of systems consisting of neutral atoms and an electron. These states have large dimensions (up to several tens of angstroms) and are a reflection of the effects of three-body dynamics at low pairwise binding energies. For example, by directly analyzing the Faddeev equations[1] the effective interaction potential of two neutral atoms in the presence of an electron was found. This potential is local over the range[1)]

$$\max(r_0)\ll r\ll \min(\kappa^{-1}) \tag{1}$$

(r_0 is the range of the pairwise forces, and κ is the wave number of the bound or virtual state of the pair) and contains not only long-range components of the $1/r^2$ type (which is characteristic of the Efimov effect[4]) but also terms of the quasi-Coulomb form (of the $1/r$ type), which provide the main contribution to the spectrum of negative diatomic molecular ions with a moderate binding energy of the electron and atom (the binding may be real or virtual). In particular, the experimental data on the scattering of an electron by a helium atom suggest that even in the absence of an atom–atom interaction, a bound state of the He_2^- system can exist.[2] Furthermore, a study of a system consisting of three neutral atoms of alkali metals and an electron[3] has shown that the effective interaction generated by the electron ensures the existence of more than 1000 bound states. It was assumed that such system can serve as a starting point for the formation of clusters in highly rarefied gases. However, the analysis proved to be incomplete in the presence of an atom–atom interaction with a binding energy higher than the binding energy of the three-particle system. In this case the system can disintegrate into a molecule and a free electron, with the bound states of the three-particle complex transforming into resonances with lifetimes determined by the resonance widths.

The present work studies these widths within the scope of three-body scattering of a light particle by a bound pair of heavy particles. Here, as in Refs. 1 and 2, we examine systems in which the pairwise scattering length is much greater than the ranges of the pairwise forces, a condition corresponding to that for the Efimov effect:

$$\kappa r_0 \ll 1. \tag{2}$$

The binding energy of the atom is assumed finite, while the binding energy in the electron–atom pairs is assumed to vanish, so as to simplify analytic calculations. Thus, the transformation threshold for the system coincides with the breakup threshold, and the three-particle spectrum[1,2] mentioned earlier is a spectrum of subthreshold Efimov resonances. Such resonances (but only below the excitation threshold) were examined in Ref. 5 for a three-boson system. There it was noted that both the positions of the resonances and their widths are equidistant on a logarithmic scale.

This paper examines the case in which the mass of one of the particles is much less than the masses of the other two. This means that we a dealing with almost classical motion of the heavy particles,[1] which technically makes an accurate numerical solution of the Faddeev equations for the electron-to-atom mass ratio impossible. Hence, as in Ref. 1, in the present work the numerical calculations were done for mass ratios not exceeding 1/100 in order to verify the validity of the analytic relationship linking the widths and positions of resonances, a relationship that can be derived for very narrow ranges of the pairwise forces. We show that the analytic expression provides a satisfactory description of the extremely complicated mass dependence of the resonance widths, which makes it possible to use this expression in estimating resonance widths for actual electron-to-atom mass ratios.

2. INTEGRAL EQUATIONS

We examine a system consisting of three spinless particles with masses m_i, where the subscript $i=1,2,3$ numbers the particles, and pairwise potentials v_i, where the subscript indicates the particle absent from the pair system. We assume that the identical particles 2 and 3 with masses

$m_2=m_3=m$ interact via the potential v_1 and form a bound state with an energy $\varepsilon_1=-\kappa_1^2/2m_{23}$, where m_{ij} is the reduced mass of particles i and j. Particle 1 with mass m_1 interacts with particles 2 and 3 via the potentials v_3 and v_2, respectively ($v_2=v_3=v$), which generate pair states (real or virtual) with very low binding energies ($\kappa_2=\kappa_3=\kappa\to 0$). For the potentials v_i (between particles j and k) we take the separable Yamaguchi potential acting only in the S-wave:

$$v_i(p,p')=-\frac{4\pi}{m_{jk}}\frac{\beta_i(\beta_i+\kappa_i)^2}{(\beta_i^2+p^2)(\beta_i^2+p'^2)}. \tag{3}$$

The sign of the wave number κ determines whether the pair forms a bound state ($\kappa>0$) or a virtual state. The parameter β_i determines the ranges of the forces in pair i. In particular, when $\beta_i \gg \kappa_i$, the expansion of the effective range in the pair i in powers of the relative-motion momentum p ($p\cot\delta_i=\kappa_i+r_{\text{eff}}p^2/2+\cdots$) yields $r_{\text{eff}}=3/\beta_i$, which makes it possible to use condition (2) in the form $\kappa_i \ll \beta_i$. For brevity, we write the separable potential in the form of projection operators, $v_i=|\nu_i\rangle\langle\nu_i|$.

We examine the scattering of particle 1 by the bound pair of identical particles 2 and 3 (2,3) with an energy $E_k=k_0^2/2m_{1,23}$ of relative motion lower than the system's transformation energy $\varepsilon_2-\varepsilon_1$. Here $m_{i,jk}$ is the reduced mass of particle i and pair (j,k). The momentum $\mathbf{k}_i$ will always be the momentum of relative motion of particle i and pair (j,k), and the momentum $\mathbf{p}_i$ will refer to motion within the pair (j,k). The subscripts will be dropped if this does not cause a mix-up of notation.

To write the system of Faddeev integral equations (see, e.g., Ref. 6) we need an expression for the product of the pairwise t-matrix in the three-particle space and the free Green's function $G^0(Z)=(Z-h_i^0-h_{(i)}^0)^{-1}$. The Hamiltonian of the free motion of three particles is represented by two terms: the Hamiltonian h_i^0 of free motion of particle i and pair (j,k) and the Hamiltonian $h_{(i)}^0$ of relative motion within the pair. The total energy of the three-particle system is $Z=E_k+\varepsilon_1+i0$. For separable potentials this product can be written in the form[7] $tG^0(Z)=v_i|\varphi_i\rangle g_i^0(Z-\varepsilon_i)\langle\tilde{\varphi}_i|$, which incorporates the wave function $|\varphi_i\rangle$ of the bound state of the pair and the function $\langle\tilde{\varphi}_i(Z)|$ defined in the three-particle space. Here and below we use expressions for the pairwise Green's functions $g_i^0(x)$ and $g_{(i)}^0(x)$ corresponding to the pairwise Hamiltonians h_i^0 and $h_{(i)}^0$, respectively. In this notation, the projection of $\langle\tilde{\varphi}_i|$ on $\langle\mathbf{k}_i|$ has the simple form

$$\langle\tilde{\varphi}_i|=R_i\langle\varphi_i|v_iG^0(Z-k_i^2/2m_{i,jk}),$$

$$R_i=-\frac{1}{2m_{jk}}\frac{(\beta_i+a_i)^2(a_i+\kappa_i)}{(2\beta_i+\kappa_i+a_i)\langle\varphi_i|v_i|\varphi_i\rangle},$$

$$a_i=\sqrt{-2m_{jk}(Z-k_i^2/2m_{i,jk})}\,. \tag{4}$$

Note that the pairwise t-matrix corresponding to a separable potential of the Yamaguchi type has two poles. The near pole is at the point of the bound state of the pair, $p_i=\pm i\kappa_i$ (on the physical sheet for a real bound state and on the unphysical sheet for a virtual state). The distant pole on the unphysical sheet, $p_i=-i(2\beta_i+\kappa_i)$, is responsible for the range of the pairwise force. The zero-range limit is reached when β goes to infinity. Only one pole remains in this limit, and hence it is unimportant which pairwise potential generated the pole—local or nonlocal. We use a separable potential in this paper solely because it simplifies the Faddeev integral equations. The t-matrices remain finite as $\beta\to\infty$.

If we know the pairwise t-matrices, we can write the Faddeev equations for the scattering of particle 1 by the bound pair of identical particles 2 and 3 in the form

$$T^{\text{el}}=V_{12}g_2^0(Z-\varepsilon_2)T^r,$$

$$T^r=2V_{21}+2V_{21}g_1^0(Z-\varepsilon_1)T^{\text{el}}+V_{23}g_3^0(Z-\varepsilon_3)T^r, \tag{5}$$

where $V_{ij}=R_i^{1/2}\langle\varphi_i|v_iG_0(Z)v_j|\varphi_j\rangle R_j^{1/2}$, and the transition matrices T^{el} and T^r are linked to the physical elastic and inelastic scattering amplitudes through the relationships

$$f^{\text{el}}=-\frac{m_{1,23}}{2\pi}T^{\text{el}}(\mathbf{k}_1^{\text{out}},\mathbf{k}_1^{\text{in}}),$$

$$f^r=-\frac{\sqrt{m_{1,23}m_{2,13}}}{2\pi}T^r(\mathbf{k}_2^{\text{out}},\mathbf{k}_1^{\text{in}}),$$

where the incoming and outgoing momenta, $\mathbf{k}^{\text{in}}$ and $\mathbf{k}^{\text{out}}$, are located on the energy surfaces.

The system of complex integral equations (5) can be transformed into a system of real (below the transformation threshold) equations (see, e.g., Ref. 8). For example, a simple substitution done after the partial expansions

$$f^{\text{el}}(k,k_0)=-\frac{m_{1,23}}{2\pi}K^{\text{el}}(k,k_0)(1+ik_0f^{\text{el}}(k_0,k_0)),$$

$$f^r(k,k_0)=-\frac{\sqrt{m_{1,,23}m_{2,13}}}{2\pi}K^r(k,k_0)(1+ik_0f^{\text{el}}(k_0,k_0)) \tag{6}$$

have been performed yields a system of equations for the real functions K^{el} and K^r. The new equations were solved numerically. To distinguish between the real equations containing principal value integrals, we introduce additional notation for the total energy, $z=\text{Re}Z$. In this notation the real integral equations for K^{el} and K^r coincide with Eqs. (5) if T^{el} is replaced by K^{el}, T^r by K^r, and Z by z, respectively:

$$K^{\text{el}}=V_{12}g_2^0(z-\varepsilon_2)K^r,$$

$$K^r=2V_{21}+2V_{21}g_1^0(z-\varepsilon_1)K^{\text{el}}+V_{23}g_3^0(z-\varepsilon_3)K^r. \tag{7}$$

These equations for K^{el} and K^r were solved numerically. Below we use Eqs. (7) to discuss a scheme for generating subthreshold resonances.

3. SUBTHRESHOLD RESONANCES

To see how subthreshold resonances emerge, we transform Eqs. (7) to

$$K^{\text{el}}=V_{12}g_2^0(z-\varepsilon_2)K^r,$$

$$K^r=2V_{21}+Vg_3^0(z-\varepsilon_3)K^r, \tag{8}$$

where the effective potential of the energy-closed (inelastic) channel has the form

$$V=V_{23}+2V_{21}g_1^0(z-\varepsilon_1)V_{12}. \tag{9}$$

The system of equations (8) makes it possible to describe the two-channel scattering of particle 1 by the pair (2,3) in terms of the single-channel interaction of particle 2 (or 3) and the pair of particles 1 and 3 (or 2). To do this, we use the system of equations (8) and express K^{el} in terms of the Green's function of the closed (inelastic) channel, $g_v(x)=(x-h_2^0-V)^{-1}$:

$$K^{\mathrm{el}}=2V_{12}g_v(z-\varepsilon_2)V_{21}, \tag{10}$$

whose spectrum determines the features of the elastic channel. For example, resonances in elastic scattering correspond to the points in the spectrum E_t of the Hamiltonian $h_v=h_2^0+V$ for $E_t>\varepsilon_1-\varepsilon_2$. Notwithstanding the obvious nature of this statement, we present a method for building the S-matrix in the resonant case because of some special features of resonances in the system considered here. To this end we examine the case in which the energy z is close to ε_2+E_t and specify the singular part in the Green's function g_v explicitly:

$$g_v(z-\varepsilon_2)=\frac{|\Psi_t\rangle\langle\Psi_t|}{\omega}+g_R,$$

where $\omega=z-\varepsilon_2-E_t$, Ψ_t is the wave function corresponding to the eigenvalue E_t, and g_R is the regular residual term in the Green's function (this term is usually dropped). Then, using the expression (10) and the relationship (6) between the physical amplitude and K^{el}, we arrive at an expression for the S-matrix ($S=1+2ik_0f^{\mathrm{el}}(k_0,k_0)$):

$$S=\frac{1-iB}{1+iB}\,\frac{\omega+\dfrac{1}{2}\dfrac{\Gamma B}{1+B^2}-i\dfrac{1}{2}\dfrac{\Gamma}{1+B^2}}{\omega+\dfrac{1}{2}\dfrac{\Gamma B}{1+B^2}+i\dfrac{1}{2}\dfrac{\Gamma}{1+B^2}},$$

where the width

$$\Gamma=2\frac{k_0m_{1,23}}{\pi}|\langle k_0|V_{12}|\Psi_t\rangle|^2 \tag{11}$$

generated by the singular part of the Green's function varies due to the presence of the regular part

$$B=\frac{k_0m_{1,23}}{\pi}\langle k_0|V_{12}g_RV_{21}|k_0\rangle.$$

The size of the shift of the resonance from E_t is also determined by this part. It is the value of B that determines background scattering far from resonance, where $S=(1-iB)/(1+iB)$. Introducing the background scattering phase $\delta_f=-\arctan B$, we arrive at an expression for the S-matrix:

$$S=\exp\{2i\delta_f\}\,\frac{\omega-i\dfrac{\Gamma}{4}-\dfrac{\Gamma}{4}(\sin(2\delta_f)+i\cos(2\delta_f))}{\omega+i\dfrac{\Gamma}{4}-\dfrac{\Gamma}{4}(\sin(2\delta_f)-i\cos(2\delta_f))}. \tag{12}$$

Note that Eq. (12) was derived without using any approximation schemes and is simply a convenient way to express the S-matrix. However, by introducing the Breit–Wigner parametrization, i.e., by separating out the resonance energy and resonance width, we impose a constraint on the widths: they must be small so that we can identify a resonance. Here the ''true'' resonance width $\tilde{\Gamma}$ is determined not only by Γ but also by the background scattering phase: $\tilde{\Gamma}=\Gamma\cos^2\delta_f$. This explains why the expression for Γ contains plane waves rather that the wave functions of the scattering state in a background potential. Below we use the expression for the width Γ, which is always greater than or equal to $\tilde{\Gamma}$. The other features of representing the S-matrix in the form (12) are fairly obvious, and their discussion lies outside the scope of the present paper.

4. WAVE FUNCTION OF THE CLOSED CHANNEL

Here we are interested in the series of resonances corresponding to the Efimov effect, a situation where the condition (2) is met and the spectrum of the Hamiltonian h_v becomes denser near zero. In this case the potential V is simplifies substantially, and to study subthreshold resonances we only need to consider the S-wave part of the effective potential at low energies or, more precisely, in the range of momenta $k\gg\sqrt{-2m_{2,13}z}$, where the energy dependence of the solutions of the Schrödinger equation is weak. Under such an additional condition ($\beta\to\infty$ as $z\to0$) the terms in the effective potential (9), which are the ''exchange'' potential $V_{\mathrm{ex}}\equiv V_{23}$ and two ''triangular'' (with internal integration) terms $V_{\mathrm{tr}}\equiv V_{21}g_1^0(z-\varepsilon_1)V_{21}$, can be written

$$V_{\mathrm{ex}}^0(k,k')=-\frac{\pi}{2\sqrt{kk'}\,\lambda_1}\,\frac{1}{\sqrt{m_{3,12}m_{12}}}\ln\frac{k^2+k'^2+2\lambda_1kk'}{k^2+k'^2-2\lambda_1kk'}, \tag{13}$$

$$V_{\mathrm{tr}}^0(k,k')=-\frac{1}{\sqrt{kk'}\,(2\lambda_1)^2}\,\frac{m_{12}}{m_{23}^2}\,\mathrm{PV}\int_0^\infty\frac{dt}{t-a}L(t,k)L(t,k'), \tag{14}$$

$$L(t,k)=\ln\frac{\gamma t^2+k^2+2\lambda_2kt}{\gamma t^2+k^2-2\lambda_2kt},\qquad a=\kappa_1\sqrt{\frac{m_{1,23}}{m_{32}}},$$

$$\lambda_i=\sqrt{\frac{m_{ij}m_{ik}}{m_i^2}},\qquad \gamma=\sqrt{\frac{m_{23}}{m_{12}}}.$$

The upper index 0 in this notation indicates that the effective interaction is taken at $z=0$. The exchange potential V_{ex}^0 corresponds to the scattering of particle 2 (or 3) by the bound pair of particles 1 and 3 (or 2) at zero binding energy in the pair and is the ''classical'' potential generating Efimov states. The solutions in the field of such a potential, studied in Ref. 7 for a three-boson system, coincide with the solution in the field of a local potential of the type $-(\mu^2+0.25)/2m_{2,13}\rho^2$ (ρ is the Jacobi potential of the relative motion of the particle and the complex) with the coupling

TABLE I. Dependence of μ_1, μ_2, and μ_{as} on the masses.

m_2/m_1	10	20	30	50	70	100
μ_1	1.379 051	1.893 909	2.284 713	2.906 777	3.415 826	4.061 110
μ_2	1.468 174	1.919 444	2.293 865	2.908 434	3.416 219	4.061 172
μ_{as}	1.430 016	1.907 906	2.289 978	2.908 096	3.416 437	4.061 489
$\Pi_{\mathrm{tr}}/\Pi_{\mathrm{ex}}$	1.12×10^{-1}	2.20×10^{-2}	6.41×10^{-3}	9.01×10^{-4}	1.81×10^{-4}	2.38×10^{-5}

constant μ satisfying a transcendental equation. Below we derive such an equation for the total interaction potential. Note that the potential V^0_{ex} behaves as $1/k$ and allows for a solution in the form k^s. The potential V^0_{tr} is more complicated and allows for such solutions only in the momentum range where it is possible to ignore a in the denominator of the integrand of (14). To clarify this aspect, we look for the solutions of the Schrödinger equation with zero energy in the form $\Psi_t = k^{i\mu-5/2}$. We define Π as

$$\Pi(k) = -\frac{2m_{2,13}}{k^{i\mu-1/2}}\int V^0(k,t)\,t^{i\mu-5/2}\frac{d^3t}{(2\pi)^3}$$

and attach a subscript that corresponds to the specific potential. The meaning of Π is simple: it is the ratio of the potential energy to the kinetic energy in the Schrödinger equation. Then the Schrödinger equation can be written in the form $\Pi_{\mathrm{ex}}+2\Pi_{\mathrm{tr}}=1$. The contribution of the exchange potential can be expressed in terms of an integral:

$$\Pi_{\mathrm{ex}} = \frac{1+\zeta_1^2}{\zeta_1} I_1,$$

$$I_1 = \frac{1}{2\pi}\int_0^\infty dx\, x^{i\mu-1}\ln\frac{1+2\lambda_1 x+x^2}{1-2\lambda_1 x+x^2}$$

($\zeta_i = \sqrt{m_j m_k/m_i M}$, and M is the total mass), which exists when $-1<\mathrm{Im}\,\mu<1$ and can be evaluated by the method of residues after being integrated by parts:

$$I_1 = \frac{\sinh(\mu\arctan\zeta_1)}{\mu\cosh\dfrac{\pi}{2}\mu}.$$

Introducing the function

$$\Phi_i(\mu) = \frac{1+\zeta_i^2}{\zeta_i}\,\frac{\sinh(\mu\arctan\zeta_i)}{\mu\cosh\dfrac{\pi}{2}\mu},$$

we find that $\Pi_{\mathrm{ex}} = \Phi_1(\mu)$.

After simple transformations, the function $\Pi_{\mathrm{tr}}(k)$ can be expressed in terms of the function Φ defined earlier and a new function $\tilde{\Phi}$:

$$\Pi_{\mathrm{tr}}(k) = \Phi_2(\mu)\tilde{\Phi},$$

$$\tilde{\Phi} = \frac{1+\zeta_2^2}{2\pi\zeta_2}\mathrm{PV}\int_0^\infty dx\,\frac{x^{i\mu-1}kx}{kx-a\gamma}\ln\frac{1+2\lambda_2 x+x^2}{1-2\lambda_2 x+x^2}.$$

Analyzing this integral, we note that it can easily be transformed into a contour integral about two logarithmic cuts, and the integration contours can always be chosen in such a way that either $|x|\gg 1$ or $|x|\ll 1$. This makes it possible to expand the denominator of the integrand in a series either for high momenta ($k>a\gamma$) or for low momenta ($k<a\gamma$). In the first case we get the leading term $\tilde{\Phi}=\Phi_2$ and in the second, $\tilde{\Phi}\propto k/\kappa_1$. Hence we can write the transcendental equations for μ in two asymptotic regions. Allowing for the fact that $\Pi_{\mathrm{ex}}+2\Pi_{\mathrm{tr}}=1$, we obtain

$$\Phi_1(\mu)=1,\quad k\ll\sqrt{\frac{m_{1,23}}{m_{12}}}\kappa_1, \tag{15}$$

$$\Phi_1(\mu)+2\Phi_2^2(\mu)=1,\quad k\gg\sqrt{\frac{m_{1,23}}{m_{12}}}\kappa_1. \tag{16}$$

Since Φ is an even function of μ, we conclude that the wave function Ψ_t can be represented by a linear combination of $k^{\pm i\mu-5/2}$, which coincides with the Fourier transform of the wave function in the field of the local potential

$$V_{\mathrm{eff}}(\rho) = -\frac{\mu^2+0.25}{2m_{2,13}\rho^2}, \tag{17}$$

when $|V|\gg|z|$, with different coupling constants for small and large ρ. The solutions in the field of a type-$1/\rho^2$ potential are well known. In particular, the energy levels obey the relation $E_{n-1}/E_n=\eta$, where η depends only on μ,

$$\eta = \exp(2\pi/\mu), \tag{18}$$

and either go to minus infinity, which corresponds to a fall to the center (noted by Thomas[9] as long ago as 1935) or become denser at zero, which amounts to the Efimov effect.[4] In our case the deep levels correspond to small distances and are determined by the values of μ_2 satisfying Eq. (16). On the other hand, shallow levels with $z>\varepsilon_1$ in the scattering channel generate resonances and are determined by the values of μ_1 satisfying Eq. (15).

A simple analysis of Eqs. (15) and (16) shows that μ_2 tends to μ_1 as $m_1/m_2\to 0$ with an exponential rate, and μ_1 in this limit is described well by the expansion

$$\mu_{\mathrm{as}} = c\zeta_1+\frac{c}{(c+1)\zeta_1}+O\left(\frac{1}{\zeta_1^3}\right) \tag{19}$$

($\zeta_1=\sqrt{m_2 m_3/m_1 M}\gg 1$), coinciding with the limit found earlier in Ref. 1. The constant $c=0.5671\cdots$ used above satisfies the equation $c=\exp(-c)$. To avoid having to write involved formulas representing calculations of secondary importance, we list values of μ as a function of the ratio of the masses of the light and heavy particles in Table I. The rapid convergence of μ_2 to μ_1 indicates that the second term in Eq. (16) becomes negligible (the difference in solutions is in the sixth decimal place at $m_2/m_1=100$), and this makes it possible to discard the second term in V when $m_1/m_2 \ll 1$. To demonstrate this smallness, in the last row of Table I we list values of the ratio of the potential energies $\Pi_{\mathrm{ex}}(\mu_1)$ and $\Pi_{\mathrm{tr}}(\mu_1)$.

Thus, we can assume that for a very light particle 1 the wave function Ψ_t is represented by a linear combination of the functions $k^{\pm i\mu-5/2}$ with $\mu=\mu_1$ in a broad range of momenta $(k \gg \sqrt{-2m_{2,13}z})$. Knowing this momentum asymptotic behavior is sufficient for calculating the matrix elements in the definition of Γ, since the range of momenta $\sim\kappa_1$ provides the main contribution to these integrals. The difficulty of normalizing such a function can be overcome by a simple trick. As noted earlier, our wave function corresponds to motion in the effective potential (17). The solutions in the field of such a potential are well known: $\Psi_t(\rho)\sim K_{\mu i}/\sqrt{\rho}$, i.e., they can be expressed in terms of modified Bessel functions. We can normalize Ψ_t, obtain an expression for its Fourier transform, and find the coefficients of the leading asymptotic terms $k^{\pm i\mu-5/2}$. This scheme is somewhat tedious but yields a reasonable result, which in the limit of small masses of particle 1 coincides with the exact result. On the other hand, the exact result (for zero pairwise forces) can be obtained much faster. The thing is that we can find a solution of the integral equation with the potential V_{ex} in the entire momentum range, i.e., without putting z to zero, by using the Mellin transform, as done by Danilov[10] and Minlos and Faddeev[11] in their studies of the properties of the Skornyakov–Ter-Martirosyan equation.[12] Furthermore, Minlos and Faddeev[11] found an approximate expression for the wave function of a system of three bosons with zero pairwise forces, an expression that becomes exact when the binding energies in the subsystems vanish. On other hand, after introducing the dimensionless variable $t=k/\kappa_t$ $(\kappa_t=\sqrt{-2m_{2,13}z})$ into the Lippmann–Schwinger equation for the function $\psi=k(k^2-2m_{2,13}z)^{3/4}\Psi_t$ with a potential V_{ex} in the entire momentum range, we see that the equation differs from the Skornyakov–Ter-Martirosyan equation for three bosons only by a factor of 2 and by the values of the mass constants λ_1 and $\alpha=\sqrt{m_{13}/m_{2,13}}$:

$$\psi(t)=\frac{1}{2\pi\lambda_1\alpha}\int_0^\infty \ln\frac{t^2+t'^2+2\lambda_1 tt'+\alpha^2}{t^2+t'^2-2\lambda_1 tt'+\alpha^2}\,\frac{dt'\,\psi(t')}{\sqrt{t'^2+1}}. \tag{20}$$

Hence Eq. (20) also has an analytic solution, which can be obtained by using the Mellin transform. Here we simply write the final result:

$$\psi(t)=A\sin(\mu\ln(t+\sqrt{1+t^2})),$$

whose validity can be checked by plugging ψ into Eq. (20). As before, the constant μ satisfies Eq. (15), and the constant A is determined by normalization. The integral equation (20) has a solution at all energies, but the spectrum can be fixed by imposing a constraint on the coefficients of $\sin(\mu\ln k)$ and $\cos(\mu\ln k)$ for large values of k (see Refs. 10 and 11):

$$\sin(\mu\ln\kappa_t)=b\cos(\mu\ln\kappa_t).$$

The arbitrary constant b uniquely determines the wave number $\kappa_t^{(n)}$ for a level with number n:

$$\kappa_t^{(n)}=\exp(\pi n/\mu+\arctan(b)/\mu).$$

A spectrum obtained in this way satisfies Eq. (18) and demonstrates both the effect of falling to the center for large positive values of n (noted in Ref. 11) and the Efimov effect (logarithmic crowding of levels near $z=0$) for large negative values of n (discovered later and by another method by Efimov[4]).

Normalizing the wave function Ψ_t over the entire momentum space, we arrive at the final expression for the wave function of the closed channel:

$$\Psi_t(k)=\frac{2\pi}{\sqrt{\kappa_t^3(1-\pi\mu/\sinh\pi\mu)}}\,\frac{1}{t(1+t^2)^{3/4}}\times\sin(\mu\ln(t+\sqrt{1+t^2})). \tag{21}$$

5. WIDTHS OF SUBTHRESHOLD RESONANCES

As noted earlier, the resonance width $\tilde{\Gamma}$ depends on the background scattering phase and Γ specified by Eq. (11), which is determined by the matrix elements of the potential V_{12} linking the open and closed elastic scattering channels. The background phase depends on the details of the pairwise forces at small distances and cannot be examined in the limit of zero-range forces (since it has no limit). The quantity Γ, which is an upper bound on the resonance width, can be calculated in this limit (as $\beta\to\infty$). Note that the fact that $E_k+\varepsilon_1=E_r$ implies that for small $|E_r|$ (i.e., small compared to $|\varepsilon_1|$) we can put $k_0\to\kappa_1\sqrt{m_{1,23}/m_{23}}$. Moreover, because of the simple relationship $E_r=E_t+\Gamma\sin(2\delta_f)/4$ [see Eq. (12)] and the small values of Γ in the relations we now discuss, we can put $E_r=E_t$. Then, allowing for the definition of V_{12} and for the expression (11) for the width and performing simple transformations, we arrive at

$$\frac{\Gamma}{|E_t|}=32\pi\left(\frac{m_{2,13}}{m_{23}}\right)^2|J|^2, \tag{22}$$

where J can be defined in terms of an integral over the entire momentum space:

$$J=\frac{\kappa_1}{\kappa_t}\int\frac{d^3k}{(2\pi)^3}\,\frac{(k^2+k_t^2)^{1/4}}{\kappa_1^2+k^2-\mathbf{k}_0\cdot\mathbf{k}+k_0^2/4}\,\Psi_t(k).$$

Using the explicit form of the wave function of the closed channel [Eq. (21)], we can evaluate this integral analytically. Since this is a moderately involved procedure, we only discuss the problem schematically. We perform trivial angular integration, reduce the result to dimensional form by introducing $t=k/\kappa_t$, and change variables, $t=(x^2-1)/2x$. The last step gets rid of radicals and changes the interval of integration from $(0,\infty)$ to $(1,\infty)$. The integrand is invariant under the transformation $x\to 1/x$, which allows us to return to the integration interval $(0,\infty)$. After integrating by parts we get rid of the logarithmic function. The resulting integral I of type

$$I=\int_0^\infty x^{i\mu}\frac{Q_1(x)}{Q_2(x)}dx,$$

where Q_1 and Q_2 are polynomials that guarantee convergence at zero and infinity, can be evaluated by a standard trick: the integral along the upper side of the power-law cut is expressed in terms of an integral along a contour surrounding the cut. The contour can be closed at infinity and the integral can be expressed in terms of the residues at the zeroes of the polynomial Q_2.

The entire procedure of integration leads to an expression for J:

$$J=\frac{1}{2\sqrt{1-\pi\mu/\sinh\pi\mu}}\frac{\sinh(\mu\arctan\zeta_2)}{\mu\zeta_2}\times\frac{\sin(\mu\ln 2\sqrt{\varepsilon_1/E_t})}{\cosh(\pi\mu/2)}. \tag{23}$$

Equations (22) and (23) have remarkable properties. First, the relative width $\Gamma/|E_t|$ is independent of the number of the resonance, since the positions of the resonances are determined by (18), which leads to a situation in which the values of J for resonances with numbers n_i and n_{i+k} differ only by the factor $(-1)^k$. Hence on the logarithmic scale not only the positions of resonances are equidistant but so are the resonance widths. Second, the relative resonance width decreases exponentially with the mass of particle 1. To verify this we write the limiting expression ($m_1/m_2 \ll 1$) for the relative width. With allowance for (19) and the fact that

$$\left(\frac{\sinh(\mu\arctan\zeta_2)}{\mu\zeta_2}\right)^2\to 1.024\ldots,$$

we can write this limiting expression as

$$\frac{\Gamma}{|E_t|}\simeq 32\pi 1.024\exp\left\{-\pi c\sqrt{\frac{m_2}{2m_1}}\right\}\sin^2\left(\frac{\mu}{2}\ln 4\frac{\varepsilon_1}{E_t}\right). \tag{24}$$

Since the observed width $\tilde{\Gamma}$ is always less than Γ, we can place an upper bound on the width of any subthreshold resonance:

$$\tilde{\Gamma}<103.0|E_r|\exp\left(-1.260\sqrt{\frac{m_2}{m_1}}\right), \tag{25}$$

where we have returned to the notation E_r, since the difference between E_t and E_r is exponentially small. Formula (25) allows us to make an estimate on the lifetime of highly excited negative ions of diatomic molecules. Already for a proton-to-electron mass ratio and a clearly overvalued scale of 1 eV for the binding energy of the complex ($|E_r|$), the resonance lifetime $1/\tilde{\Gamma}$ is longer than 10^6 s and exceeds all possible relaxation times for gases. What is of interest, however, is systems with a low energy of electron–atom affinity. As noted in the Introduction, negative ions of molecules of alkali metals are such systems. Plugging in the mass ratio for lithium, we arrive at a lifetime of 10^{40} s, which is longer than the lifetime of the universe.

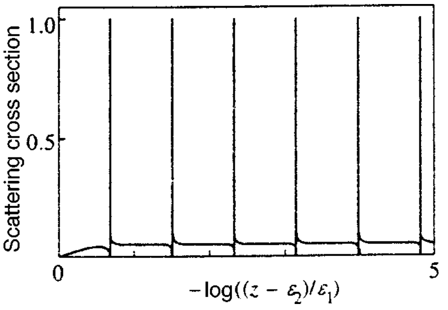

FIG. 1. Scattering cross section at $m_2/m_1=65$.

6. NUMERICAL SOLUTIONS

To demonstrate the validity of the analytic calculations, the Faddeev integral equations (7) describing the scattering of particle 1 by a pair of identical particles (2,3) were solved numerically. The calculation procedure was the one described in Ref. 5. This made it possible to examine scattering as close to the reaction threshold as possible. Actually, these calculations do not require that we come very close to the threshold, since the coupling constant μ is large. This resulted in an interval ranging from zero kinetic energy ($z=\varepsilon_1$) to an energy that was 10^{-8} MeV distant from the transformation threshold ($z-\varepsilon_2=-10^{-8}$ MeV). Here the transformation threshold differed from the three-particle threshold by 10^{-15} MeV ($\varepsilon_2=-10^{-15}$ MeV) and the binding energy of the pair (2,3) was $\varepsilon_1=-10^{-3}$ MeV. The masses of particles 2 and 3 equaled the nucleon mass, and the constant β determined by the ranges of the forces was chosen equal to $0.72\,\mathrm{fm}^{-1}$ for all pairwise potentials. The units of energy (MeV) and length (fm) reflect the use of Yamaguchi potentials with parameters of a nuclear problem (masses and ranges) and a characteristic potential energy of several tens of MeV. Thus, small but finite values of the ranges of the pairwise forces and binding energies in the pairs were selected for the calculations. Since the problem is determined entirely by the ratio of dimensional quantities, the results are also valid in atomic units. For example, with the atomic scale of the radius of forces amounting to 1 Å ($r_0\simeq 3/\beta$), the condition (2) for pairs,

$$\kappa_1/\beta_1\simeq 6.7\times 10^{-3},\qquad \kappa_2/\beta_2=10^{-6}\sqrt{m_1/m_2}\,\kappa_1/\beta_1,$$

corresponds to an atom–atom scattering length ($1/\kappa_1$) of about 50 Å.

Figure 1 depicts the cross sections normalized to the unitary limit in the S-wave: $\sigma_{\mathrm{un}}=4\pi/k_0^2$ for a mass ratio $m_2/m_1=65$. Since the resonances crowd in toward $z=0$, the energy is presented on a logarithmic scale. The fact that resonances are positioned equidistantly is clearly visible. Unfortunately, even the maximum width in the range of masses of particle 1 (see Fig. 2) remains very small, and the resonances look essentially like straight lines. Table II lists the positions of the first six resonances, the ratios of the energy of the previous resonance to the energy of the given resonance, i.e., the values of η_{calc}, and the relative resonance widths. For comparison, here are the values of μ_1 and μ_2 that are solutions of Eqs. (15) and (16), respectively:

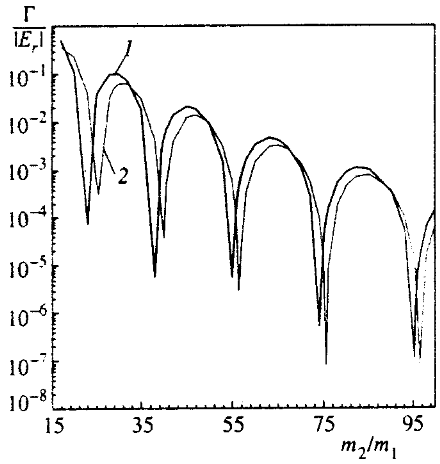

FIG. 2. Relative resonance widths. Curve *1* represents the numerical solution of the Faddeev equation, and curve *2* represents the analytic dependence.

$$\mu_1 = 3.29589\ldots, \quad \mu_2 = 3.29654\ldots.$$

The corresponding values of η are [see Eq. (18)]

$$\eta(\mu_1) = 6.728\,23\ldots, \quad \eta(\mu_2) = 6.726\,08\ldots.$$

Table II shows that already the second-to-third resonance energy ratio lies within the range of analytic values of η with an accuracy to four decimal places. Note that the first resonance is far from the threshold. The last column in Table II shows that the relative widths are almost independent of the number of the resonance. For example, starting with the third resonance, variations occur only in the fourth decimal place. Thus, Fig. 1 and Table II support fairly well the conclusion that the positions of resonances are logarithmically equidistant, and so are the resonance widths, which translates into constant relative resonance widths.

To demonstrate to what extent the expressions (22) and (23) describe the analytic relationship between resonance widths and positions of resonances for small mass ratios m_1/m_2, Fig. 2 depicts the relative widths obtained by solving the integral equations numerically (curve *1*) and by using the expression (22) with values of the energies of resonances found numerically (curve *2*). Clearly, the curve representing the analytic dependence of the resonance widths on mass ratio follows fairly closely the curve representing the results of numerical calculations, although there exists a small phase shift. Note some features of the dependence of the widths on the mass of the light particle 1. First, the resonance width is not a monotonic function of mass. Second, the peaks in the resonance widths exponentially decrease with the mass of particle 1, correctly reflecting the estimate (24). For example, in the range of mass variations (m_2/m_1) depicted in the picture, i.e., from 15 to 100, the resonance widths decrease by a factor of 1000. The marked nonmonotonic behavior of the relative widths suggests that there is a difference between analytic estimates and the results of numerical calculations. If the estimates rely on the sharp local minima, the ''observed'' phase shift amounts to 2 mass units for the first minimum and monotonically decreases to 1.25 mass unit near the last local minimum $(m_2/m_1 \sim 95)$. Thus, the difference between analytic and numerical estimates decreases with the mass m_1.

TABLE II. Parameters of the first six resonances at $m_2/m_1 = 65$.

N	E_r/ε_1	η_{calc}	$\Gamma/\lvert E_r\rvert$
1	2.1236×10^{-1}		4.413×10^{-3}
2	3.1480×10^{-2}	6.746	4.246×10^{-3}
3	4.6787×10^{-3}	6.728	4.208×10^{-3}
4	6.9540×10^{-4}	6.728	4.202×10^{-3}
5	1.0338×10^{-4}	6.727	4.201×10^{-3}
6	1.5369×10^{-5}	6.726	4.200×10^{-3}

7. CONCLUSION

In studying the problem of estimating the lifetimes of the Efimov states of negative ions of diatomic molecules, we were able to reduce it to the problem of the scattering of a light particle (an electron) by a bound pair of two heavy particles (a molecule), provided that condition (2) is met. Such a three-body statement of the problem makes it possible, on the one hand, to derive analytic relationships linking the widths and positions of resonances and, on the other, to verify these relationships directly by a numerical solution of the Faddeev equations. Here for the first time we calculated the Efimov resonances below the transformation threshold of a three-particle system.

The conclusion that can be drawn from the present study is that the lifetimes of the novel molecular states discussed in Refs. 1–3 suggest that these are bounds states for all types of physical processes.

Note that the assumption that there are real systems obeying condition (2) rests on the known values of the electron–atom and atom–atom scattering lengths, which may be as large as tens and even hundreds of angstroms. Here the experimental values of the effective interaction ranges are unknown, and only dimensional considerations (atomic sizes) and the successful use of zero-range potentials in atomic physics (see, e.g., Demkov and Ostrovskiĭ's monograph[13]) support the hypothesis that condition (2) operates in real molecular systems.

This research was carried out within the framework of the K-40 ISTC Project.

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

[1)] In this paper the Planck constant $\hbar = 1$.

---

[1] F. M. Pen'kov, Zh. Éksp. Teor. Fiz. **106**, 1046 (1994) [JETP **79**, 568 (1994)].

[2] F. M. Pen'kov, Zh. Éksp. Teor. Fiz. **109**, 721 (1996) [JETP **82**, 387 (1996)].

[3] F. M. Pen'kov, Zh. Éksp. Teor. Fiz. **111**, 1229 (1997) [JETP **84**, 678 (1997)].

[4] V. Efimov, Yad. Fiz. **12**, 1080 (1970) [Sov. J. Nucl. Phys. **12**, 589 (1971)].

[5] F. M. Pen'kov and N. Zh. Takibaev, Yad. Fiz. **57**, 1300 (1994) [Phys. At. Nucl. **57**, 1232 (1994)].

[6] S. P. Merkur'ev and L. D. Faddeev, *Quantum Scattering Theory for Several-Particle Systems* [in Russian], Nauka, Moscow (1985), p. 69 [English transl.: L. D. Faddeev and S. P. Merkur'ev, Kluwer, Boston (1993)].

[7] N. Zh. Takibaev and F. M. Pen'kov, Yad. Fiz. **50**, 373 (1989) [Sov. J. Nucl. Phys. **50**, 234 (1989)].

[8] V. B. Belyaev, *Lectures on the Theory of Few-Body Systems* [in Russian], Énergoatomizdat, Moscow (1986), p. 37.

[9] L. H. Thomas, Phys. Rev. **47**, 903 (1935).

[10] G. S. Danilov, Zh. Éksp. Teor. Fiz. **40**, 498 (1961) [Sov. Phys. JETP **13**, 349 (1961)].

[11] R. A. Minlos and L. D. Faddeev, Zh. Éksp. Teor. Fiz. **41**, 1850 (1961) [Sov. Phys. JETP **14**, 1315 (1962)].

[12] G. V. Skornyakov and K. A. Ter-Martirosyan, Zh. Éksp. Teor. Fiz. **31**, 775 (1956) [Sov. Phys. JETP **4**, 648 (1957)].

[13] Yu. N. Demkov and V. N. Ostrovskiĭ, *Zero-Range Potentials and their Applications in Atomic Physics*, Plenum Press, New York (1988).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Multiphoton ionization of molecules under the conditions of strong field-induced perturbation of Rydberg states

G. K. Ivanov and G. B. Golubkov*)

N. N. Semenov Institute of Chemical Physics, Russian Academy of Sciences, 117977 Moscow, Russia

(Submitted 9 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 1987–2000 (June 1999)

Multiphonon ionization of the H_2 molecule under the action of a weak (probe) field, which provides the initial population of the low-lying (working) level, and intense monochromatic linearly polarized radiation is studied. The multiphoton ionization process occurs under the conditions of strong field perturbation of two intermediate Rydberg series, $np0(^1\Sigma_u^+)$ and $np2(^1\Pi_u)$, of the optical $R(0)$branch which have different ionization potentials. The series are occupied simultaneously as a result of single-photon absorption by an excited H_2^* molecule in the working state $4s\sigma H'\,^1\Sigma_g^+$ $(\nu=0)$. As a result of the irregularity in the arrangement of the intermediate levels from a large group of states that are combined in the multiphoton ionization process a sharp and irregular change occurs in the dependence of the shifts and widths Γ_n of the levels on the intensity f of the strong field in a transition from one level to another. It is shown that for field intensities f such that the level widths remain much less than the splitting between the levels $(\Gamma_n \ll 1/n^3)$ the stabilizing effect (i.e., the field-induced narrowing of the levels as $f \rightarrow \infty$) in the form $\Gamma_n \propto 1/f^2$ (as happens in atoms with a structureless core) is not observed in molecular systems. *© 1999 American Institute of Physics.* [S1063-7761(99)00606-X]

1. INTRODUCTION

The development of a theory of multiphoton ionization of atoms and molecules taking account of strong field-induced perturbations of the system under study is a key problem of modern laser chemistry. Although the need to introduce field-dressed states arose at the beginning of the 1970s,[1,2] specific theoretical eleborations for doing so are still very limited. This applies completely to Rydberg molecules, where the effects of strong nonadiabatic coupling of the electronic and nuclear motions are pronounced. These effects are manifested in the fact that each molecular Rydberg state is a superpositional-type state, whose individual components correspond to different levels of the rotational N and vibrational v excitations of the ion. The quantum defects μ of the Rydberg levelsand the coefficients in the expansion of the wave functions acquire a strong dependence on the principal quantum number n (Ref. 3).The regularity of the arrangement of the levels, which greatly simplifies the analysis of processes in which highly excited atoms participate, is thereby destroyed. The specific properties of Rydberg molecules have a direct effect on the processes in which interference of the contributions of large groups of Rydberg states which are drawn into a strong interaction with an external electromagnetic field plays an important role. Specifically, here the interference suppression of the decay of Rydberg states (the stabilization effect), the study of which up to now has been limited to atomic systems,[4–9] should be impeded.

Stabilization consists in a decrease of the level widths Γ_n with increasing external field intensity f (after a certain critical value has been reached) and is manifested in two fundamentally different forms. For superstrong fields $f \sim f_a$, where f_a is the intensity of the field of a hydrogen atom in the ground state), stabilization occurs when the widths Γ_n of the Rydberg states mixed as a result of Λ transitions become equal to the splitting between neighboring levels, i.e., $\Gamma_n \sim 1/n^3$ (Ref. 7).

The second mechanism producing the stabilization of the function $\Gamma_n(f)$ is due to V transitions from a low-lying resonant level at lower field intensities f, so that the shifts of the perturbed levels become comparable to $1/n^3$ (the widths Γ_n themselves remain small compared to the splitting between the levels, i.e., $\Gamma_n \ll 1/n^3$).[8,9] In addition, for a monochromatic field with frequency ω_f it can be assumed that1):

$$\eta = f\omega_f^{-5/3} \ll 1. \tag{1}$$

The present paper is devoted to a study of the characteristic features of multiphoton ionization spectra that are due to strong nonadiabatic coupling with rotation. Light molecules with a small moment of inertia are of the main interest from this standpoint. The most inviting such molecule is the hydrogen molecule, whose optical properties are well known. For this reason, it is worthwhile to investigate the following photoionization scheme. For the $X\,^1\Sigma_g^+$ ground state of para-H_2 the population of the working level ''0'' should be studied in the two-photon (2Ω) absorption regime with excitation of the singlet $s\sigma$ state. In what follows, the state $4s\sigma H'\,^1\Sigma_g^+$ $(\nu=0)$ is chosen as the working state. Then, in an intense monochromatic laser field (with frequency ω_f) the classical $np0(^1\Sigma_u^+)$ and $np2(^1\Pi_u)$ two-channel Fano system of the optical branch $R(0)$ will be populated at the intermediate stage.[3] This makes it possible to analyze quite simply and clearly the role of nonadiabatic coupling with

1063-7761/99/88(6)/8/$15.00

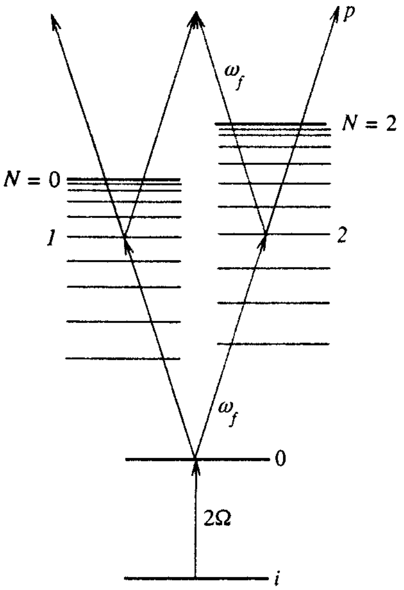

FIG. 1. Scheme of multiphoton ionization of a molecule with excitation (at an intermediate step) of a two-component $R(0)$ branch of the $np0(^1\Sigma_u^+)$ and $np2(^1\Pi_u)$ Rydberg series, converging to the ionization thresholds $N=0$ and $N=2$ and labeled by the indices 1 and 2, respectively; i — initial state of the molecule, 0 — position of the $4s\sigma H'\ ^1\Sigma_g^+\ (\nu=0)$ working level populated as a result of the absorption of two weak-field photons.

rotation in field-induced stabilization, since all structural features of the spectrum can be represented here in analytic form. Figure 1 shows the scheme of the process of interest

$$H_2(J_i=N_i=0)\xrightarrow{2\Omega}H_2^*(J''=l''=N''=0)\xrightarrow{\omega_f}H_2^{**}\times\left(J'=L'=1,\ \begin{matrix}N'=0\\ N'=2\end{matrix}\right)\xrightarrow{\omega_f}e^-+H_2^+. \tag{2}$$

Here J_i, J'', and J' are the total angular momenta of the initial and intermediate states, respectively, l is the orbital angular momentum of the optical electron, and N is the rotational quantum number.

Transitions through the Rydberg series $np0$ and $np2$ are studied in the continuous excitation regime from a lower-lying ''0'' level, which is populated in advance by a weak (probe) field with frequency Ω. Under these conditions it is possible to implement a quite simple and rigorously validated computational procedure based on a stationary variant of the theory of Refs. 8 and 9 where the wave functions of the continuous spectrum of the system $e^-+\ XY^+$, which are found by taking account of the strong field-induced mixing of large groups of states, are expressed in terms of the elements of the **T** matrix for radiative collisions of the particles e^- and XY^+ in a strong radiation field with frequency ω_f. The expressions obtained make it possible to find not only the photoelectron spectra but also the energy eigenvalues of highly excited molecules in a strong electromagnetic field. We note that the question of the behavior of highly excited molecules in a laser radiation field in the presence of strong nonadiabatic coupling with rotation has not been previously discussed in the literature and is studied here for the first time.

2. GENERAL FORMULATION OF THE PROBLEM

Since for the scheme considered here the weak (probe) field can be taken into account in perturbation theory, the basic problem reduces to determining the continuum wave functions Ψ_p of the system e^-+XY^+ taking account of the strong electromagnetic field, where perturbation theory is inapplicable in principle. The idea of the method proposed in Ref. 8 consists in the fact that under the condition (1) an electron strongly interacts with the radiation field only at small distances from the ion core. For this reason, the formal apparatus of multichannel scattering theory can be used to construct the wave functions Ψ_p.

Energy transfer in the system e^-+XY^+ is determined by ρ-vibronic transitions in the ion core XY^+ and, taking account of the effect of the field, also by the possibility of induced absorption $(k<0)$ or emission $(k>0)$ of external-field photons, for which the energy of the system changes by the amount $k\omega_f$. The index k characterizes the change in the number of photons in an external electromagnetic field, assumed below to be linearly polarized, with the vector **f** oriented along the z axis in the laboratory coordinate system.

The problem can be solved as follows. Basis channel wave functions $|q_k\rangle=|Jl\nu Nk\rangle$ of the zeroth-order Hamiltonian, which correspond to different values of k and fixed vibrational ν and rotational N states of the ion, are introduced into the theory to describe the intermediate Rydberg states. These functions take account of only the Coulomb interaction of the particles e^- and XY^+. We recall that for symmetric Rydberg molecules X_2^{**} the electron orbital angular momentum l is a good quantum number. In what follows, the perturbation of the Coulomb states $|q_k\rangle$, which is due to both the interaction of an electron with the ion core (which distorts the Coulomb field at short distances) and the interaction $\mathbf{V}^f=\mathbf{f}\cdot\mathbf{D}/2$ with the radiation field (**D** is the dipole moment operator), is investigated in the formalism of modified Lippmann–Schwinger equations.

The ionization of the H_2 molecule according to the scheme shown in Fig. 1 can be described by the following expression for the transition amplitude into the final continuum state $|p\rangle$ of the system $e^-+\ H_2^+$:

$$M_{ip}=A_{i0}\frac{1}{E-E_0}T_{0p}, \tag{3}$$

where $E=E_i+2\Omega$, E_i is the energy of the initial state measured from the spectrum limit, and E_0 is the position of the working ''0'' level. The quantity A_{i0} is the transition amplitude to the zero level as a result of absorption of two weak-field photons Ω.

The structural features of the process (2) are contained in the **T** matrix for radiative collisions, which satisfies the fundamental equation[8]

$$\mathbf{T}=\mathbf{t}+\mathbf{t}\sum_{q_k}|q_k\rangle\langle q_k|\cot\nu_{qk}\mathbf{T}. \tag{4}$$

Here $\nu_{qk}=[2(E_q+k\omega_f-E)]^{-1/2}$ is the effective principal quantum number in the q_k channel $(k<0)$ and E_q is the excitation energy of the ion. The summation indices in Eq. (4) include the open p and closed c channels, i.e., q_k

$=\{c_k,p_k\}$. The states $|q_k\rangle$ are normalized to a delta function of the energy $\langle q_k(\varepsilon)|q_k(\varepsilon')\rangle=\pi\delta(\varepsilon-\varepsilon')$. For a finite number of strongly coupled states the operator equation (4) reduces to a system of algebraic equations. The matrix elements appearing in it have a transparent physical meaning. For example, the diagonal elements $t_{c_kc_k}$ can be expressed in terms of the characteristic quantum defects (including corrections responsible for the Stark shifts of the levels); the quantities $t_{c_kp_{k'}}$ are related to the definition of the amplitudes of natural (for $k=k'$) and field-induced (for $k\neq k'$) decays; and the elements $t_{c_kc_{k'}}$ $(k'=k\pm1)$ are responsible for the field-induced interaction of the discrete states.

The weakly energy-dependent **t** operator includes the electrostatic interaction $\mathbf{V}^e$, responsible for the nonCoulomb part of the potential, and the interaction $\mathbf{V}^f$, which under the condition (1) must be treated as weak. Taking account of the terms which are linear and quadratic in the interaction $\mathbf{V}^f$ it has the form

$$\mathbf{t}=\mathbf{t}^e+\mathbf{t}^f,\quad \mathbf{t}^f=\boldsymbol{\Omega}^e(\mathbf{V}^f+\mathbf{V}^f\mathbf{G}\mathbf{V}^f)\boldsymbol{\Omega}^e, \tag{5}$$

where the operator describing the interaction of the electron with the ion core

$$\mathbf{t}^e=\mathbf{V}^e+\mathbf{V}^e\mathbf{G}\mathbf{V}^e=\mathbf{V}^e\boldsymbol{\Omega}^e=\boldsymbol{\Omega}^e\mathbf{V}^e$$

is expressed in terms of the part of the Green's function G of the Rydberg molecule that is a smooth function of energy and is found by taking account of $\mathbf{V}^e$. In addition, the operator $\boldsymbol{\Omega}^e$, which describes the distortion due to interaction of an electron with the ion core, appears in the expression for the field-interaction matrix $\mathbf{t}^f$.

3. MATRIX ELEMENTS OF THE OPERATORS $\mathbf{t}^e$ AND $\mathbf{t}^f$

We shall now determine the explicit form of the electrostatic $\mathbf{t}^e$ and field $\mathbf{t}^f$ interaction operators. Since the H_2 molecule in the initial ${}^1\Sigma_g^+$ state and the low-lying $4s\sigma H'\ {}^1\Sigma_g^+$ states is not vibrationally and rotationally excited, the vibronic coupling in the intermediate Rydberg states $np0({}^1\Sigma_u^+)$ and $np2({}^1\Pi_u)$ is weak and can be neglected to a first approximation. This makes it possible to neglect vibrational transitions.

Then in the total angular momentum (J) representation, where J is the total angular momentum of the system $e^-+H_2^+$, the basis wave functions of the zeroth-order Hamiltonian are given by

$$|q_k\rangle=\varphi_{lk}^{JN}=g_{lk}^N(r)\Phi_{lN}^{JM}(\hat{r},\hat{R}). \tag{6}$$

For short distances between the electron and the ion core

$$(l+1/2)^2\ll r\ll|\varepsilon_{Nk}|^{-1}$$

(which make the main contribution to the transition dipole matrix elements) the radial part of the Coulomb wave function $g_{lk}^N(r)$, which is regular at the origin, is

$$g_{lk}^N(r)=\left(\frac{2}{r^3}\right)^{1/4}\sin\alpha_{lk}^N(r). \tag{7}$$

The corresponding semiclassical phase for $l\ll|\varepsilon_{Nk}|^{-1/3}$ is

$$\alpha_{lk}^N(r)=\sqrt{8r}+(\varepsilon_{Nk}/3)\sqrt{2r^3}-\pi l-\pi/4, \tag{8}$$

$$\varepsilon_{Nk}=BN(N+1)+k\omega_f-E$$

where B is the corresponding rotational constant.

The complete angular wave function of the system has the form

$$\Phi_{lN}^{JM}(\hat{r},\hat{R})=\sum_m Y_{lm}(\hat{r})Y_{N,M-m}(\hat{R})(lNmM-m|JM), \tag{9}$$

where $Y_{lm}(x)$ are spherical harmonics, $(lNmM-m|JM)$ are vector coupling coefficients, $\hat{r}$ and $\hat{R}$ are angular variables giving the directions along the radius vector of the electron $\mathbf{r}$ and along the molecular axis $\mathbf{R}$ in the laboratory coordinate system, and M is the projection of the total angular momentum J of the system. The $\mathbf{t}^e$-operator elements appearing in Eq. (4) are defined in the Coulomb basis $|q_k\rangle$ and are related by a unitary transformation

$$t_{lN,l'N'}^{e(J)}=-[\mathbf{U}\tan(\pi\boldsymbol{\mu})\mathbf{U}^T]_{lN,l'N'}$$

with a diagonal matrix of characteristic phases $\tan(\pi\boldsymbol{\mu})$, whose values $\tan(\pi\mu_{l\Lambda}^{(J)})$ are taken at the equilibrium position R_e of the nuclei and correspond (for fixed J) to all possible projections Λ of the angular momentum l of the electron on the axis of the molecule. For homo-atomic molecules this relation is given by the expression[10]

$$\begin{aligned}t_{lN,l'N'}^{e(J)}=-\sum_{\Lambda,j}&(-1)^{J+\Lambda}[(2j+1)(2N+1)]^{1/2}\\&\times(ll',\Lambda-\Lambda|j0)(jN,00|N'0)W(ll'NN';jJ)\\&\times\tan(\pi\mu_{l\Lambda}^{(J)}),\end{aligned} \tag{10}$$

where $W(ll'NN';jJ)$ are Racah coefficients.

The rotation matrix $\mathbf{U}$ gives rise to a transition from the adiabatic representation $l\Lambda$ to the nonadiabatic lN representation. For example, for the two-channel Fano system its elements are

$$U_{N\Lambda}^{Jl}=\langle lN|l\Lambda\rangle=(-1)^{l+\Lambda}(2-\delta_{\Lambda0})^{1/2}(lJ\Lambda-\Lambda|N0). \tag{11}$$

The elements of the reaction matrices are then given by the expressions

$$\begin{aligned}t_{10,10}^{e(1)}&=-\frac{1}{3}(\tan(\pi\mu_{p\sigma}^{(1)})+2\tan(\pi\mu_{p\pi}^{(1)})),\\t_{10,12}^{e(1)}&=-\frac{\sqrt{2}}{3}(\tan(\pi\mu_{p\sigma}^{(1)})-\tan(\pi\mu_{\mu\pi}^{(1)})),\\t_{12,12}^{e(1)}&=-\frac{1}{3}(2\tan(\pi\mu_{p\sigma}^{(1)})+\tan(\pi\mu_{p\pi}^{(1)})).\end{aligned} \tag{12}$$

The field matrix elements, according to Eq. (5), must be calculated using the modified functions

$$\Psi_{lk}^{JN}=\boldsymbol{\Omega}^e\varphi_{lk}^{JN}=\varphi_{lk}^{JN}-\sum_{N'}t_{lN,lN'}^{e(J)}\tilde{\varphi}_{lk}^{JN'}, \tag{13}$$

which contain, in addition to the regular Coulomb functions, the irregular Coulomb functions

$$\tilde{\varphi}_{lk}^{JN}=\tilde{g}_{lk}^N(r)\Phi_{lN}^{JM}(rR), \tag{14}$$

where

$$\tilde{g}_{lk}^{N}(r)=\left(\frac{2}{r^3}\right)^{1/4}\cos\alpha_{lk}^{N}(r).$$

The elements of the matrix $\mathbf{t}^f$ must be found using the condition (1) and the linear approximation in Eq. (5), i.e., they must be represented in the form

$$\mathbf{t}^f=-\frac{1}{2}\mathbf{\Omega}^e\mathbf{f}\cdot\mathbf{r}\mathbf{\Omega}^e=-\frac{2\pi}{3}f\mathbf{\Omega}^e r\sum_s Y_{1s}^{*}(\mathbf{n}_f)Y_{1s}^{*}(\mathbf{n}_r)\mathbf{\Omega}^e,$$

where $\mathbf{n}_f$ and $\mathbf{n}_r$ are unit vectors. Since the selection rule allows in this case only transitions with $\Delta M=0$, choosing the z axis so that one component of the vector $\mathbf{f}$ is nonzero (corresponding to $s=0$), we set $M=0$. Then the matrix elements of the field interaction are

$$\begin{aligned}V_{lNk,l'N'k'}^{JJ'}=-f\sqrt{\frac{\pi}{3}}\Big\{&\langle\varphi_{lk}^{JN}Y_{10}(\mathbf{n}_r)|r|\varphi_{l'k'}^{J'N'}\rangle\delta_{NN'}\\&-t_{lN,lN'}^{e(J)}\langle\tilde{\varphi}_{lk}^{JN'}Y_{10}(\mathbf{n}_r)|r|\varphi_{l'k'}^{J'N'}\rangle\\&-t_{l'N',l'N}^{e(J')}\langle\varphi_{lk}^{JN}Y_{10}(\mathbf{n}_r)|r|\tilde{\varphi}_{l'k'}^{J'N}\rangle\\&+\sum_{N''}t_{lNlN''}^{e(J)}t_{l'N''l'N'}^{e(J')}\langle\tilde{\varphi}_{lk}^{JN''}Y_{10}(\mathbf{n}_r)|r|\tilde{\varphi}_{l'k'}^{J''N''}\rangle\Big\}.\end{aligned}\tag{15}$$

Next, using the well-known properties of the Coulomb wave functions (6) and (11) and the value of the radial integral

$$\begin{aligned}r_{lk,l'k'}^{NN'}=\langle g_{lk}^{N}|r|g_{l'k'}^{N'}\rangle=\frac{1}{\pi}\Gamma\left(\frac{2}{3}\right)\\\times\left(\frac{4}{3}\right)^{1/3}\omega_f^{-5/3}\cos\left[\pi\left(\Delta n_{lNk,l'N''k'}+\frac{1}{6}\right)\right],\end{aligned}$$

where $\Delta n_{lNk,l'N'k'}$ is the difference of the corresponding principal quantum numbers of the final and initial states, we have to within a sign

$$\begin{aligned}V_{lNk,l'N'k'}^{JJ'}=\frac{f\omega_f^{-5/3}}{2\sqrt{\pi}}\Gamma\left(\frac{2}{3}\right)\left(\frac{4}{3}\right)^{1/3}\Big[&A_{NN}^{J'l'}S_{lNk,l'Nk'}^{JJ'}\delta_{NN'}\\&-\sum_{N''}B_{N'N''}^{J'l'}t_{lN,lN''}^{e(J)}S_{lN''k,l'N''k'}^{JJ'}\Big],\end{aligned}\tag{16}$$

where the coefficients are

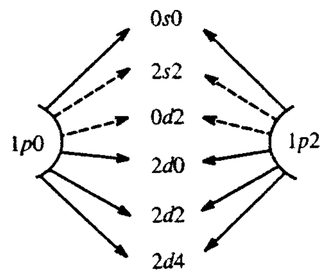

FIG. 2. Scheme giving rise to the population of $P(1)$ and $R(1)$branches by single-photon dipole transitions into the ionization continua.

$$A_{NN}^{J'l'}=1+\frac{t_{l'N,l'N}^{e(J')}}{\sqrt{3}}\quad\text{and}\quad B_{N'N''}^{J'l'}=\frac{\delta_{N'N''}}{\sqrt{3}}-t_{l'N',l'N''}^{e(J')}$$

and the corresponding angular parts are

$$\begin{aligned}S_{lNk,l'Nk'}^{JJ'}&=\langle\Phi_{lN}^{J0}(\hat{r},\hat{R})Y_{10}(\mathbf{n}_r)|\Phi_{l'N}^{J'0}(\hat{r},\hat{R})\rangle\\&=\left[\frac{3(2l+1)(2J+1)}{4\pi}\right]^{1/2}(l100|l'0)\\&\quad\times(1J00|J'0)W(1lJ'N;l'1).\end{aligned}\tag{17}$$

It is evident from the expression obtained that transitions with a change in N are due to the second term in Eq. (16), which contains the off-diagonal elements of the $\mathbf{t}^e$ reaction matrix, which give the nonadiabatic coupling with the rotation. The angular parts remain strictly diagonal in N. Therefore two Rydberg series $N=0$ and $N=2$ are excited from the lower-lying ''0'' level (in Fig. 1 these series are denoted by the indices 1 and 2, respectively), though the ''0'' level itself belongs to the state with $N=0$.

We note that the matrix elements of the field interaction (16) change sign depending on ω_f, since $\Delta n\sim n^3\omega_f$. It is also important to emphasize that Λ transitions through the ionization continua are taken into account in Eq. (4) on an equal basis.

4. MULTIPHOTON IONIZATION OF THE H_2 MOLECULE

For single-photon absorption, seven final states corresponding to three optical branches can be populated simultaneously from the intermediate Rydberg states in the process (2). They are the following:

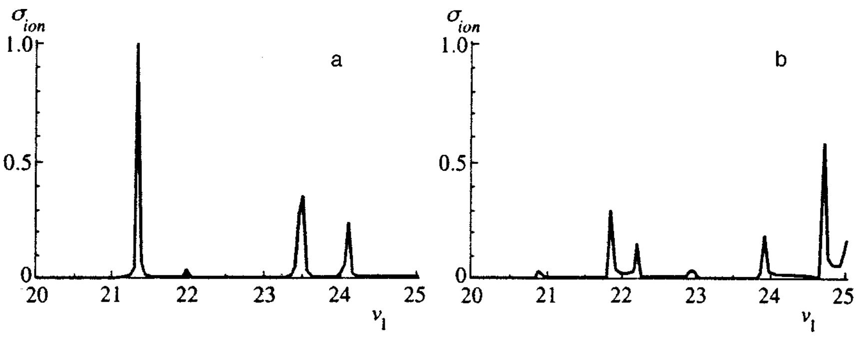

FIG. 3. Ionization cross sections (2), normalized to unity and calculated using Eq. (20) with data from Refs. 11 and 12 for field intensities $f=10^{-3}$ a.u. (a) and $f=10^{-2}$ a.u. (b).

a) the $P(1)$ branch with formation of the total angular momentum $J=0$, where the $l=N=0$ and $l=N=2$ series are mixed;

b) the $Q(1)$ branch with total angular momentum $J=1$, where one series $l=N=2$ is populated; and

c) the $R(1)$ branch, corresponding to the total angular momentum $J=2$. Here three different situations, depending on ω_f, are possible. One is the Butler region of the spectrum $0<E+\omega_f<6B$, which contains the three closed channels ($l=0$, $N=2$), ($l=N=2$), and ($l=2, N=4$), where the ($l=2$, $N=4$) channel is populated only as a result of rescattering of an electron by the ion core with the transition $N=2\rightarrow N=4$. In the next region $6B<E+\omega_f<20B$ one channel ($l=2$, $N=4$) is closed. Finally, for $20B<E+\omega_f$ all channels are open and the coupling between them is broken. This is the physical situation that we shall study below.

The scheme resulting in population of the $P(1)$ and $R(1)$ branches is shown in Fig. 2. Here six final JlN states, which have four general ionization continua ($0s0$, $2d0$, $2d2$, and $2d4$), corresponding to V transitions, form. The states $2s2$ and $0d2$ make a negligible contribution to the process under study. The optical $Q(1)$ branch is also weak and will be neglected below. Thus a six-channel scheme taking account of the P and R branches must be taken into account.

The expression

$$T_{0p}=\frac{(z_1+i\gamma_{11})V^f_{2p}V^f_{20}+(z_2+i\gamma_{22})V^f_{1p}V^f_{10}+(t_{12}-i\gamma_{12})(V^f_{2p}V^f_{10}+V^f_{1p}V^f_{20})}{D}, \tag{18}$$

is obtained for the transition amplitude T_{0p} into a state of the p continuum, i.e., into one of the six states enumerated above. For convenience we introduce here the following notation (the indices J, l, and N are dropped):

$$z_1=\tan(\pi\nu_1)-t_{11}, \quad z_2=\tan(\pi\nu_2)-t_{22},$$

$$\gamma_{ss'}=\sum_p V^f_{sp}V^f_{ps'}, \quad s=1,2,$$

where

$$\nu_1=[-2(E+\omega_f)]^{-1/2} \quad \text{and} \quad \nu_2=[2(6B-E-\omega_f)]^{-1/2}$$

are the effective principal quantum numbers. The matrix elements $t_{ss'}$ describe the nonadiabatic coupling with rotation and are given by Eqs. (12). The field coupling is takes place through the elementsV^f_{0s} and V^f_{sp}, which are given by Eqs. (16) and (17). The denominator in the expression (18), i.e., the determinant obtained from the system of algebraic equations (4), is

$$D=(z_1-a_{11}+i\gamma_{11})(z_2-a_{22}+i\gamma_{22}) -(t_{12}+a_{12}-i\gamma_{12})^2. \tag{19}$$

The coefficients $a_{ss'}$ are proportional to f^2 and are

$$a_{ss'}=V_{s0}V_{0s'}\varepsilon, \quad \text{where} \quad \varepsilon=E-E_0.$$

According to Eq. (3) the photoionization cross section (to within unimportant factors) is

$$\sigma_{\text{ion}}(\varepsilon)\propto\Sigma_p A^2_{i0}\frac{1}{\varepsilon^2}|T^{(p)}_{0p}(\varepsilon)|^2. \tag{20}$$

The summation in Eq. (20) extends over all possible final p states of the system (see Fig. 2). The ionization spectrum is characterized by the presence of alternating maxima (resonances) and minima. In addition, the position of the maxima is determined by the zeros of the real part Re $D=0$ of the determinant (19). The minima are a result of interference between the contributions of the direct and resonant transitions. Near the points $\nu_1=n_1\pm 1/2$ and $\nu_2=n_2\pm 1/2$, where $n_{1(2)}$ are integers, the nonadiabatic coupling between the series 1 and 2 is broken and photoionization with population of a given isolated series proceeds independently of the presence of the other series. To illustrate this, the ionization spectrum of the H_2 molecule in the interval $\nu_1=20-25$ is presented in Fig. 3 for two values of the field intensity f. The calculation was performed with the following values of the parameters of the multichannel quantum defect theory:[11,12]

$$\mu_{s\sigma}=-0.120, \quad \mu_{p\sigma}=0.191, \quad \mu_{d\sigma}=0.022,$$

$$\mu_{p\pi}=-0.078, \quad \mu_{d\pi}=\mu_{d\delta}=0$$

for external field frequency $\omega_f=0.029$ corresponding to the direct transition

$$4s\sigma H'\,{}^1\Sigma^+_g(\nu=0)\rightarrow 20p0\,{}^1\Sigma^+_u(\nu=0).$$

It is evident that the position of the minima is essentially independent of the field intensity. A much sharper dependence on f appears in the position and characteristics of the resonance peaks, which display the structure of the intermediate Rydberg states of the H_2 molecule. They are a superposition of the states $np0({}^1\Sigma^+_u)$ and $np2({}^1\Pi_u)$of the two-channel Fano system, which contains information about the characteristics of these states and their decay.

The quasistationary (in the presence of a field) levels of the system $e^-+H_2^+$ can be found from the equation $D=0$, finding successively first their position and then their width. The equation Re $D=0$, written in the form

$$(\tan(\pi\nu_1)-t_{11}-a_{11})(\tan(\pi\nu_2)-t_{22}-a_{22})=(t_{12}+a_{12})^2, \tag{21}$$

is formally identical to the two-component Fano equation[3] in the problem of the energy eigenvalues of the optical $R(0)$ branch of the H_2^{**} molecule under conditions of strong nonadiabatic electron–rotational coupling. In contrast to Ref. 3, Eq. (21) also includes the interaction characteristics (a_{11}, a_{22},a_{12}) that depend on the external electromagnetic field.

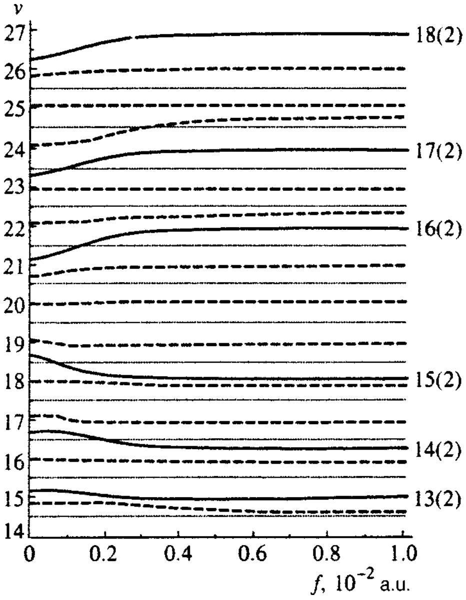

FIG. 4. Positions of the Rydberg levels in the intermediate series 1 and 2 of the optical $R(0)$ branch of the H_2 molecule as a function of the external field strength f which are roots of the secular equation (21). The broken lines show the levels of the series 1 for $N=0$; solid lines— series 2 for $N=2$, dots — limiting positions of the series-1 levels with quantum defects $\mu_{n1}(f)=1/2$.

For the region of the spectrum $|\varepsilon|\gg 1/n^3$, to determine the positions of the levels within the characteristic Rydberg ranges $\Delta\varepsilon\propto 1/n^3$ the variation of the interaction characteristics can be neglected. Stepwise analysis of the equation $D=0$ is made possible by the fact that the quantities $\gamma_{ss'}$ (s, $s'=0,2$) are quadratic in the small [by virtue of the condition (1)] interaction with the field, while in the matrix elements $a_{ss'}$ the presence of the energy denominator $|\varepsilon|\ll 1$ can compensate this smallness.

The dependence of the positions of the Rydberg levels in the series 1 and2 on the external field strength is displayed in Fig. 4. It is evident that in the range $f\sim 10^{-3}$–10^{-2} the field-induced perturbation of the Rydberg states is indeed large, since the change in the quantum defects of the levels $\mu_{ns}(f)=n(s)-\nu_s(f)$ is a measure of this perturbation. A substantial irregularity from level to level is observed in the $\mu_{ns}(f)$ dependence: Together with the strongly perturbed levels (such as 15(1), 16(1), 13(2), 14(2), and others) there are levels whose position remains almost unchanged (20(1), 23(1), and so on). In addition, the series 2 is perturbed much more strongly (the Coulomb scale factor $1/\nu_s^3$, which is much smaller for this series, comes into play here).

The structure of the spectrum of perturbed levels exhibits a variety of features. The first one is a repulsion of the entire collection of levels relative to the level 20(1) located at the center of the absorption line. At the same time, the resulting picture is supplemented by the existence of regions of ''strong'' convergence of individual pairs of levels. They occur in pairs 15(1)–13(2) and 18(1)–15(2) near $f\sim 10^{-3}$. Convergence is due to the field-induced interaction of a large group of states, though it occurs only when the splitting between the initial levels is small in the absence of a field (strong nonadiabatic mixing). It is important to emphasize that in the series 1 a term can cross with the limit of its Rydberg range

$$E_{\nu_1(\infty)}=-\frac{1}{2(n_1+1/2)^2}.$$

This is ruled out for the perturbation of Rydberg series with the general ionization potential. Such crossing for a positive-definite detuning $\varepsilon>0$ can occur if

$$E_{\nu_1(\infty)}<E_{\nu_2(\infty)} \qquad \text{and } 12B+\frac{1}{\nu_1(0)^2}<\frac{1}{\nu_2(0)^2},$$

where $\nu_s(0)$ are the effective principal quantum numbers of the series 1 and 2 and are roots of Eq. (21) for $f=0$. This condition essentially reduces to the requirement that in the absence of a field the corresponding level $E_{\nu_1(0)}$ of series 1, together with its limiting value $E_{\nu_1(\infty)}$, must be contained in the interval $[E_{\nu_2(0)};\ E_{\nu_2(\infty)}]$. An upward shift of the level E_{ν_2} with increasing external field strength f can cause the level E_{ν_1} to be expelled and to cross its own limiting value. The possibility of crossing is accidental, since this is determined by the specific values of the elements of the reaction matrix (12), which are individual characteristics of the H_2 molecule and do not depend on the field strength. In the range of n_1 and f presented in Fig. 4 this situation is observed only for the level $n_1=24$.

These features of the spectrum of intermediate Rydberg states of the H_2 molecule give rise to unusual behavior) (compared to atomic systems with a structureless core) of the autodecay widths $\Gamma_n(f)$, which are determined by the expression

$$\Gamma_\nu=\frac{2}{\pi}\left[\frac{\gamma_1(z_2-a_{22})+\gamma_2(z_1-a_{11})+2\gamma_{12}a_{12}}{\nu_1^3(z_2-a_{22})(1+z'^2_1)+\nu_2^3(z_1-a_{11})(1+z'^2_2)}\right] \tag{22}$$

(where $z'_s=\tan(\pi\nu_s)$, $s=1,2$) and are a functional of the positions of the levels in the Rydberg molecules XY**. The f dependence of the autoionization widths for Rydberg levels of atoms (excited due to V transitions from a lower-lying state) is directed related to the external field induced quantum defect $\mu_n(f)$ and has the simple form[8,9]

$$\Gamma_n(f)=\frac{2\gamma}{\pi n^3}\cos^2(\pi\mu_n(f))\simeq\frac{Af^2}{1+cf^4}, \tag{23}$$

where A and C are numerical factors. The presence of relatively extended sections where the level widths $\Gamma_n\propto 1/f^2$ decrease monotonically with increasing f, is clear evidence of field-induced stabilization. The most characteristic indicator here is the fact that the Rydberg levels occupy positions between the Coulomb levels (i.e., $\nu\rightarrow n\pm 1/2$ as $f\rightarrow\infty$). This dependence is typical for atoms with a structureless core and occurs for excitation of both one and two Rydberg series.[9]

In molecular systems the regularity in the arrangement of the levels is destroyed (as a result of the interaction of Rydberg series corresponding to different excited states of the ion core) and the Γ_ν dependence (f) can be very diverse. Specifically, the simple relation (23) between the width and

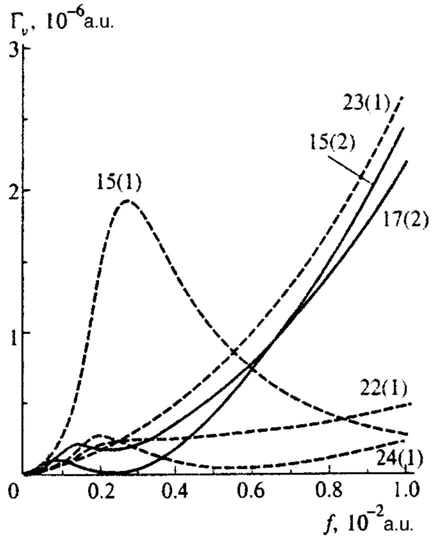

FIG. 5. Autoionization widths of the Rydberg levels of the optical series 1 and 2 versus the external field strength.

position of a level no longer holds. Here, just as in systems with few levels,[13–15] interference stabilization is determined not only by the relative arrangement of the levels but also by a special relation between the interaction parameters. In addition, a simple crossing of the limit of a Rydberg interval of one series ($\nu_1(\infty)=n_1\pm 1/2$) by a level belonging to a different series (for example, ν_2) does not result in stabilization. However, it is interesting to note that near a crossing point the decay characteristics of the term are determined by one of the quantities γ_1 or γ_2 (in the present case γ_2):

$$\Gamma_\nu=\frac{2\gamma_2}{\pi}\,\frac{1}{\nu_1^3(t_{12}+a_{12})^2+\nu_2^3(t_{22}+a_{22})^2}. \tag{24}$$

This result follows from Eq. (22), provided that the solution of Eq. (21) near this point, where $\tan(\pi\nu_1)=\infty$ and $\tan(\pi\nu_2)=t_{22}+a_{22}$, is used.

Most states corresponding to the term picture presented in Fig. 4 are not stabilized at all: Their widths increase monotonically with f (an example of such behavior is the form of the width of the level 23(1), presented in Fig. 5). Nonetheless some of them (for example, the states $\nu_1=24$, $\nu_2=14$, 15, 16, and 17) show a tendency toward a form of stabilization that appears in the presence of sections with a small decrease of Γ_f. However, subsequently this decrease is replaced by growth with increasing field strength, though the growth is more gradual than on the initial section of the curve (where $\Gamma\propto f^2$). Only the state 15(1) exhibits behavior similar to that described by Eq. (23). It can therefore asserted be that for moderate field strengths, so that in accordance with the condition (1) the parameter $\eta\ll 1$ and the field-dependent widths of the Rydberg levels are much less than the splitting between the levels, the stabilization effect in the form $\Gamma(f)\propto 1/f^2$ as $f\rightarrow\infty$ is essentially not observed.

This is explained by the fact that when the regularity in the arrangement of the levels as a function of n is destroyed, interference phenomena in large groups of Rydberg states are impeded. Therefore complete interference expulsion of an electron from the region of strong interaction with the ion core does not occur in molecules.

5. CONCLUSIONS

We have presented he foundations of the theory of multiphoton ionization of molecules taking account of strong perturbations introduced into the intermediate highly excited states by an external monochromatic field. Specifically, the large change in the structure of the Rydberg levels, the hybridization of interacting states, and the sharp dependence of the decay characteristics of the states on the level number were discussed. These phenomena are important for practical applications, since only optically allowed series participate in direct transitions in Rydberg states, while quite high field intensitis are required to study of multiphonon or cascade processes.

The photoionization of the H_2 molecule with the participation of the $np0$ and $np2$ series of the optical $R(0)$ branch is a striking illustration of strong nonadiabatic electron–rotational coupling. The field-induced perturbation of states with different rotational quantum numbers ($N=0$ and $N=2$) and, correspondingly, with two ionization potentials is described here by the transcendental equation (21), which is formally identical to the two-component Fano equation,[3] but contains additional interaction characteristics that depend on the external field. Essentially, this opens up the possibility of deliberate action, using an external field, on the structure of the spectrum of the excited states and control of the processes in which these states participate.

We also investigated the decay characteristics of these states and showed that the disruption of the regularity in the arrangement of the levels (as a result of the strong interaction of the Rydberg states corresponding to different ionization potentials) impedes the stabilization effect in Rydberg systems. Most states do not stabilize at all and are characterized, as usual, by level broadening quadratic in the field strength f ($\Gamma_n\propto f^2$). Nonetheless, for certain states a nonmonotonic dependence $\Gamma_n(f)$ is observed and is expressed in the fact that $\Gamma(f)\approx Af^2$ for small f, when the parameter $a_{ss'}\ll 1$. Subsequently, partial narrowing of the levels occurs and then changes to gradual growth $\Gamma_n(f)\approx Bf^2$ with coefficient of proportionality B much less than A. This $\Gamma_n(f)$ dependence is typical for the Rydberg states of molecules excited as a result of quasiresonant V transitions from a lower-lying level in quite moderate fields (so that $\Gamma_n\ll 1/n^3$). At the same time, it differs substantially from the corresponding behavior for the hydrogen atom or simple atoms with a structureless core

This work was supported by the Russian Fund for Fundamental Research (Grant No. 96-03-34113).

*) E-mail: golubkov@phch.ras.ru

1) Here and below we use the atomic system of units $\hbar=m_e=e=1$, in which the field intensity is $f_a=1$.

---

[1] H. Samble, Phys. Rev. A **7**, 2203 (1973).
[2] Ya. B. Zeldovich, Usp. Fiz. Nauk **110**, 139 (1973) [Sov. Phys. Usp. **16**, 427 (1973)].

[3] U. Fano, Phys. Rev. **2**, 353 (1970).
[4] M. Gavrila and J. Z. Kaminski, Phys. Rev. Lett. **52**, 613 (1984).
[5] R. M. Potvliege and R. Shakeshaft, Phys. Rev. A **40**, 306 (1985).
[6] Q. Su, H. Eberly, and J. Javanien, Phys. Rev. Lett. **64**, 862 (1990).
[7] A. V. Movsesyan and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **99**, 411 (1991) [Sov. Phys. JETP **72**, 228 (1991)].
[8] G. K. Ivanov and G. V. Golubkov, Zh. Éksp. Teor. Fiz. **99**, 1404 (1991) [Sov. Phys. JETP **72**, 783 (1991)].
[9] G. K. Ivanov, G. V. Golubkov, and D. M. Manakov, Zh. Éksp. Teor. Fiz. **106**, 1306 (1994) [JETP **79**, 707 (1994)].
[10] G. V. Golubkov and G. K. Ivanov, Zh. Éksp. Teor. Fiz. **80**, 1321 (1981) [Sov. Phys. JETP **53**, 674 (1981)].
[11] G. Herzberg and Ch. Jungen, J. Mol. Spectrosc. **41**, 425 (1972).
[12] K. Nakashima, H. Takagi, and H. Nakamura, J. Chem. Phys. **86**, 726 (1987).
[13] M. Aymar, J. Phys. B **18**, L763 (1985).
[14] H. Frederich and D. Wintgen, Phys. Rev. A **32**, 3231 (1985).
[15] N. P. Poluéktov and M. V. Fedorov, Zh. Éksp. Teor. Fiz. **114**, 821 (1998) [JETP **87**, 445 (1998)].

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Removal of single-atom optical bistability by quantum fluctuations

A. V. Kozlovskiĭ*) and A. N. Oraevskiĭ†)

P. N. Lebedev Physics Institute, Russian Academy of Sciences, 117924 Moscow, Russia

(Submitted 14 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 2001–2013 (June 1999)

The dynamics of the quantum-statistical properties of the radiation of an atom in a cavity interacting with an external coherent field is investigated. A high level of quantum fluctuations of the field in the cavity is shown to destroy optical bistability in the multiatom case. Photon-number fluctuations and the spectral dispersion of the canonically conjugate quadratures of the field inside and outside a cavity are calculated. It is found that in contrast to the multiatom case quadrature squeezing and squeezing of the radiation intensity of a single atom are negligible inside and outside the cavity. © *1999 American Institute of Physics.* [S1063-7761(99)00706-4]

1. INTRODUCTION

The radiation of an active medium consisting of two-level atoms in an optical cavity in an external coherent field exhibits the property of bistability for certain ratios of the dissipation rates, frequency detuning, and external field strength. The microscopic quantum theory of this phenomenon shows that under bistability conditions the radiation can be in a nonclassical quadrature-squeezed state.[1–12] It has been found[1–3] that in the multiatom case ($N_{\rm at}\gg 1$) squeezing inside a cavity can reach 50% on the lower stable branch. Analysis of the squeezing of radiation leaving the cavity showed that it can reach 90%.[4–12]

The phenomenon of bistability for one atom in a cavity has been investigated in Refs. 13 and 14. The Fokker–Planck equation obtained in these works for the field and atomic variables on the basis of a linear approximation in the field fluctuations made it possible to conclude that bistability is present if the rate γ of the cavity losses is much greater than the spontaneous decay rate Γ of the levels, $\gamma\gg\Gamma$, and that bistability is absent for $\gamma<\Gamma$. The presence of fluctuations near the turning points in the latter case destroys optical bistability in the weak-fluctuation case.[15]

In the present paper we analyze the dynamics of the quantum-statistical properties of the radiation of an atom in the optical bistability region. In the reduced density operator method for the ''atom + field'' system in a basis consisting of the Fock states of the field, the theory employed is valid for arbitrary ratios of the parameters characterizing the coherent pumping and dissipation processes, simulated by appropriate reservoirs interacting with the atom and the radiation field. We note that our investigations of the dynamics of the field and of the statistical properties of the radiation do not employ the small-fluctuation approximation, in which ordinarily an approximation linear in the fluctuations is used. The statistical properties of the field inside and outside the cavity are analyzed.

2. MODEL OF OPTICAL BISTABILITY

The Hamiltonian of the ''atom + field + reservoir'' system includes the energy H_F of a discrete mode of the electromagnetic field inside the cavity, the energy H_A of the two-level atom, the atom–field interaction V_{A-F}, the energy R_F of the reservoir of the continuous spectrum of the thermal modes of the field (the thermostat), the interaction V_{A-R} of the atom with the thermostat reservoir, the interaction V_{F-R} of the cavity field and the thermostat, and the interaction V_{F-I} of a field oscillator inside the cavity with an external coherent field $\mathscr{E}=Ee^{-i\omega_0 t}$, $E=\sqrt{n_I}e^{i\phi I}$ (Refs. 1–15):

$$\begin{aligned}\frac{H}{\hbar}&=\frac{1}{\hbar}(H_F+H_A+V_{A-F}+R_F+V_{A-R}+V_{F-R}+V_{F-I})\\&=\omega_c a^+a+\frac{1}{2}\omega_A\sigma^z+g(a^+\sigma^-+\sigma^+a)+\sum_{j=1}^{\infty}\omega_j b_j^+b_j\\&\quad+\sum_{j=1}^{\infty}g_j(b_j^+\sigma^-+b_j\sigma^+)+\sum_{j=1}^{\infty}k_j(a^+b_j+b_j^+a)\\&\quad+(a^+\mathscr{E}+\mathscr{E}a).\end{aligned}\tag{1}$$

Here $a(a^+)$ are annihilation (creation) operators for the electromagnetic field of a discrete cavity mode with frequency ω_c and $b_j(b_j^+)$ are annihilation (creation) operators of the reservoir of the continuous spectrum of the thermal modes of the field which are present inside the cavity due to the partial transparency of the mirrors. These operators satisfy commutation relations for Bose particles:

$$\begin{aligned}&[a,a^+]=1,\quad [a^+,a^+]=[a,a]=0,\\&[b_j,b_i^+]=\delta_{ji},\quad [b_j^+,b_i^+]=[b_j,b_i]=0.\end{aligned}\tag{2}$$

In Eq. (1) the interaction of the laser radiation field with the reservoir modes $\{b_j\}$, which is responsible for establishing thermodynamic equilibrium between the cavity modes and the reservoir–thermostat field, which enters through the mirrors, with an average number of photons per mode[16]

$$n_T\equiv\overline{n}(\omega=\omega_c,T)=[\exp(\hbar\omega_c/kT)-1]^{-1}.\tag{3}$$

The atomic polarization operators (σ^-,σ^+) and the inversion operator σ^z of the populations of the two-level atom,

$$\sigma^- = |\downarrow\rangle\langle\uparrow|, \quad \sigma^+ = |\uparrow\rangle\langle\downarrow|, \quad \text{and } \sigma^z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|, \tag{4}$$

satisfy

$$2\sigma^\mp\sigma^\pm = 1 \mp \sigma^z \quad \text{and} \quad \sigma^z\sigma^z = 1 \tag{5}$$

and the commutation relations for Fermi particles

$$\{\sigma^+,\sigma^-\} = |\uparrow\rangle\langle\uparrow| + |\downarrow\rangle\langle\downarrow| = 1,$$
$$\{\sigma^+,\sigma^+\} = \{\sigma^-,\sigma^-\} = 0. \tag{6}$$

In Eqs. (4) and (6) $|\downarrow\rangle$ and $|\uparrow\rangle$ are the lower and upper energy states of the active electron and $\{\,\ldots\}$is an anticommutator.

The interaction between an atomic electron and the field $\{b_j\}$ of the harmonic oscillators in the thermostat establishes thermal equilibrium of the quantum average atomic operators with temperature *T*.

The interaction between the radiation field and the polarization of an atom is proportional to the interaction constant *g*, characterized by the transition dipole moment $d_{\uparrow\downarrow}$ and the cavity volume *V*,

$$g = d_{\uparrow\downarrow}\sqrt{\frac{2\pi\omega_A}{\hbar V}}, \tag{7}$$

where ω_A is the atomic transition frequency. The constants $g_j$of the interaction between the thermostat field and an atomic electron are found similarly. The parameter k_j appearing in the operator V_{F-R} is a constant characterizing the interaction between the thermostat modes and a cavity mode.

We shall use the reduced density operator of the ''atom + single-mode field'' system in the basis of Fock states:

$$\rho(t) = \sum_{i,j=\{\uparrow,\downarrow\}}\ \sum_{n,m=0}^{\infty} \rho_{n,i;m,j}(t)|i\rangle|n\rangle\langle m|\langle j| \tag{8}$$

to analyze the quantum stochastic dynamics of a two-level single-atom laser. In the interaction representation and the Born–Markov approximation[16] the reduced density operator (8) of the ''atom + field'' system interacting with a reservoir satisfies the Liouville equation

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \mathrm{Tr}_R(\sigma_{R\oplus S}) = -i\frac{\Delta_A}{2}[\sigma^z,\rho] - i\Delta_C[a^+a,\rho] \\ &\quad - ig[(a^+\sigma^- + \sigma^+a),\rho] + \frac{\gamma}{2}(n_T+1)(2a\rho a^+ \\ &\quad - a^+a\rho - \rho a^+a) + \frac{\gamma}{2}n_T(2a^+\rho a - aa^+\rho - \rho aa^+) \\ &\quad + \frac{\Gamma}{2}(N_T+1)(2\sigma^-\rho\sigma^+ - \sigma^+\sigma^-\rho - \rho\sigma^+\sigma^-) \\ &\quad + \frac{\Gamma}{2}N_T(2\sigma^+\rho\sigma^- - \sigma^-\sigma^+\rho - \rho\sigma^-\sigma^+) - i[E(a^+\rho \\ &\quad - \rho a^+) + E^*(a\rho - \rho a)],\end{aligned} \tag{9}$$

where the detuning of the frequency ω_0 of the external field from the atomic frequency ω_A and from the oscillator frequency ω_c is $\Delta_A = \omega_A - \omega_0$ and $\Delta_c = \omega_c - \omega_0$, and γ and Γ are, respectively, the field-loss rate at the mirrors and the spontaneous-emission rate. These dissipation constants can be expressed in terms of the correlation functions of the corresponding reservoir operators $\{b_j\}$ (see, for example, Ref. 16).

The average number N_T of excitations of the atomic reservoir that appears in Eq. (9) for the case where the reservoir modes $\{b_j\}$ are in thermodynamic equilibrium at temperature *T* is (see, for example, Ref.16):

$$\begin{aligned}N_T &\equiv \overline{N}(\omega=\omega_A,T) = \langle b_j^+ b_j\rangle_{\omega_j=\omega_A} \\ &= \left[\exp\left(\frac{\hbar\omega_A}{kT}\right) - 1\right]^{-1}.\end{aligned} \tag{10}$$

Since at optical frequencies we have $\hbar\omega_A$, $\hbar\omega_c \gg kT$, we shall neglect n_T and N_T in the calculations below.

Using the correspondence between quantum-mechanical operators and *c*-number variables, a generalized Fokker–Planck equation for the quasiprobability distribution function[3,13,14,17] corresponding to the density operator can be obtained from the Liouville equation (9). In turn, the Fokker–Planck equation so obtained is equivalent to a system of Langevin stochastic differential equations of motion for the *c*-number field and atomic variables. The solution of this system of equations for stationary conditions makes it possible to obtain a relation between the external field intensity and the quantum-mechanical average intensity of the field inside the cavity (the equation of state). As a result of the nonlinearity of the atom–field interaction in the two-level model of an atom the stationary equation of state has the form

$$\begin{aligned}Y^2 &\equiv \left|\frac{2E}{\gamma}\right|^2\frac{1}{n_s} = X^2\left[\left(1 + \frac{2C}{1+X^2+\delta_A^2}\right)^2\right. \\ &\quad \left. + \left(\delta_c - \frac{2C\delta_A}{1+X^2+\delta_A^2}\right)^2\right],\end{aligned} \tag{11}$$

$$X^2 \equiv \frac{\langle a^+a\rangle}{n_s}, \quad \delta_A = 2\frac{\omega_A - \omega_0}{\Gamma}, \quad \delta_c = 2\frac{\omega_c - \omega_0}{\gamma},$$

where the cooperativity parameter *C* and the number n_s of saturation photons have been introduced as

$$C = \frac{2g^2}{\Gamma\gamma} \quad \text{and} \quad n_s = \frac{\Gamma^2}{8g^2}. \tag{12}$$

The relation (11) can also be obtained in a deterministic theory using the Maxwell–Bloch equations.[18] Analysis of the equation of state (11) shows that absorptive bistability is present if $C>4$ and $\delta_A = \delta_c = 0$. Absorptive–dispersive bistability is present if $C^2 > 27\delta_A^2/4$ ($C \gg 1$), $\delta_A \neq 0$, and $\delta_c = 0$ or if $C^2 > 27\delta_c^2/4$, $\delta_c \neq 0$, and $\delta_A = 0$ and also if $C > 4|\delta_c\delta_A|$ ($C \gg 1$) for $\delta_c\delta_A < 0$ or $C > 0.5\delta_c\delta_A$ ($C \gg 1$) for $\delta_c\delta_A > 0$.

The following system of coupled differential equations for the density matrix elements follows from the Liouville equation (9):[19,20]

$$\rho_{1n,m}(t) \equiv \langle\uparrow|\rho_{n,m}|\uparrow\rangle + \langle\downarrow|\rho_{n,m}|\downarrow\rangle,$$

$$\rho_{2n,m}(t)\equiv\langle\uparrow|\rho_{n,m}|\uparrow\rangle-\langle\downarrow|\rho_{n,m}|\downarrow\rangle, \tag{13}$$

$$\rho_{3n,m}(t)\equiv\langle\uparrow|\rho_{n,m}|\downarrow\rangle, \quad \rho_{4n,m}(t)\equiv\langle\downarrow|\rho_{n,m}|\uparrow\rangle.$$

Using the following relations for the creation–annihilation operators in the Fock basis

$$a|n\rangle=\sqrt{n}\,|n-1\rangle, \quad a^{+}|n\rangle=\sqrt{n+1}\,|n+1\rangle, \tag{14}$$

$$\langle n|a=\sqrt{n+1}\,\langle n+1|, \quad \langle n|a^{+}=\sqrt{n}\,\langle n-1|,$$

and

$$\sigma^{z}\sigma^{+}=\sigma^{+}, \quad \sigma^{+}\sigma^{z}=-\sigma^{+}, \quad \sigma^{-}\sigma^{z}=\sigma^{-}, \tag{15}$$

as well as Eqs. (2) and (4)–(7), we find directly from the Liouville equation (9) the equation of motion for the elements of the form (13) of the density matrix (8). The result is

$$\dot{\rho}_{1,n,m}=ig(\sqrt{m+1}\,\rho_{3,n,m+1}-\sqrt{n}\,\rho_{3,n-1,m}+\sqrt{m}\,\rho_{4,n,m-1}$$

$$-\sqrt{n+1}\,\rho_{4,n+1,m})+\frac{\gamma}{2}L_{\gamma}\rho_{1}+L_{E}\rho_{1},$$

$$\dot{\rho}_{2,n,m}=ig(\sqrt{m+1}\,\rho_{3,n,m+1}+\sqrt{n}\,\rho_{3,n-1,m}-\sqrt{m}\,\rho_{4,n,m-1}$$

$$-\sqrt{n+1}\,\rho_{4,n+1,m})+\frac{\gamma}{2}L_{\gamma}\rho_{2}-\Gamma\rho_{1,n,m}-\Gamma(2N_{T}$$

$$+1)\rho_{2,n,m}+L_{E}\rho_{2},$$

$$\dot{\rho}_{3,n,m}=i\Delta\rho_{3,n,m}+i\frac{g}{2}(\sqrt{m}\,\rho_{1,n,m-1}-\sqrt{n+1}\,\rho_{1,n+1,m}$$

$$+\sqrt{m}\,\rho_{2,n,m-1}+\sqrt{n+1}\,\rho_{2,n+1,m})+\frac{\gamma}{2}L_{\gamma}\rho_{3}$$

$$+L_{E}\rho_{3}-\frac{\Gamma(2N_{T}+1)}{2}\rho_{3,n,m},$$

$$\dot{\rho}_{4,n,m}=i\Delta\rho_{4,n,m}+i\frac{g}{2}(\sqrt{m+1}\,\rho_{1,n,m+1}-\sqrt{n}\,\rho_{1,n-1,m}$$

$$-\sqrt{m+1}\,\rho_{2,n,m+1}-\sqrt{n}\,\rho_{2,n-1,m})+\frac{\gamma}{2}L_{\gamma}\rho_{4}$$

$$+L_{E}\rho_{4}-\frac{\Gamma(2N_{T}+1)}{2}\rho_{4,n,m}, \tag{16}$$

where the terms that are common to all four equations and are due to cavity losses through the mirrors and the interaction of the cavity field with the external coherent field can be written in the general form

$$L_{\gamma}\rho_{j}=(n_{T}+1)[2\sqrt{(n+1)(m+1)}\,\rho_{j,n+1,m+1}-(m$$

$$+n)\rho_{j,n,m}]+n_{T}[2\sqrt{nm}\,\rho_{j,n-1,m-1}-\rho_{j,n,m}(n+m$$

$$+2)], \quad j=1,2,3,4, \tag{17}$$

$$L_{E}\rho_{j}=-i\{[E\sqrt{n}\,\rho_{j,n-1,m}-\sqrt{m+1}\,\rho_{j,n,m+1}]$$

$$+E^{*}[\sqrt{n+1}\,\rho_{j,n+1,m}-\sqrt{m}\,\rho_{j,n,m-1}]\}$$

$$-i\Delta_{c}(n-m)\rho_{j,n,m}, \quad j=1,2,3,4. \tag{18}$$

We solved the system of equations (16) of dimension $4(n_{\max}+1)(n_{\max}+1)$ numerically using the fourth-order Runge–Kutta method ($n_{\max}$ is the size of the Fock basis). In the general case the field was initially in an arbitrary mixed state and the atom was in the lower state. Therefore the density matrix of the noninteracting atom and field at time $t=0$ is

$$\rho(0)=\rho_{a}\otimes\rho_{f}, \tag{19}$$

where

$$\rho_{a}=|\downarrow\rangle\langle\downarrow|, \quad \rho_{f}=\sum_{n,m=0}^{\infty}c_{n}c_{m}^{*}|n\rangle\langle m|.$$

Specifically, if the field is initially in a coherent state $|\alpha\rangle$, then the coefficients in the expansion in terms of the Fock states are

$$c_{n}^{*}c_{m}=\frac{\alpha^{*n}\alpha^{m}}{\sqrt{n!m!}}\exp(-|\alpha|^{2}). \tag{20}$$

If the initial state is a pure Fock state $|n_{0}\rangle$, then $c_{n}^{*}c_{m}=\delta_{n,m}$. For a thermal state we have $c_{n}^{*}c_{m}=[n_{T}^{n}/(1+n_{T})^{n+1}]\delta_{n,m}$.

The average number of photons, the average inversion, the fluctuations (variance) of the photons, and the average field can be found, using a grouping of the form (13) for the matrix elements of the system, as

$$\langle n(t)\rangle=\mathrm{Tr}[\rho(t)a^{+}a]=\sum_{n=0}^{\infty}n\rho_{1n,n}(t), \tag{21}$$

$$\langle D(t)\rangle=\mathrm{Tr}[\rho(t)\sigma^{z}]=\sum_{n=0}^{\infty}\rho_{2n,n}(t), \tag{22}$$

$$\mathrm{Var}(n(t))\equiv\langle(\Delta n(t))^{2}\rangle=\sum_{n=0}^{\infty}(n-\langle n(t)\rangle)^{2}\rho_{1n,n}(t), \tag{23}$$

$$\langle a^{+}(t)\rangle=\sum_{n=0}^{\infty}\sqrt{n+1}\,\rho_{1n,n+1}(t),$$

$$\langle a(t)\rangle=\sum_{n=1}^{\infty}\sqrt{n}\,\rho_{1n,n-1}(t). \tag{24}$$

The variances of the conjugate quadratures $X_{+}(t)=[a^{+}(t)+a(t)]/2$ and $X_{-}(t)=[a^{+}(t)-a(t)]/2i$ can be expressed in terms of the matrix elements of the density operator as

$$\langle(\Delta X_{\pm})^{2}\rangle=\frac{1}{4}\left\{\sum_{n=0}^{\infty}(2n+1)\rho_{1n,n}(t)\right.$$

$$\pm\sum_{n=2}^{\infty}\sqrt{n(n-1)}\,\rho_{1n,n-2}(t)$$

$$\pm\sum_{n=0}^{\infty}\sqrt{(n+1)(n+2)}\rho_{1n,n+2}(t)$$

$$\mp\left[\sum_{n=0}^{\infty}\sqrt{n+1}\,\rho_{1n,n+1}(t)\right.$$

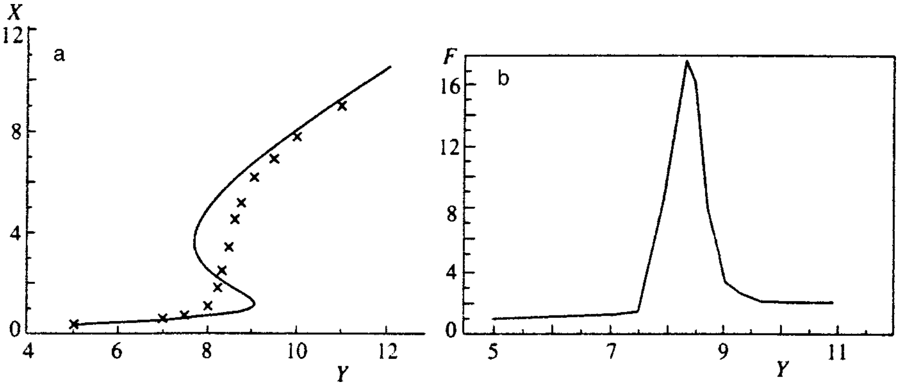

FIG. 1. a) Dimensionless field X inside a cavity versus the dimensionless amplitude Y of the external field for the absorptive bistability: solid curve — semiclassical equation of state (11); crosses — quantum-mechanical calculation. $C=8$, $n_s=1$, $\delta_c=\delta_A=0$. b) Fano factor $F=\mathrm{Var}(n)/\langle n\rangle$ for the number of in-cavity photons versus the dimensionless external field.

$$\pm\sum_{n=1}^{\infty}\sqrt{n}\,\rho_{1n,n-1}(t)\Bigg]^2\Bigg\}. \tag{25}$$

3. OPTICAL BISTABILITY AND CAVITY FLUCTUATIONS

We performed numerical calculations of the dynamics of the quantities (21)–(25) to investigate the effect of quantum fluctuations on the occurrence of optical bistability of the radiation of a single atom. We calculated the cases of a high-Q cavity ($\gamma \ll \Gamma$) and a high-loss cavity ($\gamma \gg \Gamma$) for the absorptive and absorptive–dispersive bistabilities. A coherent-vacuum state and a coherent pure state were used as the initial state of the field at $t=0$. A numerical investigation of the system (16) for various values of the parameters of the atom, the external field, and the reservoir showed that for the stationary radiation of an atom both forms of optical bistability predicted by the semiclassical theory[1–12] are absent. The dependence $X(Y)$ (Fig. 1) obtained for the purely absorptive bistability in the semiclassical theory [the expression (11)] in the region of bistability between the upper ($tp2$) and lower ($tp1$) turning points, $Y_{tp2}<Y<Y_{tp1}$, differs sharply from the results of our quantum calculation. In this region the stationary solution $X(Y)$ does not depend on the initial conditions, i.e., there is no hysteresis, while the theory neglecting quantum fluctuations (or taking such fluctuations into account in the linear approximation[13,14]) predicts bistability and hysteresis, i.e., a dependence of the stationary state of the field on the initial conditions. The dependence of the Fano factor $F=\mathrm{Var}(n)/\langle n\rangle$ in the cavity on the external field strength Y is displayed in Fig. 1b. It follows from Fig.1b that a sharp spike of photon-number fluctuations which destroys bistability occurs in the bistability region.

The data presented in Fig. 1a can be compared with the computational results obtained in Ref. 14 in the adiabatic approximation in the atomic variables and in the linear approximation in the fluctuations for the same values of the parameters C and n_s (see Fig. 2a from Ref. 14). The comparison shows that the small-fluctuation approximation is inapplicable for the single-atom optical bistability, at least in the region $Y_{tp2}<Y<Y_{tp1}$ where the fluctuations are very large (see Fig. 1b). It is also found that much more time is required to establish a stationary state in the region of semiclassical bistability ($gt\sim 3000$) than outside this region ($gt\sim 100$). Figure 2 shows the average stationary number n of photons versus Y for the absorptive bistability for a larger value of the cooperativity parameter C and a smaller number n_s of saturation photons. Just as in the preceding case, the Fano factor (Fig. 2b) increases sharply in the region of semiclassical bistability, and hysteresis is completely absent.

The photon number distribution function $P(n)=\rho_{1n,n}$ in the region of semiclassical bistability is of special interest. It is evident from Fig. 3 that the photon distribution function possesses two maxima, i.e., the state of the field for $Y_{tp2}<Y<Y_{tp1}$ is a superposition of two states corresponding to the lower and upper stable branches of the semiclassical equation of state. Each peak in the distribution function can be approximated well by a Gaussian function, i.e., the quantum state of the field is a superposition of two coherent states: $|\psi\rangle=c_1|\alpha_1\rangle+c_2|\alpha_2\rangle$.

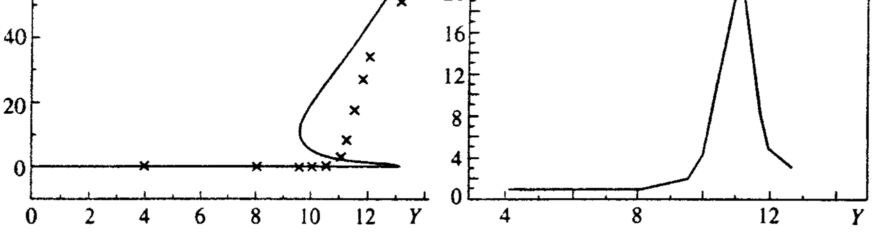

FIG. 2. a) Average stationary number n of photons versus the dimensionless external field Y for $C=12$ and $n_s=0.5$. b)In-cavity Fano factor under the same conditions.

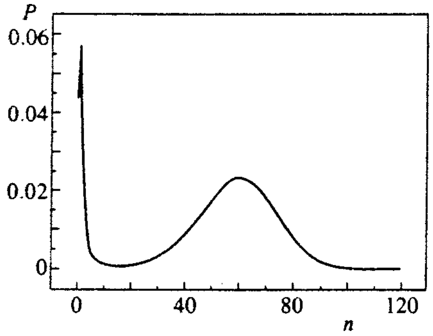

FIG. 3. Photon distribution function for the absorptive bistability with $C=80$, $n_s=8$, and $Y=8.5$.

Figure 4a shows $n(Y)$ for absorptive–dispersive bistability ($\delta_c\delta_A>0$ for $C\gg 1$). Just as for the purely absorptive bistability, quantum fluctuations completely remove bistability in the one-atom case under study. It is evident from Fig. 4a that the small-fluctuation approximation also breaks down outside the region of bistability of the semiclassical curve: The results of the exact calculation differ substantially from the semiclassical data for $Y>Y_{tp1}$ also. Almost linear growth of the Fano factor occurs for large Y (Fig. 4b). The computational results for the same value of the cooperativity parameter C and detuning δ_c and δ_A but for small n_s are presented in Fig. 5. This case corresponds to a large ratio $\gamma/\Gamma>1$, i.e., high cavity losses. Just as in the high-Q case, quantum fluctuations remove the bistability in the single-atom case.

Our calculations showed that for parameters C, δ_c, and δ_A for which the greatest squeezing of the field quadratures outside the cavity was found in Ref. 4 for the multiatom case, in our single-atom case quadrature squeezing inside the cavity is negligible (less than 10%) and is present only for small $Y<Y_{tp2}$. The Fano factor for small Yalso differs just as little from the shot-noise level. Increasing the number n_s of saturation photons for the same value of C, leading to an increase in the average number of photons in the cavity, does not appreciably increase the squeezing.

Figure 6a shows $n(Y)$ for the case where semiclassical bistability is absent. In this case the results of exact calculations are essentially identical to the semiclassical results. As follows from Fig. 6b, the peak of the fluctuations that arises in the region of bistability is absent here. The Fano factor increases monotonically with Y, and the small-fluctuation approximation likewise breaks down for large Y (see Fig. 6a).

4. FLUCTUATIONS IN THE QUADRATURE OF THE FIELD AT THE CAVITY EXIT

We assume that inside the laser cavity the electromagnetic field is in a state with discrete frequencies (photon energies), while outside the cavity the field possesses a continuous spectrum. As a result, it must be assumed that the temporal fluctuations of the field inside the cavity are sources of fluctuations of the frequency spectrum of the radiation exiting through the cavity mirror. The field outside the cavity can be represented as a sum of the laser radiation field exiting through the mirrors and the reservoir–thermostat noise field incident on the mirror, i.e.,[21–23]

$$a^{\mathrm{out}}(t)=b^{\mathrm{in}}(t)+\sqrt{\gamma}\,a(t),$$

$$b^{\mathrm{in}}(t)\propto\sum_j\sqrt{\omega_j}\exp(-i\omega_j t)b_j.$$

The Heisenberg operator $a^{\mathrm{out}+}(t)a^{\mathrm{out}}(t)$ is the operator of the number of photons exiting through the mirror per unit time. The commutation relations for the field operators forming a continuous spectrum outside the cavity (out) have the form[21–23]

$$[a^{\mathrm{out}}(t+\tau),a^{\mathrm{out}+}(t)]=\delta(\tau). \tag{26}$$

We introduce the canonically conjugate phase quadratures of the field outside the cavity, $X_+^{\mathrm{out}}(\theta)$ and $X_-^{\mathrm{out}}(\theta)$, for a local oscillator of arbitrary phase θ:

$$a^{\mathrm{out}}=e^{-i\theta}[X_+^{\mathrm{out}}(\theta)+iX_-^{\mathrm{out}}(\theta)],$$

$$a^{\mathrm{out}}=e^{i\theta}[X_+^{\mathrm{out}}(\theta)-iX_-^{\mathrm{out}}(\theta)],\quad [X_+^{\mathrm{out}}(\theta),X_-^{\mathrm{out}}(\theta)]=\frac{i}{2}. \tag{27}$$

For a single transmitting mirror the correlation functions of the field of a discrete radiation mode inside the cavity are related to the correlation functions of the fields of the continuous spectrum outside the cavity have the form[20–22]

$$\langle a^+(t+\tau)a(t)\rangle^{\mathrm{out}}=\gamma\langle a^+(t+\tau)a(t)\rangle. \tag{28}$$

The quantity characterizing the statistics of radiation passing through the exit mirror is the stationary spectrum of fluctuations[4–12]

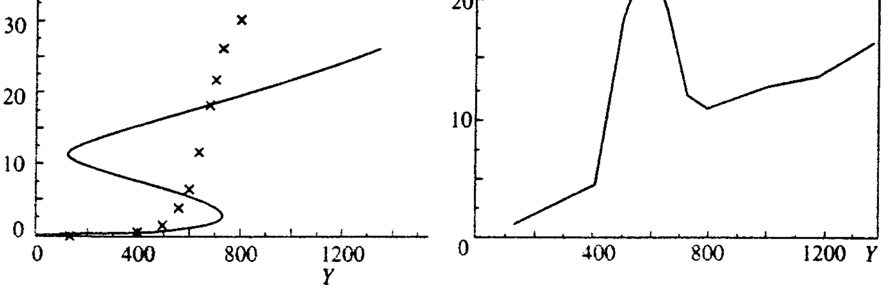

FIG. 4. Same as Fig. 2 for absorptive–dispersive bistability with $C=2200$, $n_s=0.001$, $\delta_A=100$, and $\delta_c=20.6$.

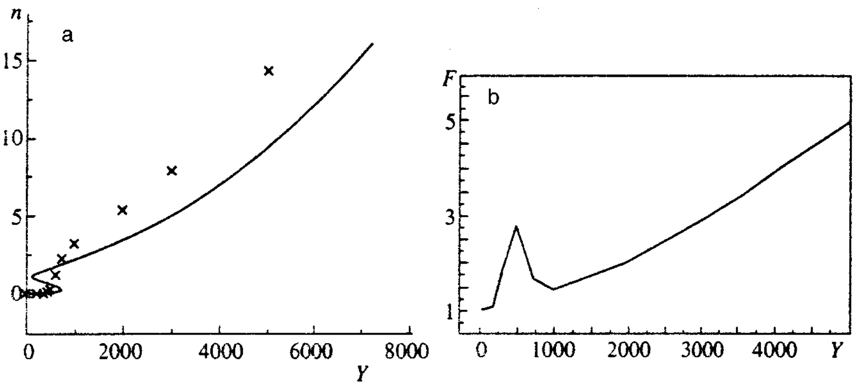

FIG. 5. Same as Fig. 4 with $n_s = 0.0001$.

$$V_{\pm}^{\text{out}}(\omega,\theta)=\lim_{t\to\infty}\int_{-\infty}^{\infty}d\tau\, e^{-i\omega\tau}[\langle X_{\pm}^{\text{out}}(t+\tau)X_{\pm}^{\text{out}}(t)\rangle - \langle X_{\pm}^{\text{out}}(t+\tau)\rangle\langle X_{\pm}^{\text{out}}(t)\rangle]. \qquad (29)$$

The fluctuations (spectral dispersion) (29) outside the cavity can be calculated using the in-cavity field operators[21–23] and the relation (28):

$$V_{\pm}^{\text{out}}(\omega,\theta)=\frac{1}{4}+\frac{\gamma}{4}[S_{12}(\omega)+S_{12}(-\omega) \pm 2(e^{2i\theta}S_{11}(\omega)+\text{c.c.})], \qquad (30)$$

where

$$S_{12}(\omega)=\int_{-\infty}^{\infty}d\tau\, e^{-i\omega\tau}[\langle a^{+}(t+\tau)a(t)\rangle - \langle a^{+}(t+\tau)\rangle\langle a(t)\rangle]=S_{21}(-\omega), \qquad (31a)$$

$$S_{11}(\omega)=\int_{-\infty}^{\infty}d\tau\, e^{-i\omega\tau}[\langle a^{+}(t+\tau)a^{+}(t)\rangle - \langle a^{+}(t+\tau)\rangle\langle a^{+}(t)\rangle]=S_{22}^{*}(\omega). \qquad (31b)$$

Since under stationary conditions the two-time correlation functions are even functions of t, the cosine Fourier transform was used in Eq. (31) for calculations.

As follows from Eq. (30), the spectrum of the photon-number fluctuations in the field at the cavity exit consists of a shot-noise term [the first term in Eq. (30)] and the normally and chronologically ordered fluctuation spectrum [second term in Eq. (30)]. Maximum squeezing $V(\omega,\theta)<1/4$ occurs for the optimal phase of a local oscillator for each frequency:

$$e^{-2i\theta}=\frac{S_{11}(\omega)}{|S_{11}(\omega)|}.$$

Then

$$V^{\text{out}}(\omega)=\frac{1}{4}+\frac{\gamma}{4}[S_{12}(\omega)+S_{12}(-\omega)-2|S_{11}(\omega)|]. \qquad (32)$$

Under stationary conditions it is easy to obtain the following expression, which follows from the quantum regression theorem (see, for example, Ref. 16), for the correlation functions of the cavity field operators:

$$\langle a^{+}(t+\tau)a(t)\rangle_{ss}=\text{Tr}[a^{+}\tilde{\rho}(\tau)], \qquad (33a)$$

$$\langle a(t+\tau)a(t)\rangle_{ss}=\text{Tr}[a\tilde{\rho}(\tau)], \qquad (33b)$$

where the operator $\tilde{\rho}(\tau)\equiv\tilde{\rho}(t+\tau)$ satisfies the Liouville equation (9) with the initial $(t=0)$ condition

$$\tilde{\rho}_{j,n,m}(0)=\sqrt{n+1}\,\rho_{j,n+1,m}(t_{ss}), \quad j=1,2,3,4. \qquad (34)$$

As shown in Ref. 4 for the absorptive bistability, appreciable squeezing of the field quadratures does not occur outside the cavity for $\gamma\gg\Gamma$. For absorptive–dispersive bistability substantial squeezing occurs if $\delta_c\delta_A>0$ on the lower stable branch. The optimal conditions, obtained in Ref. 4, for squeezing occurring for large detuning and small $X=\sqrt{\langle n\rangle/n_s}$ are $X^2/\delta_A^2\ll 1$ and $X^4/\delta_A^3\ll 1$. To compare with the quadrature squeezing outside the cavity in the multiatom case[4] we performed calculations of the stationary fluctuation spectrum for $C=1100$, $\delta_A=100$, and $\delta_c=20.6$, for which multiple squeezing was found in Ref. 4 in a wide range of values of $X(Y)$. The calculations showed that in contrast to

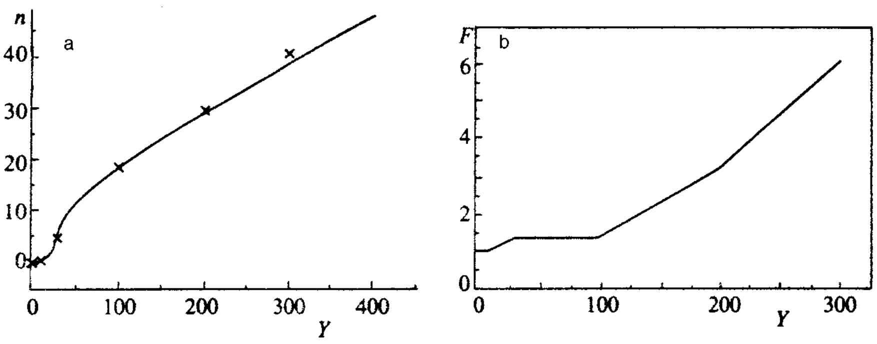

FIG. 6. Same as Fig. 4 but with $C=1100$ and $n_s=0.01$.

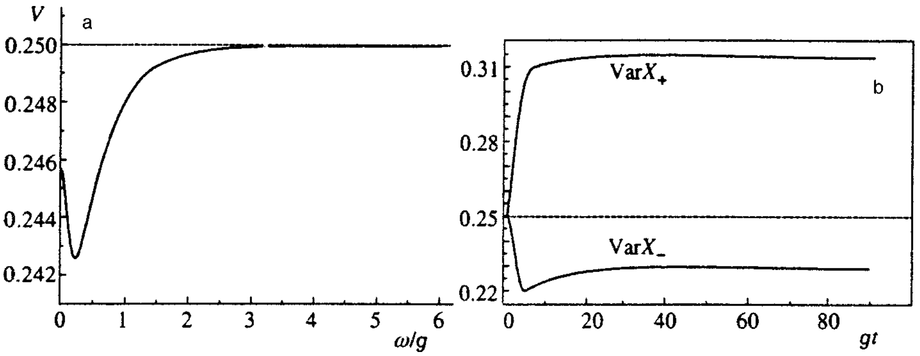

FIG. 7. a) Spectrum of squeezing of laser radiation outside the cavity for $C=8$ $n_s=1$, $Y=7$, and $\delta_c=\delta_A=0$. b)Time dependences of the variances of the canonically conjugate quadratures X_+ and X_-.

the multiatom case the degree of squeezing in the single-atom case is extremely low and is only several percent for small Y (Fig. 7). As Y increases, squeezing vanishes for both the absorptive and absorptive–dispersive cases.

5. CONCLUSIONS

We have analyzed the dynamics of the quantum-statistical properties of the radiation of an atom in an external field without using adiabatic elimination of atomic variables and without assuming quantum fluctuations to be small. We have shown that for a high-Q cavity as well as for large cavity losses bistability in the single-atom case is destroyed by large quantum fluctuations of the field. A sharp spike of the quantum photon-number fluctuations is present in the region of bistability (hysteresis region).

It was shown that the field in the region of the semiclassical bistability is in a quantum state consisting of a superposition of two coherent states corresponding to the upper and lower stable branches of the semiclassical equation of state.

It was found that in contrast to the multiatom case, where substantial quadrature squeezing of the field occurs inside and outside the cavity, in the single-atom case squeezing is negligible under the same conditions.

It was shown that the small-fluctuation approximation is inapplicable under optical bistability conditions, both in the hysteresis region and on the upper stable branch of the semiclassical curve of the state corresponding to large values of the external field.

*⁾E-mail: kozlovsky@neur.lpi.msk.su

†⁾E-mail: oraevsky@sci.lebedev.ru

---

[1] G. J. Milburn and D. F. Walls, Opt. Commun. **39**, 401 (1981).

[2] L. A. Lugiato and G. Strini, Opt. Commun. **41**, 67 (1982).

[3] P. D. Drummond and D. F. Walls, Phys. Rev. A **23**, 2563 (1981).

[4] M. D. Reid and D. F. Walls, Phys. Rev. A **32**, 396 (1985).

[5] C. M. Savage and H. J. Carmichael, IEEE J. Quantum Electron. **QE-24**, 1495 (1988).

[6] L. A. Lugiato, Phys. Rev. A **33**, 4079 (1986).

[7] M. J. Collett and D. F. Walls, Phys. Rev. A **32**, 2887 (1985).

[8] L. A. Orozco, M. J. Raizen, Min Xiao, R. J. Brecha, and H. J. Kimble, J. Opt. Soc. Am. B **4**, 1490 (1987).

[9] M. D. Reid and D. F. Walls, Phys. Rev. A **34**, 4929 (1986).

[10] D. A. Holm and M. Sargent III, Phys. Rev. A **35**, 2150 (1987).

[11] Min Xiao, H. J. Kimble, and H. J. Carmichael, J. Opt. Soc. Am. B **4**, 1546 (1987).

[12] D. M. Hope, D. E. McClelland, and C. M. Savage, Phys. Rev. A **41**, 5074 (1990).

[13] C. Wang and R. Vyas, Phys. Rev. A **51**, 2516 (1995).

[14] C. Wang and R. Vyas, Phys. Rev. A **54**, 4453 (1996).

[15] S. Sarkar and J. S. Satchell, Europhys. Lett. **3**, 797 (1987).

[16] C. W. Gardiner, *Handbook of Stochastic Methods*, edited by H. Haken (Springer-Verlag, Berlin, 1985) [Russian translation, Mir, Moscow, 1986].

[17] A. V. Kozlovskiĭ and A. N. Oraevskiĭ, Zh. Éksp. Teor. Fiz. **109**, 1524 (1996) [JETP **82**, 820 (1996)].

[18] B. Daeubler, H. Risken, and L. Schoendorff, Phys. Rev. A **46**, 1654 (1992).

[19] A. V. Kozlovskiĭ and A. N. Oraevskiĭ, JETP Lett. **67**, 796 (1998).

[20] H. Haken, *Light* (Elsevier Science Publishers, Amsterdam, 1985, Vol. 2).

[21] C. W. Gardiner and M. J. Collett, Phys. Rev. A **31**, 3761 (1985).

[22] M. J. Collett and C. W. Gardiner, Phys. Rev. A **30**, 1386 (1984).

[23] H. J. Carmichael, J. Opt. Soc. Am. B **4**, 1588 (1987).

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Multiple ionization of a Thomas–Fermi cluster by a strong electromagnetic field

M. B. Smirnov and V. P. Kraĭnov*)

Moscow Physics and Technology Institute, 141700 Dolgoprudnyĭ, Moscow Region, Russia

(Submitted 4 February 1999)
Zh. Éksp. Teor. Fiz. **115**, 2014–2019 (June 1999)

We develop a new model of a Thomas–Fermi cluster that describes the distribution of electrons in alkaline clusters with many atoms. We examine the classical multiple ionization of such a cluster by a strong electromagnetic field. Finally, we calculate the degree of ionization as a function of the field strength. © *1999 American Institute of Physics.* [S1063-7761(99)00806-9]

1. INTRODUCTION

The multiple ionization of a Thomas–Fermi atom (and molecule) by intense low-frequency laser light has been studied theoretically by a number of researchers.[1–4] These researchers modeled a multielectron atom in the Thomas–Fermi approximation,[5] and assumed that the laser light transforms the atom into an ion by stripping it of its electrons, starting with the outer shells and proceeding down to electrons for which above-barrier ionization (classical ionization) ceases to exist. Quantum tunneling of the remaining atomic electrons under the effective potential barrier during the the laser pulse has a probability that is much lower than that of above-barrier ionization, so that it can be completely ignored in this model. Moreover, in view of the very strong dependence of the ionization probability on the intensity of the electromagnetic field, ionization occurs only near the intensity maximum, which means that the ionizing electric field may be assumed constant. This quasi-steady-state approach to the ionization problem is valid at small values of the Keldysh parameter,[6] which correspond to high values of the electromagnetic-field intensity in the optical frequency range. Fields generated by high-power titanium–sapphire lasers fully meet these conditions.

The aim of the present work is to generalize the Thomas–Fermi model to alkaline clusters containing many atoms and to describe the classical (above-barrier) multiple ionization of such clusters in a strong electromagnetic field. Here the alkaline clusters proper are considered in the standard jellium model (see the review cited in Ref. 7). This model presupposes that the density of the atomic ions is constant and does not depend on the radial coordinate. The cluster electrons self-consistently adjust to the ion distribution specified in this manner. They fill shells, just as they do in atoms. When the number of electrons is large, the shell picture is replaced in the Thomas–Fermi model (as in an ordinary atom) by a continuous electron-density distribution.

2. THOMAS–FERMI MODEL

The Poisson equation for the electrostatic potential φ in a cluster has the simple form

$$\nabla^2\varphi=4\pi(n_e-n_i), \tag{1}$$

where n_e and n_i are the electron and ion concentrations, respectively. Throughout this paper we use natural units, with $\hbar=e=m_e=1$. In the jellium model, the ion concentration is constant and has the form

$$n_i=\frac{Z\eta(R-r)}{(4\pi/3)R^3}, \tag{2}$$

where Z is number of atoms in the cluster, R is the radius of the ion subsystem, and $\eta(x)$ is the Heaviside step function.

According to the Thomas–Fermi model for ions,[8] the potential φ is simply related to the electron concentration n_e:

$$n_e=\frac{(2(\varphi-\varphi_e))^{3/2}}{3\pi^2}, \tag{3}$$

where

$$\varphi_e=\varphi(r_e) \tag{4}$$

is the electrostatic potential at the boundary r_e of the electron distribution.

Combining Eqs. (1)–(3), we obtain a self-consistent equation for the electrostatic potential at $r<r_e$:

$$\nabla^2\varphi=\frac{1}{r}\frac{d^2}{dr^2}(r\varphi)=4\pi\left[\frac{1}{3\pi^2}(2(\varphi-\varphi_e))^{3/2}-\frac{Z\eta(R-r)}{(4\pi/3)R^3}\right]. \tag{5}$$

The boundary conditions for this second-order differential equation are

$$\frac{d\varphi}{dr}(r=0)=0, \quad \varphi(r=r_e)=\varphi_e. \tag{6}$$

The first corresponds to a vanishing electrostatic field at the coordinate origin of a spherical cluster (in contrast to the Thomas–Fermi atom), and the second corresponds [according to (3)] to the boundary r_e of the electron distribution of a charged cluster ion (the charge results from ionization by the electric field).

Outside the electron distribution $(r>r_e)$, the solution depends on the ratio of the electron and ion distributions. If $r_e>R$, the problem is simple. In this case, for $r>r_e$, the potential of the cluster ion outside the electron distribution is determined by the Coulomb field of the charges Z and Z':

$$\varphi(r)=\frac{Z-Z'}{r}. \tag{7}$$

Multiplying both sides of Eq. (5) by r^2 and integrating with respect to the radial component, we can easily verify that at the boundary r_e of the electron distribution not only is the potential continuous, but so is its radial derivative, i.e.,

$$\varphi'(r_e)=-\frac{Z-Z'}{r_e^2}. \tag{8}$$

This equation makes it possible to express Z' in terms of r_e, since the left-hand side of Eq. (8) is known from the solution of Eq. (5).

The situation is somewhat more difficult when $r_e<R$. The solution (7) is valid only in the region $r>R$. In the region $r_e<r<R$, the solution for the potential has the form

$$\varphi(r)=-\frac{Z'}{r}-\frac{Zr^2}{2R^3}+\frac{3Z}{2R} \tag{9}$$

(making use of the electrostatic solution for an ionized sphere with uniform bulk charge density). If we match this solution at the boundary of the electron distribution to the interior solution (5), we can express the number of electrons Z' in terms of the radius r_e. Instead of (8) we obtain

$$\varphi'(r_e)=-\frac{Z(r_e/R)^3-Z'}{r_e^2}. \tag{10}$$

We solved Eq. (4) numerically with the boundary conditions (5) for a typical cluster containing $Z=100$ sodium atoms. The ion concentration was taken equal to the concentration of solid crystalline sodium,[9] i.e., $n_i=2.652\times10^{22}\,\mathrm{cm}^{-3}$. According to (2), the ion radius R of such a cluster is 18.25 a.u.

The typical electron potential energy $-\varphi$ for a positively charged cluster (the energy is expressed in atomic units) as a function of the radial coordinate (also in atomic units) is shown by curve *1* in Fig. 1 for $Z'=54$ (a roughly half-ionized cluster). In this case, $r_e=16.25\,\mathrm{a.u.}<R=18.25\,\mathrm{a.u.}$ Clearly, over most of the electron distribution (the hatched region) the potential is constant, i.e., there is no field. The value $\varphi(0)=2.82\,\mathrm{a.u.}\sim77\,\mathrm{eV}$ provides an estimate for the binding energy of the interior electrons in such a cluster. The Fermi limit (the vertical dotted line) corresponds to electrons with an energy 2.70 a.u.∼73 eV. Note that for a neutral cluster ($Z=Z'$) the binding energy of the interior electrons is much lower: $\varphi(0)=0.12\,\mathrm{a.u.}=3.26\,\mathrm{eV}$.

Figure 2 depicts the radial dependence of ion concentration n_i given by (2) (dotted line), and the curves represent the electron concentrations n_e calculated by (3) in terms of the potential (established earlier) for various values of cluster ion charge (the case depicted in Fig. 1 corresponds to curve *3* in Fig. 2). Clearly, the radius of the electron distribution in a neutral cluster (curve *1* in Fig. 2) is close to the ion distribution radius. However, the smearing of the electron distribution is much more pronounced than that of the ion distribution. Generally, the electron distribution is normalized to the electron number Z' in the cluster:

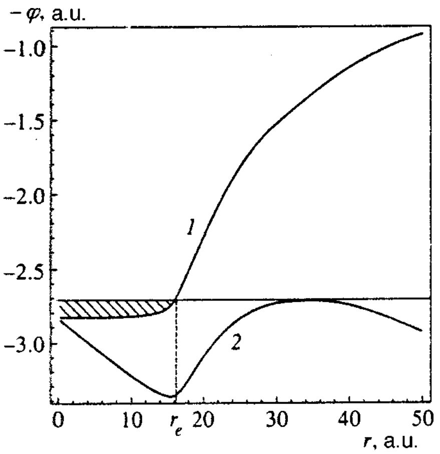

FIG. 1. Unperturbed potential energy $-\varphi(r)$ of a cluster of 100 sodium atoms containing 54 electrons (curve *1*). The hatched region is the one occupied by electrons. The effective potential energy (curve *2*) was calculated with Eq. (12) for an electric field strength $F=0.04$ a.u., which corresponds to (13). The horizontal line represents the Fermi limit for electrons. The vertical dotted line corresponds to the boundary of the electron distribution.

$$4\pi\int_0^\infty n_e(r)r^2\,dr=Z'\leqslant4\pi\int_0^\infty n_i(r)r^2\,dr=Z. \tag{11}$$

For curves *1–5* in Fig. 2, the values $r_e=\infty$, 21.2, 16.25, 12.2, and 8 a.u. correspond to $Z'=100$, 92.3, 54.35, 22.7, and 4.9. The shape of these curves suggests that in all cases there is an interior region in the electron distribution where the electron concentration is independent of the radial coordinate, i.e., an electron is not attracted to the center.

Of course, if the number of electrons is small, the cluster ion is unstable due to the strong mutual repulsion of the ions, and it decays rapidly .

3. MULTIPLE IONIZATION OF A CLUSTER

Up to this point we examined the behavior of a cluster in a vanishing external field. Now, in accordance with the ideas

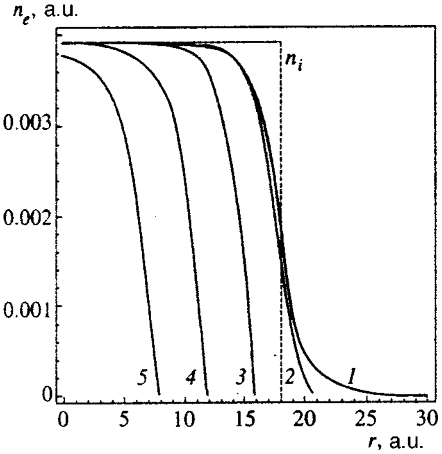

FIG. 2. Electron number density n_e in cluster ions consisting of 100 sodium atoms as a function of the radial coordinate: curve *1*, neutral cluster ($Z=Z'=100$); curve *2*, $Z'=92.3$; curve *3*, $Z'=54.35$; curve *4*, $Z'=22.7$; and curve *5*, $Z'=4.9$. The dotted line represents the ion number density n_i as a function of r.

advanced in the Introduction relating to the quasi-steady-state nature of a low-frequency electromagnetic field, we bring a constant electric field (field strength F) into the picture. Such a field ionizes the neutral cluster (which becomes a multiply-charged cluster ion) by detaching a classically well-defined number of electrons from the cluster. The field strength F is related to the charge $Z-Z'$ that remains after electron detachment in a manner that follows from the expression for the effective potential (see curve *2* in Fig. 1):

$$V=-\varphi(r)-Fr\approx-\frac{Z-Z'}{r}-Fr. \tag{12}$$

If we find the maximum of this expression (which lies outside the electron and ion distributions) and equate it to the maximum effective potential of the Fermi energy, φ_e, we obtain the electric field strength F that leads to a cluster with a given degree of ionization (this dependence is similar to the Bethe formula for atomic ionization; see Eq. (54.2) in Ref. 10):

$$F=\frac{(\varphi(r_e))^2}{4(Z-Z')}. \tag{13}$$

Of course the electric field is actually directed along some Cartesian axis, rather than along the radial variable. We assume that in the event of rapid ''ionization'' (atomic times) of the first electrons on the axis along which the external field acts, subsequent electrons rapidly reoccupy those locations and are then also ''ionized'' (or detached) along the same axis, with the result that a multiply-charged cluster ion is produced. The process is self-consistent, i.e., the cluster-ion potential is altered by ionization.

For our example of a sodium cluster with 100 atoms, we calculated the degree of ionization, defined as $\alpha=(Z-Z')/Z$, as a function of F. The result is depicted in Fig. 3. Clearly, multiple ionization of the cluster occurs at field strengths up to 0.14 a.u. For the scale of field strengths shown in Fig. 3, the Keldysh parameter is small compared to unity, so the ionization can be reliably considered quasi-steady-state (see the Introduction). The ejected electrons are aligned in the direction of the effective field vector, i.e., in the direction of polarization of linearly polarized laser light.[11]

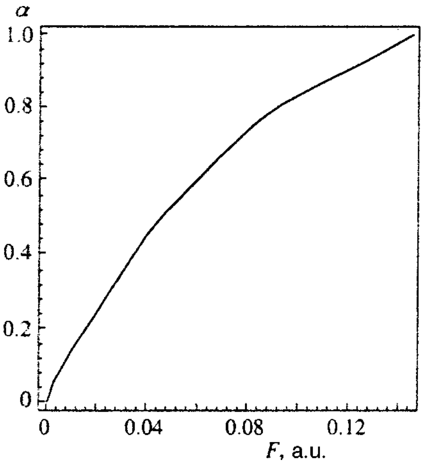

FIG. 3. Degree of ionization of a cluster of 100 sodium atoms as a function of electric field strength F (in atomic units).

After termination of the laser pulse, the remaining electrons and the ions fly apart in a spherically symmetric way due to Coulomb repulsion (Coulomb explosion of a cluster ion).

In conclusion we note that our numerical example is typical, so that all phenomena described in this paper occur in clusters of other atoms and with other numbers of particles.

This work was supported by a grant from the Russian Fund for Fundamental Research (Grant No. 99-02-17810).

*)E-mail: krainov@theory.mipt.ru

---

[1] V. P. Kraĭnov and É. A. Manykin, Ukr. Fiz. Zh. **25**, 400 (1980).

[2] V. Kresin, Phys. Rep. **220**, 1 (1992).

[3] M. Brewczyk, K. Rzążewski, and Ch. W. Clark, Phys. Rev. Lett. **78**, 191 (1997).

[4] M. Brewczyk, Ch. W. Clark, M. Lewenstein, and K. Rzążewski, Phys. Rev. Lett. **80**, 1857 (1998).

[5] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Nonrelativistic Theory*, 3rd ed., Pergamon Press, Oxford (1977).

[6] L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1945 (1964) [Sov. Phys. JETP **20**, 1307 (1965)].

[7] B. M. Smirnov, Usp. Fiz. Nauk **162**(12), 97 (1992) [Sov. Phys. Usp. **35**, 1052 (1992)].

[8] S. Flügge, *Practical Quantum Mechanics II*, Springer, Berlin (1971).

[9] C. Kittel, *Introduction to Solid State Physics*, 5th ed. Wiley, New York (1976).

[10] H. A. Bethe and E. E. Salpeter, *Quantum Mechanics of One- and Two-Electron Systems*, Springer, Berlin (1958).

[11] N. B. Delone and V. P. Kraĭnov, Usp. Fiz. Nauk **168**, 531 (1998) [Phys. Usp. **41**, 469 (1998)].

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Effect of a dust component on the rates of elementary processes in low-temperature plasmas

V. V. Ivanov,*) T. V. Rakhimova, A. O. Serov, and N. V. Suetin

D. V. Skobeltsyn Scientific Research Institute of Nuclear Physics, 115899 Moscow, Russia

A. F. Pal'

Troitsk Institute of Innovative and Thermonuclear Research, 142092 Troitsk, Moscow Region, Russia

(Submitted 7 October 1998)
Zh. Éksp. Teor. Fiz. **115**, 2020–2036 (June 1999)

Elementary processes in dusty, beam-driven plasma discharges are studied experimentally and theoretically for the first time. A theoretical model is constructed for a beam-driven plasma containing macroscopic particles. The effect of macroscopic particles on the electron energy distribution function is estimated assuming a Coulomb field for the particles. The resulting rate of electron–ion recombination on the macroscopic particles is compared with the electron loss constant calculated from the electron energy distribution function with an electron absorption constant in the orbital-motion approximation. This approximation, which is valid in the collisionless case, is found to work satisfactorily beyond its range of applicability. The distributions of the charged particles and electric fields created by macroscopic particles in a helium plasma are determined. The experimental data demonstrate the importance of secondary emission by high-energy electrons. *© 1999 American Institute of Physics.* [S1063-7761(99)00906-3]

1. INTRODUCTION

Dusty gas plasmas, that is, plasmas containing submicron- and micron-sized dust particles (macroscopic particles) in addition to electrons, ions, and gas atoms and molecules, have been studied widely in natural phenomena in astrophysics, such as the tails of comets, interstellar dust, planetary rings, etc. Although the first laboratory studies of dusty plasmas were done early in the century,[1] the greatest interest in research on these plasmas has arisen in recent years in connection with the rapid development of microtechnologies and technologies for producing new materials in plasma reactors.[2–4] On one hand, these plasmas are viewed as one of the most efficient methods for synthesizing nanoparticles with unique physical properties, and this determines their widespread practical importance. On the other hand, research on the processes taking place in dusty plasmas and the dynamic behavior of the structures formed in them is of deep fundamental interest, both for plasma physics and for solid state physics.

It should be noted that because of the greater mobility of the electrons, particles in these plasmas acquire a negative charge, but if secondary electron emission or photoemission are predominant, then the charge can also be positive. A system of dust particles of like charge can form a stable structure when long-range attractive forces are present owing, for example, to the interaction of the macroscopic particles with the plasma particles.

The experimental and theoretical studies of plasmas containing macroscopic particles have dealt mainly with their formation,[5,6] coagulation,[7] and spatial distribution,[8] the particle temperatures,[9,10] and the formation of ordered structures (''Coulomb'' crystals).[3,4] Most papers on observations and studies of Coulomb crystals concern rf discharges, as they are the most widely used in microtechnology. Thus, several similarity laws have been formulated for the structures which develop in the electrode sheaths of rf discharges[4,11–13] (reduced interparticle distance in the lattice of a structure with increasing discharge power, reduced lattice stability as the particle size increases, the possible escape of particles from the lattice resulting in dislocations).

The successful use of dc discharges in the reactors employed for plasma deposition of thin films has led to interest in research on dusty plasmas in these plasmas.[14,15] However, the presence of particles in the plasma can change its properties substantially because they represent an additional absorber (and sometimes, emitter) of electrons and ions.

There is special interest in the mechanism for nucleation of macroscopic particles and their effect on the microscopic and macroscopic characteristics of the plasma. Dust particles affect the electron energy distribution function, charge composition, charged particle densities, and chemical composition of a plasma.

Unfortunately, at present there are no detailed experimental and theoretical studies of the properties of plasmas containing macroscopic particles. A number of theoretical papers[16–18] indicate that macroscopic particles do have an effect on the electron distribution function and, therefore, on the transport coefficients. In order to construct a theory of plasmas with macroscopic particles, some study of elementary processes (excitation, recombination, etc.) in the plasmas will be needed. One of the main processes determining the charged particle balance in the plasma is bulk recombination. When macroscopic particles are present, additional electron–ion recombination takes place on the particle surfaces. This recombination is one of the principal ways macroscopic dust particles affect the plasma.

Direct measurements of the rate constants for this kind of recombination have, however, not been made. In this paper we propose a method for measuring the rate of loss of electrons on dust particles. Measurements of this type are made for the first time. The results are presented inthe form of rate constants for volume recombination. Usually the recombination rate of a plasma is determined from its decay characteristics, i.e., somehow the rate at which the charged particle density (as a rule, the electron density) is measured when the source by which the plasma is sustained is turned off. This method, however, is rather difficult andit is hard to interpret the results. In addition, the rate of loss ofelectrons in a decaying plasma can differ from the standard case, e.g., because of differences in the electron energy distribution function. Measurements of the plasma parameters as a function of the electric field strength in stationary self-sustained discharges are difficult because of the strong coupling of the field with the ionization rate. We believe that non-self sustained gaseous discharges, i.e., discharges maintained by an external ionization source, offer unique prospects for measuring the efficiency of recombination processes, including those on dust particles.

Indeed, in this case, the rate of production of charged particles is essentially independent of the plasma parameters and is determined by the external ionizer, and when it is constant the charged particledensity is determined exclusively by the processes by which they are lost in the plasma. Therefore, a beam-driven gaseous discharge is a unique instrument for studying charged particle loss processes, including those on dust components. In addition, as far as we know, no studies have been made at all on dusty, beam-driven discharge plasmas.

2. DESCRIPTION OF EXPERIMENT

The experiments were done on the apparatus sketched in Fig. 1. An electron accelerator injected a beam of 125 keV electrons at a current density of up to 100 μA/cm^2 in a pulse lasting 1 ms into the discharge gap through a 55-μm-thick beryllium foil. The 0.9-cm-long discharge gap was formed by a metal grid through which the beam could pass and a continuous circular electrode with an area of 1 cm^2. A voltage pulse with a controlled amplitude, duration, and delay was applied to them. A large part of the experiments were done in a non-self sustained discharge in helium at atmospheric pressure.

The discharge gap was filled with dust in the form of a powder consisting of microscopic particles of carbon glass. A scanning electron microscope picture of the dust particles is shown in Fig. 2. It can be seen in the picture that the particles are regular spheres with diameters of $24\pm 5\ \mu$m. Powder was placed in a container (Fig. 1) with a mesh bottom, located above the discharge gap. A heater was placed in the container for heating the powder in order to remove possible impurities. The container could be shaken once or at a frequency of 100 Hz using a special vibrator. A container for collecting the used powder was placed below the discharge gap. These containers were connected by a vertical pipe. The discharge electrodes were located in the midsection of the pipe. In the perpendicular direction there are 2-mm-diam holes for the entrance and exit of the probe laser beam. The outer case of thedischarge chamber, which provided a vacuum seal, was constructed so that the optical windows were far from the dusty volume.

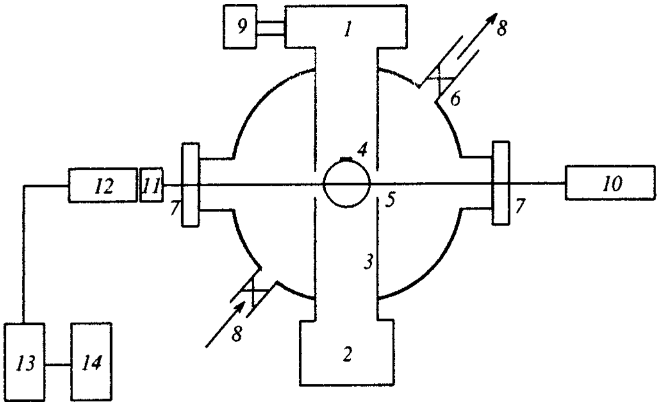

FIG. 1. Experimental arrangement: (*1*) container with powder, (*2*) container for collecting dust, (*3*) pipe in which the dust cloud propagates, (*4*) electrodes, (*5*) holes for entry and exit of probe light, (*6*) casing, (*7*) optical windows, (*8*) gas inlet and outlet ports, (*9*) vibrator, (*10*) laser, (*11*) filter, (*12*) detector, (*13*) oscilloscope, (*14*) computer.

After the dust container was attached to the vessel and before filling with the working gas, the vessel was pumped out with a roughing pump for several hours and the container was heated to 400 °C. Between discharge pulses working gas inside the vessel was continuously renewed. Gas pumping was stopped before the discharge was turned on. Following a single shake of the upper container, the settling cloud of dust occupiedthe discharge gap for a few seconds. Visual observation (without a discharge) shows that the dust completely fills the internal volume of the discharge vessel and is essentially absent outside it.

The density of dust particles in the plasma was determined from the attenuation of a laser beam in the dust cloud. The setup for these measurements is shown in Fig. 1. Light from a He–Ne laser (wavelength 0.63 μm, power 1 mW) passes through the discharge vessel with the dusty gas. The transmitted light passes through a filter to a detector (FD-24K photodiode) and is recorded on an oscilloscope (S9-8). The filter was chosen so that the detector current was below the saturation level.

The density N_d of dust particles under our conditions, where the radius r_d of the particles is much greater than the wavelength, was calculated using the following formula based on the Bouguer–Beer law:[19]

$$N_d = 0.75 \log(I_0/I) r_d^2 s, \tag{1}$$

where I and I_0 are the respective intensities of the light with and without dust in the vessel and $s=1$ cm is the depth of the dust cloud. Estimates showed that the error in measuring N_d is mainly determined by the uncertainty in the diameter of

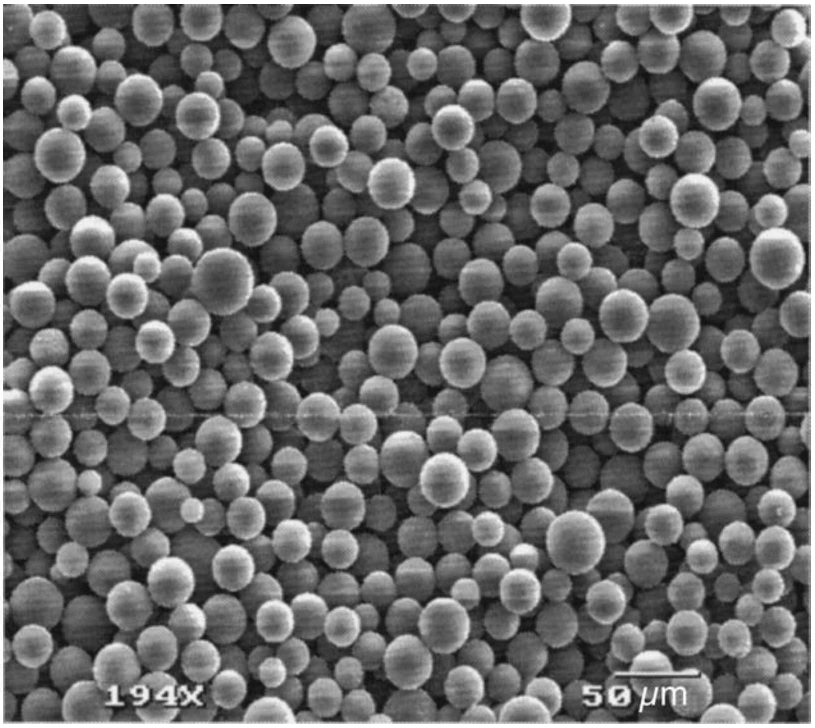

FIG. 2. A microphotograph of a dust particle.

the particles and was about 8% in our case. The range of measured dust densities for the particle diameter of $\sim 25\,\mu$m was $5\cdot 10^3 - 10^6$ cm^{-3}.

Figure 3 is an oscilloscope trace of the signal after the powder container was shaken. It can be seen that dust remains in the discharge gap for a few seconds and the dust density falls off slowly with time. Thus, it is possible to conduct the studies over a wide range of dust densities by varying the time delay before the discharge is turned on. In our experiments the discharge duration was less than 1 ms. Over this time the density of dust particles is essentially constant.

3. MEASUREMENT RESULTS

Figure 4 shows a current–voltage characteristic $j(U)$ of a discharge in pure helium for a beam current of 60 μA/cm^2. The measurements were made after the current had reached its quasistationary value. The ionization rate S was 10^{16} cm$^{-3}\cdot$s^{-1} for this beam current.[20] Given this production rateand a determination of the electron density from the discharge current, we can determine the effective rate of loss of the electrons in a dust-free plasma. In fact, when there are no macroscopic particles present, a non-self sustaining discharge in helium can be described by the following system of equations:

$$\frac{\partial n_e}{\partial t} = S - \beta n_e n_{\mathrm{He}_2^+}, \tag{2}$$

$$\frac{\partial n_{\mathrm{He}_2^+}}{\partial t} = -\beta n_e n_{\mathrm{He}_2^+} + k n_{\mathrm{He}^+} n_{\mathrm{He}}^2, \tag{3}$$

$$\frac{\partial n_{\mathrm{He}^+}}{\partial t} = S - k n_{\mathrm{He}^+} n_{\mathrm{He}}^2, \tag{4}$$

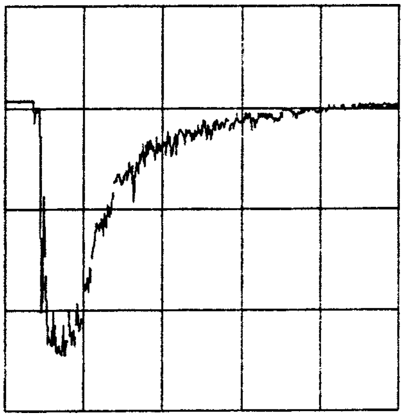

FIG. 3. Oscilloscope trace of the signal from the dust particle detector (2.5 V/div, 2 s/div).

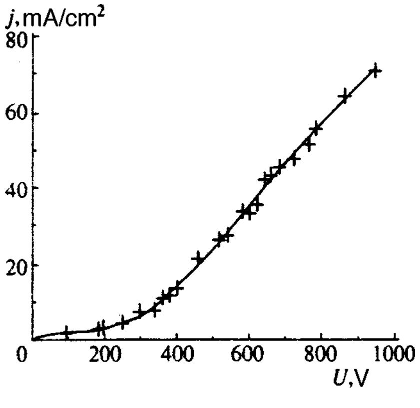

FIG. 4. Current–voltage characteristic of a beam-driven discharge in pure helium with a beam current of 60 μA/cm^2; +—experimental data. The smooth curve is an approximation using a fifth-order polynomial.

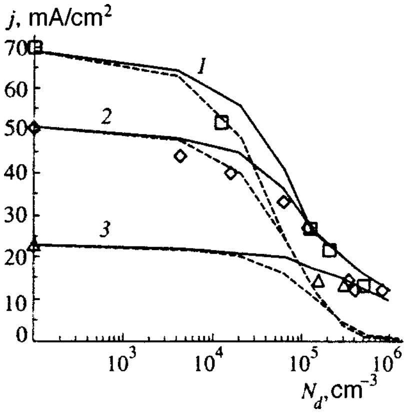

FIG. 5. Current density in a helium discharge as a function of the density of macroscopic particles and the applied voltage: (*1*) 940 V, (*2*) 720 V, (*3*) 480 V. The points (□, ◇, △) are the corresponding experimental data. Calculations neglecting (dashed curves) and including (smooth curves) a secondary emission coefficient of 1000.

where n_e, n_{He}, n_{He^+}, and $n_{\mathrm{He}_2^+}$ are the densities of electrons, helium atoms and ions, and helium molecular ions, t is time, β is the recombination coefficient for the helium dimer ions, k is the coefficient of conversion of helium ions into the dimer He_2^+, and S is the rate of ionization by the electron beam. The observed $j(U)$ curve for $U>450$ can be explained using Eqs. (2)–(4), assuming that the recombination coefficient is $\beta=2.5\cdot 10^{-8}\ \mathrm{cm}^3\mathrm{s}^{-1}$ and the cathode fall is $U_c=450$ eV. At lower voltages, the discharge description is more complicated, since the discharge becomes a Thomson discharge because of the low bulk ionization rate, and most of the voltage drop is across the cathode sheath. The recombination coefficient obtained above for the He_2^+ ions is roughly fifty times the published value. This difference may be caused either by vibrational excitation of the helium dimer ion or by the presence of impurity molecular gases. The first reason seems unlikely because of the high gas pressure and rapid vibrational relaxation of molecular ions. The second reason is better justified. In fact, because of charge exchange on impurity ions, the helium ions are neutralized-fairly rapidly and the molecular ions of an impurity gas, such as nitrogen, undergo recombination. The rate of recombination for nitrogen molecular ions is quite realistic for this process. Note that 10^{-3}% impurity nitrogen is sufficient for the N_2^+ ion to be the most common ion in the discharge.

Adding dust particles to the plasma (beginning at densities of $5\cdot 10^3\ \mathrm{cm}^{-3}$) alters the current–voltage characteristic of the discharge, specifically, reduces the current associated with a rise in the rate of loss of charged particles owing to recombination on the dust particles. Figure 5 shows plots of the current density in the discharge as a function of the density of macroscopic particles for different voltages. The figure shows that the current density decreases as the density of particles increases. The rate at which the current density decreases depends on the applied voltage. When no macroscopic particles are present, the current density is higher at a higher voltage, and as the particle density is increased this difference becomes smaller. For dust particle densities exceeding $10^5\ \mathrm{cm}^{-3}$, the current density is essentially the same for all voltages.

4. THEORETICAL MODEL

To determine the charged particle loss rate in a dusty plasma it is necessary to solve a self-consistent system of equations that includes the Boltzmann equation for the electron distribution function along with the particles in the plasma, the continuity equation for the plasma particles and macroscopic particles, and the Poisson equation for the electric field in the neighborhood of the macroscopic particles. The complete statement of the problem for a collisional plasma is rather complicated, since the electron distribution function has a dependence on the distance to a macroscopic particle and it is necessary to solve the inhomogeneous Boltzmann equation:

$$\frac{\partial f}{\partial t}+\mathbf{v}\nabla f-\frac{e\mathbf{E}}{m}\nabla_v f=\left(\frac{df}{dt}\right)_{\mathrm{coll}}, \tag{5}$$

where the symbol ∇_v denotes the gradient in velocity space, $(df/dt)_{\mathrm{coll}}$ is the collision integral, and e is the electronic charge. As a result, it is necessary to solve the complete self-consistent problem for determining the radial dependences of the electron distribution function, electric field, and charged-particle densities.

In a beam-driven discharge, the electron distribution function at low energies (on the order of the potential of the macroscopic particles) is determined mainly by the external electric field and not by the degradation spectrum of theelectron beam. The weak dependence of the electron distribution function on the primary electron beam means that the effect of the beam on it can be neglected. As part of the present effort, we have estimated the effect of macroscopic particles on electron distribution function by solving the homogeneous Boltzmann equation in a binomial approximation, with the electron absorption cross section approximated in terms of the orbital motion. This approximation works well in a nearly collisionless plasma with

$$r_d\ll\lambda\ll l, \tag{6}$$

where λ is the characteristic shielding length of the potential of a macroscopic particle and l is the characteristic distance between the particles. The question of the applicability of this cross section in the collisional case requires a separate examination.

This approximation makes it possible to separate the problem of determining the electron distribution function from that of solving the continuity and Poisson equations. The method for determining the electron distribution function in this approximation has been discussed in detail elsewhere.[16] This method is based on the assumption that all the macroscopic particles of a given size r_d in the plasma have some average charge Z_d. In the binomial approximation for the steady state, the isotropic part of the electron distributionfunction is determined from an equation which we shall transform to a form with an explicit dependence on the basic parameters:

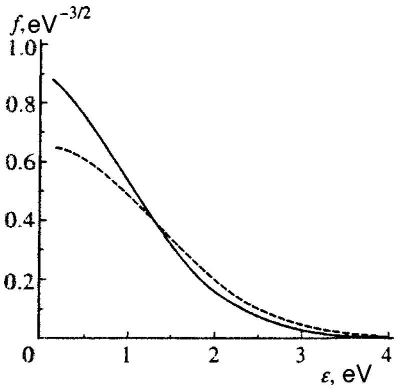

FIG. 6. Electron energy distribution function with (smooth curve) and without (dashed curve) macroscopic particles in a beam-driven atmospheric pressure discharge in helium. The electric field strength is 2 Td, the density of macroscopic particles is 10^6 cm^{-3}, and their potential is 0.75 eV.

$$\frac{\partial}{\partial\varepsilon}\left(\sqrt{\frac{2\varepsilon}{m}}\,\frac{e^2(E/N)^2}{3\sigma_t}\,\frac{\partial}{\partial\varepsilon}\,\frac{f_\varepsilon}{\sqrt{\varepsilon}}+\frac{\sqrt{2m}}{M}\,\varepsilon^{3/2}\sigma_t\right)+Q(f_\varepsilon)+S_r\left(1-\frac{U_d e}{\varepsilon}\right)\sqrt{\frac{2\varepsilon}{m}}=0, \qquad (7)$$

where E/N is the reduced electric field, $U_d=Z_d e/r_d$ is the potential of a macroscopic particle, $S_r=\xi\pi r_d^2 N_d/N$ is the reduced area of the macroscopic particles, ξ is the probability of attachment of an electron to a macroscopic particle, which was assumed equal to 1 for the calculations in this paper, and $Q(f_\varepsilon)$ is the collision integral for the plasma without macroscopic particles, normalized to the gas density.

The next term in the Boltzmann equation accounts for the loss of electrons to the macroscopic particles in the orbital approximation as a function of their energy ε. An expression for it can be obtained from the flux of electrons to a macroscopic particle,

$$Q(\varepsilon)=\xi\sigma_{\text{esc}}v,$$

where σ_{esc} is the cross section for electron capture by a macroscopic particle, given by

$$\sigma_{\text{esc}}=\begin{cases}\pi r_d^2(1-eU_d/\varepsilon), & \varepsilon\geqslant eU_d,\\ 0, & \varepsilon<eU_d,\end{cases} \qquad (8)$$

where v is the electron velocity.

Solving the Boltzmann equation showed that for a high helium pressure ($p=1$ atm), the change in the diffusion and electron drift coefficients in the weak fields $E/N<3$ Td characteristic of a beam-driven He discharge under the experimental conditions is not significant up to macroscopic particle densities $\sim 3\cdot 10^5$ cm^{-3}. At the maximum particle density in these experiments ($\sim 10^6$ cm^{-3}), the situation changes. Owing to depletion of the electron energy spectrum, the electron drift velocity and their diffusion change from 10 to 30%, depending on the applied electric field. (See Fig. 6, which shows the form of the electron distribution function without and with macroscopic particles.) It should be kept in mind that the effect of macroscopic particles on the electron distribution function in the neighborhood of the ionization potential of helium becomes important for low densities of macroscopic particles, but in a beam-driven discharge, ionization in the external field is negligible everywhere except in the cathode sheath, whose thickness is ~ 0.1 cm under our conditions.[21] The structure of the cathode sheath can be determined to a great extent by macroscopic particles, but this question lies outside the scope of this paper, which concerns processes taking place in the quasineutral positive column of a beam-driven discharge.

The system of equations for describing the interaction of the charged plasma particles with macroscopic particles in a one-dimensional (spherical) coordinate system in the diffusion-drift approximation has the following form:

$$\frac{\partial n_e}{\partial t}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_e\frac{\partial n_e}{\partial r}\right)-\frac{1}{r^2}\frac{\partial}{\partial r}(r^2 v_e^{dr}n_e)+S-\beta n_e n_{\text{He}_2^+}+S_e, \qquad (9)$$

$$\frac{\partial n_{\text{He}^+}}{\partial t}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_{\text{He}^+}\frac{\partial n_{\text{He}^+}}{\partial r}\right)-\frac{1}{r^2}\frac{\partial}{\partial r}\times(r^2 v_{\text{He}^+}^{dr}n_{\text{He}^+})+S-kn_e n_{\text{He}}^2, \qquad (10)$$

$$\frac{\partial n_{\text{He}_2^+}}{\partial t}=\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 D_{\text{He}_2^+}\frac{\partial n_{\text{He}_2^+}}{\partial r}\right)-\frac{1}{r^2}\frac{\partial}{\partial r}\times(r^2 v_{\text{He}_2^+}^{dr}n_{\text{He}_2^+})+kn_e n_{\text{He}}^2-\beta n_e n_{\text{He}_2^+}, \qquad (11)$$

with

$$v_e^{dr}=\mu_e E,\quad v_{\text{He}^+}^{dr}=\mu_{\text{He}^+}E,\quad v_{\text{He}_2^+}^{dr}=\mu_{\text{He}_2^+}E,$$

where D, v^{dr} and μ denote the diffusion coefficient, drift velocity, and mobility of the electrons and ions labeled by the different subscripts. Note that the electron diffusion coefficient D_e is determined by the magnitude of the external field E_{ext}, which is related to the external voltage by $E_{\text{ext}}=(U-U_c)/L$, where L is the interelectrode gap, rather than through the local field E. Note that, on the right hand side of Eq. (9), weinclude secondary emission of electrons from the macroscopic particles, as well as the electron source associated with the ionization of helium atoms by the electron beam.

It is known[22] that particles of nonconducting materials have a higher secondary emission coefficient. Here particles of carbon-glass were used. They have a complicated composite structure, so they may have a high emissivity σ_e. The source of the electrons associated with secondary emission was taken in a form that assumes that the cross section for interaction of a beam electron with a macroscopic particle whose electrical potential is much lower than the electron energy is given by πr_d^2:

$$S_e=\sigma_e j_f\pi r_d^2 N_d/e, \qquad (12)$$

where j_f is the beam current density.

The center of the chosen system of coordinates lies at the center of a macroscopic particle and the boundary conditions are the following: the ion density at $r=r_d$ is zero, the flux of electrons onto the macroscopic particle is limited by a maxi-

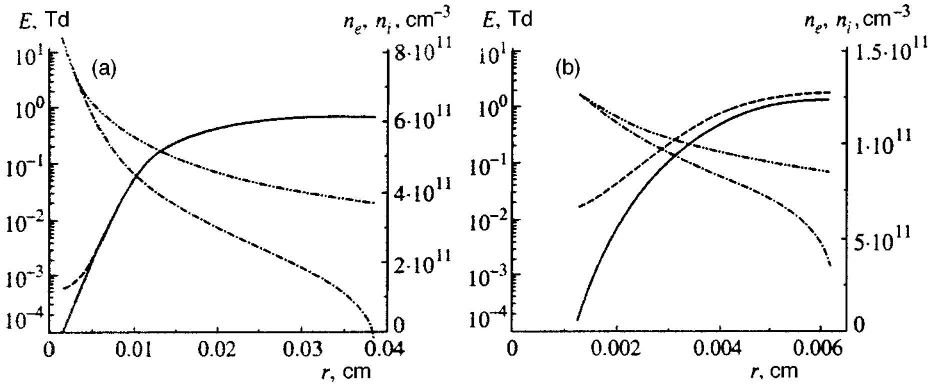

FIG. 7. Radial distributions of the electron density (smooth curves), ion density (dashed curves), and electric field strength (dot-dashed curves). The dot-dashed curves with two dots correspond to the electric field strength assuming a Coulomb potential for the macroscopic particle. The densities of macroscopic particles are 4000 cm^{-3} (a) and 10^6 cm^{-3} (b).

mum of $n_e v_e^{\rm th}/4$, where $v_e^{\rm th}$ is the electron sound speed. The boundary conditions at r equal to the radius of a sphere with the volume of a unit cell, corresponding to a single macroscopic particle, $R_d=(3/4\pi N_d)^{1/3}$, were determined from the symmetry conditions.

The system of continuity equations for all the components was closed by the Poisson equation for the electric field:

$$\frac{dE}{\mathrm{dr}}=4\pi e(n_{\mathrm{He}^+}+n_{\mathrm{He}_2^+}-n_e), \tag{13}$$

for which the boundary condition at $r=r_d$ was chosen from the quasineutrality condition for the plasma:

$$E_{r=r_d}=-\frac{eZ_d}{r_d^2}, \quad Z_d=\int_{r_d}^{R_d}4\pi e(n_{\mathrm{He}^+}+n_{\mathrm{He}_2^+}-n_e)r^2dr. \tag{14}$$

As estimates of the electron distribution function showed, the diffusion coefficients and drift velocities of the electrons can be calculated neglecting the effect of the macroscopic particles over almost all the range of densities of the macroscopic particles that was studied experimentally.

Solving the system of Eqs. (9)–(14) yields the profiles of the electric field, the electron and ion densities, and the charge and potential of the macroscopic particles. For comparison with the experimental data on the current density in the external field, we have also calculated the average electron density in a unit cell:

$$\overline{n_e}=\frac{3}{R_d^3-r_d^3}\int_{r_d}^{R_d}n_e r^2 dr. \tag{15}$$

5. COMPUTATIONAL RESULTS AND COMPARISON WITH EXPERIMENT

Figures 7a and 7b show the radial distributions of the electronand ion densities and of the electric field strength for different densities of macroscopic particles. The Coulomb field of the macroscopic particles for the same charge shown here was calculated using the formula

$$E(r)=eZ_d r_d/r^2. \tag{16}$$

The particle radius was taken to be 12 μm, which corresponds to the average radius of the carbon-glass particles. As Fig. 7a shows, for a low particle density, the region over which the particle affect the plasma density extends out to a distance of 200 μm, while the region in which the quasineutrality of the plasma is significantly disrupted is much smaller, on the order of 50 μm. Thus, the electric field is also considerably lower than the Coulomb field for a given charge on aparticle, essentially everywhere except for a narrow layer of uncompensated ionic charge near the particle itself. However, the thickness of this layer is still much grater than the particle radius, and this partially justifies our earlier assumption of an orbital motion for the macroscopic particles in calculating the electron distribution function.

When the density of the macroscopic particles is increased (Fig. 7b), the quasineutrality region remains essentially unchanged in size, while it forms an ever larger fraction of the decreasing radius of the unit cell. As a result, for $n_d=10^6$ cm^{-3}, the average electron density decreases significantly below the ion density, while the electric field distribution comes closer to the Coulomb field over a larger part of the unit cell. Note that for a low particle density, their effect on the electron distribution function is small, so that the errors associated with the enhancement of the electric field near the particles also become less important.

Figures 8a and 8b show the dependence of the average charge Z_d of a macroscopic particle and its potential U_d on the density of macroscopic particles for different strengths of the external electric field. Z_d and U_d as functions of the macroscopic particle density because the regions of influence of neighboring particles overlap; this shows up as a reduction in the plasma density and, therefore, in the flux of electrons to a particle. At the same time, the flux of electrons from a macroscopic particle owing to secondary electron emission driven by the electron beam increases in proportion to the density of macroscopic particles. As a result, the fluxes of ions and electrons come into balance at lower Z_d and U_d. The dependence of these parameters on the external field is less obvious. In our model, the reduction in Z_d and U_d with decreasing $U_{\rm ext}$ is determined by the dependence of electron diffusion coefficients and of the average electron energy on the applied external voltage. Therefore, a drop in $U_{\rm ext}$ reduces the flux of electrons onto a macroscopic particle and, with that, in Z_d and U_d.

These calculations of the radial distribution of the electric potentialand charge on the macroscopic particles provide a better estimate of the Coulomb coupling parameter between the macroscopic particles. The parameter Γ_c has been

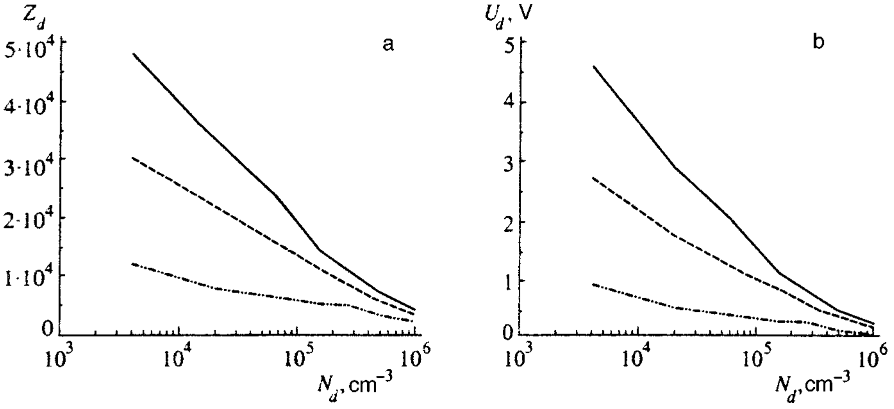

FIG. 8. The average charge on a macroscopic particle in units of the electronic charge (a) and the potential of the macroscopic particles (b) as functions of the density of macroscopic particles for external voltages of 940 (smooth curves), 720 (dashed curves), and 480 V (dot-dashed curves).

introduced in the literature to measure the degree of Coulomb coupling:[25]

$$\Gamma_c=\frac{1}{kT}\frac{Z_d e}{4\pi\varepsilon_0 d}, \tag{17}$$

where k is the Boltzmann constant, $Z_d e$ is the charge on the macroscopic particles, $d\approx 2R_d$ is the distance between the particles, and T is the kinetic temperature of the macroscopic particles. Essentially, Γ_c is the ratio of the potential energy of the macroscopic particles owing to the Coulomb interaction to their kinetic energy. Ithas been shown in Monte-Carlo calculations[23,24] that a three dimensional Coulomb liquid develops for $\Gamma_c>2$, while a Coulomb crystal develops for $\Gamma_c>170$.

The values of Γ_c given here were obtained in the approximation of a one-component plasma, so that they neglect the screening of the potential of the macroscopic particles by the ions and electrons in a real plasma, which may reduce the parameter Γ_c by several orders of magnitude. In order to take the screening into account, the formula

$$\Gamma_d=\Gamma_c\exp(-l/\lambda), \tag{18}$$

is customarily employed, where the screening length λ is taken to be the electron Debye radius.

In this paper we have calculated Γ_c, including the screening of the potential of the macroscopic particles. Table I lists the values of various parameters characterizing the beam-driven discharges with $U=940$ V formed with the maximum density $N_d=10^6$ cm^{-3} (first number) and minimum density $N_d=4\cdot10^3$ cm^{-3} (second number) for the macroscopic particles.

Note the large difference between the values of the screening parameter λ obtained in the best approximation to a real plasma potential by the Yukawa potential (see Figs. 7a and 7b) and the Debye radius r_D, as well as the fact that the changes in the density of the macroscopic particles, the Debye radius, and the screening parameter are not correlated. In most papers, it is assumed as an estimate, that λ is determined by the Debye radius (e.g., Ref. 15). It has been proposed[25] that the screening parameter for the potential of the macroscopic particles when the particles are not a plasma component is given by

$$\frac{1}{\lambda}=\frac{1}{r_{De}}+\frac{1}{r_{Di}}, \tag{19}$$

that is, under our conditions, where $n_e\approx n_i$ and $T_e\gg T_i$, the screening length λ should be equal to the ion Debye radius r_{Di}. As the table shows, the actual screening radius is not only larger than r_{Di}, but also larger than r_{De}. Note that in Ref. 15 the screening parameter is not directly related to r_{Di} or r_{De}; this question lay outside the scope of the paper. In fact, the screening parameter is related to the Debye length only in the linear approximation, while near a macroscopic particle the changes in the ion and electron densities reach 100%. In the nonlinear case, λ can be estimated on the basis of charge conservation. Let us assume that in the screening region $r<r_d+\lambda$, the ion density is equal to the ion density far from the macroscopic particle, and the electron density equals zero, while outside that region the quasineutrality condition is satisfied. Then Eq. (14) yields the following equation for λ:

TABLE I.

Parameter	n_e, 10^{11} cm^{-3}	n_i, 10^{11} cm^{-3}	l, μm	r_{De}, μm	λ, μm
Formula	–	–	$N_d^{-1/3}$	$(kT_e/8\pi n_e e^2)^{1/2}$	$U\approx eZ_d e^{(-r/\lambda)}/r$
Value	1.5/6	1.6/6	100/600	15/7.5	45/60

Parameter	Γ_c	Γ_d	Γ_D	l_{sh1}, μm	l_{sh2}, μm
Formula	$Z_d^2e^2/lT_g$	$\Gamma_c\exp(-l/\lambda)$	$\Gamma_c\exp(-l/r_{De})$	$eEZ_dt^2/2m$	$F_bt^2/2m$
Value	$2\cdot10^4/3\cdot10^5$	$2\cdot10^3/13$	$3\cdot10^{-2}/10^{-30}$	0.1/1	0.05/0.1

$$Z_d=\frac{4}{3}\pi((r_d+\lambda)^3-r_d^3)n_i. \tag{20}$$

This formula allows us to understand the noncorrelation between the screening length and the Debye radius as the density of macroscopic particles decreases. Thus, when the Debye radius is calculated for a given field, the electron temperature does not change, while at the same time the electron density, which is roughly equal to the ion density, increases by a factor of 4, so the Debye radius decreases by a factor of two. At the same time, the screening length also depends on the charge on a macroscopic particle, which increases by an order of magnitude (Fig. 8a) and more than compensates the change in the ion density. As a result, the screening length increases by roughly one and a half times.

The screening length is, therefore, not coupled directly to the electronor ion Debye radius, although in magnitude it is close to the electron Debye radius under our experimental conditions. This question is of great import for estimating the Coulomb coupling parameter. In fact, as can be seen from the table, when this parameter is estimated taking screening into account, the result is qualitatively different, depending whether λ or r_{De} is used. Estimating Γ for screening with the Debye radius yields an uncorrelated system of macroscopic particles, while using the actual screening length can give rise to crystal structures in a beam-driven plasma. Under the present experimental conditions, however, a beam-driven discharge burned stably for at most 1 ms. Over this time, the particles move a distance of less than 1 μm (much less than the distance between particles), both owing to the electrostatic forces (over a distance l_{sh1}) and owing to the forces which arise as a result of the bombardment of the surface of a macroscopic particles by the plasma particles (over a distance l_{sh2}):[26]

$$F_b=4\pi r_d^4 n_e(T_i+T_e)/R_d^2. \tag{21}$$

The calculated average electron densities in a unit cell can be used to calculate the current density in the discharge as a function of the density of macroscopic particles:

$$j=e\overline{n}_e v_e^{dr}.$$

When secondary emission is neglected, the calculated and experimental dependences of the current density on the density of the macroscopic particles differ by more than a factor of 10 at high densities of thelatter (Fig. 5). In order for these results to agree, it is necessary to assume that the secondary emission coefficient equals 1000, a magnitude more or less typical of dielectrics. That this coefficient is so large is evidence of inhomogeneities in the composition of the carbon-glass particles and of the presence of poorly conducting regions within them.

We have also examined another possibility for explaining the experimental data through additional ionization of the plasma by secondary emission electrons. A comparison with experiment in this case showed that reasonable agreement is obtained if we assume that 10^2 ionization events take place per beam electron incident on a macroscopic particle. However, at high macroscopic particle densities the calculated current densities cease to depend on the applied external voltage, which conflicts with the experimental data. (See Fig. 5.) The version of this calculation in which only secondary electron emission from the beam electrons is included is free of this shortcoming, so it was chosen as the main version.

We can describe the loss of charged particles at macroscopic particles globally by introducing a coefficient of recombination of the plasma on macroscopic particles. It can be introduced by integrating Eqs. (9)–(11) over the volume of a unit cell, noting that $\overline{n}_{\mathrm{He}_2^+}\gg\overline{n}_{\mathrm{He}^+}$. Integrating Eq. (11) and writing the flux of ions on a macroscopic particle (the integralof the first two terms on the right of Eq. (11)) in the form $\beta_d\overline{n}_{\mathrm{He}_2^+}N_d$, we obtain

$$S=\beta_d\overline{n}_{\mathrm{He}_2^+}N_d+\beta\overline{n}_e\overline{n}_{\mathrm{He}_2^+}\approx\beta_d\overline{n}_e N_d+\beta\overline{n_e^2}. \tag{22}$$

Using the quantity $\overline{n}_e$ in place of $\overline{n_{\mathrm{He}_2^+}}$ is justified by their closeness in the calculations, even for high macroscopic particle densities (Fig. 7b), and makes it possible to compare the calculated and measured values of β_d. The integral of the first two terms on the right of Eq. (9), which determines the flux of electrons onto a macroscopic particle because the charge is time independent, should equal the sum of the flux S_e of electrons from the macroscopic particle and the ion flux $\beta_d\overline{n}_e n_d$.

Figures 9a and 9b show the calculated and experimental dependences of the plasma recombination coefficient β_d on N_d and U_{ext}. Including secondary emission in Fig. 9b makes it possible to match the experimental and theoretical results, both qualitatively and quantitatively. The recombination coefficient decreases as the density of macroscopic particles increases because of a reduction in the average charge of the macroscopic particles and a corresponding reduction in the ion and electron fluxes to them. The dependence of the recombination coefficient on the applied voltage is associated with an increase in the mobility and diffusion of the electrons as the reduced electric field is increased. As a result, we find that the recombination coefficient increases as the electric field strength is raised, both inthe calculations and in the experiments.

As an illustration of the extent to which the macroscopic particles affect the rate of recombination of the plasma, Fig. 9c shows a plot of the effective recombination coefficient β_{eff} as a function of N_d and the applied voltage:

$$\beta_{\mathrm{eff}}=S/n_e^2.$$

The figure shows that macroscopic particles at densities exceeding 10^5 cm^{-3} do have a significant effect on the recombination rate of the plasma. At a density of 10^6 cm^{-3} the recombination rate is more than an order of magnitude higher.

In conclusion, we note that these results can be used to estimate the feasibility of using the orbital approximation for the cross section for absorption of electrons by a macroscopic particle (8) in a collisional plasma, i.e., far beyond the formal limits for applicability of this approximation. To do this, we calculate β_d using a distribution function with the cross section (8) for $N_d=10^6$ cm^{-3}, when the potential of the electric field near a macroscopic particle is close to the

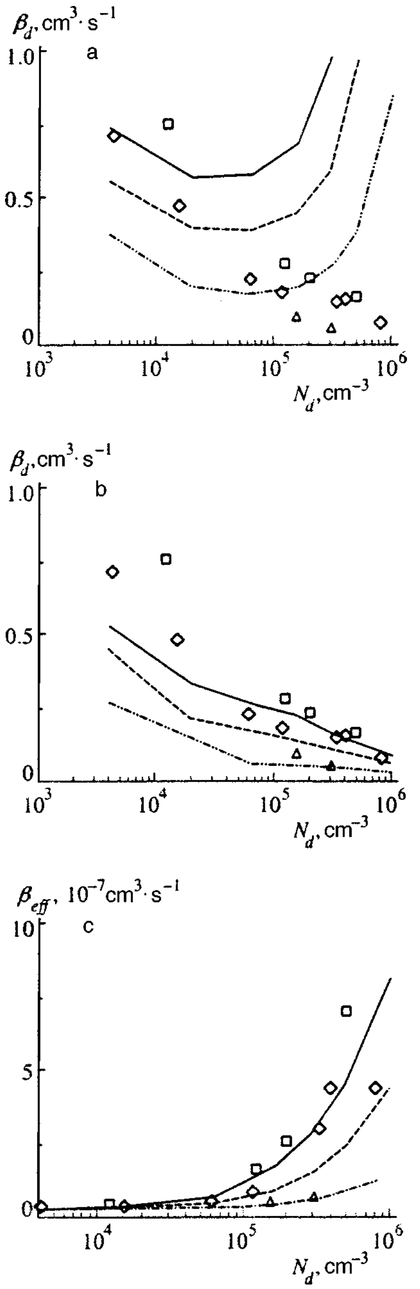

FIG. 9. Same as in Fig. 8, for the recombination coefficient of the plasma on macroscopic particles (neglecting (a) and including (b) a secondary electron emission coefficient of 1000) and for the effective recombination coefficient of the plasma on macroscopic particles (including a secondary emission coefficient of 1000) (c).

Coulomb value and $E/N = 2$ Td (which corresponds to the maximum applied voltage). These calculations yield $\beta_d = 0.8$, while the experimental value is a factor of six smaller (Fig. 9a). This shows that the orbital approximation for the cross section for electron absorption by macroscopic particles is in qualitative agreement with experiment, even far from the region where it is formally applicable.

6. CONCLUSION

In this paper we have determined the current density in a non-self sustained discharge experimentally for different voltages and injected particle densities. The discharge current density was observed to decrease as the particle density was raised when the electric field was held constant. When the field is raised, the dropoff rate increases. Based on these data, we have determined the recombination coefficient for the plasma on macroscopic particles. As the density of the particles increased up to 10^6 cm^{-3}, the recombination coefficient increases by roughly a factor of 40 compared to that in a dust-free plasma.

In this paper we have constructed a theoretical model for a non-self sustained discharge with macroscopic particles, both including their effect on the electron distribution function and plasma transport coefficients and calculating the distributions of the charged particle densities and of the potential near the macroscopic particles. A comparison of the experimental and theoretical results demonstrated the importance of secondary electron emission driven by the electron beam.

These calculations have shown that the size of the region where quasineutrality is violated may not be correlated with the electron Debye radius. The region where a macroscopic particle has an effect (i.e., the region where the plasma density differs significantly from the maximum) is considerably bigger than the region where quasineutrality is violated and varies slowly with the density of the macroscopic particles.

We have shown that the orbital approximation for the cross section for the absorption of electrons by macroscopic particles yields qualitative agreement with experiment even far from the region where it is formally applicable.

We thank G. I. Aponin, A. A. Besshaposhnikov, and D. M. Kulakov for help in measuring the macroscopic particle density.

This work was supported by the Russian Fund for Fundamental Research (Grants No. 96-02-18938 and No. 96-15-96447) and an INTAS International Grant (No. 96-235).

*)E-mail: ivv@mics.msu.su

---

[1] I. Langmuir, G. Found, and A. E. Dittmer, Science **60**, 392 (1924).
[2] G. S. Selwyn, J. Singh, and R. S. Bennet, J. Vac. Sci. Technol. A **7**, 2758 (1989).
[3] J. H. Chu and I. Lin, Phys. Rev. Lett. **72**, 4009 (1994).
[4] H. Thomas, G. E. Morfill, V. Demmel, J. Goree, B. Fewerbacher, and D. Mohlmann, Phys. Rev. Lett. **73**, 652 (1994).
[5] J. L. Dorier, Ch. Hollenstein, and A. A. Howling, J. Vac. Sci. Technol. A **13**, 918 (1995).
[6] A. A. Fridman, L. Boufendi, T. Hbid, B. V. Potapkin, and A. Bouchoule, J. Appl. Phys. **79**, 1303 (1996).
[7] V. A. Schweigert and I. V. Schweigert, J. Phys. D: Appl. Phys. **29**, 655 (1996).
[8] H. H. Hwang and M. J. Kushner, Appl. Phys. Lett. **68**, 3716 (1996).
[9] J. E. Dougherty and D. B. Graves, J. Vac. Sci. Technol. A **11**, 1126 (1993).
[10] A. P. Nefedov, A. G. Khrapak, S. A. Khrapak, O. F. Petrov, and A. A. Samaryan, Zh. Éksp. Teor. Fiz. **112**, 499 (1997) [JETP **85**, 272 (1997)].
[11] Th. Trotlenberg, A. Mezer, and A. Piel, Plasma Sources Sci. Technol. **4**, 450 (1995).
[12] J. B. Pierer, J. Goree, and R. A. Quinn, J. Vac. Sci. Technol. A **14**, 519 (1996).
[13] J. H. Chu, J. B. Du, and I. Lin, J. Phys. D **27**, 296 (1994).
[14] S. Nunomura, N. Ohno, and S. Takamura, Jpn. J. Appl. Phys., Part 2 **36**, L949 (1997).
[15] A. M. Lipaev, V. I. Molotkov, A. P. Nefedov, O. F. Petrov, V. M.

Torchinskii, V. E. Fortov, A. G. Khrapak, and S. A. Khrapak, Zh. Éksp. Teor. Fiz. **112**, 2030 (1997) [JETP **85**, 1110 (1997)].

[16] V. V. Ivanov, T. V. Rakhimova, in *Proc. of ICPIGXXXIII*, Toulouse, France (1997), Vol. I, p. 182.

[17] M. J. Mc. Caughey and M. J. Kushner, J. Appl. Phys. **69**, 6952 (1991).

[18] De-Zhen Wang, J. Q. Dong, and S. M. Mahajan, J. Phys. D: Appl. Phys. **30**, 113 (1997).

[19] G. I. Aponin, A. A. Besshaposhnikov, D. M. Kulakov, A. F. Pal, A. O. Serov, and N. V. Suetin, Aerosols C **4**, 73 (1998).

[20] C. Cason, J. F. Perkins, A. H. Werkheiser, and J. Duderstadt, AIAA J. **15**, 1079 (1977).

[21] V. V. Aleksandrov, E. P. Glotov, V. A. Danilychev, V. N.Koterov, and A. M. Soroka, Trudy FIAN **142**, 46 (1983).

[22] Yu. P. Raizer, *Gas Discharge Physics*, Springer, New York (1991) [Russian original, Nauka, Moscow (1987)].

[23] S. Ichimaru, Rev. Mod. Phys. **54**, 1017 (1982).

[24] W. L. Slattery, G. D. Doolen, and H. E. Dewitt, Phys. Rev. A **21**, 2087 (1980).

[25] V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997).

[26] Ya. K. Khodataev, R. Bingham, V. P. Tarakanov, and V. N. Tsytovich, Fiz. Plazmy **22**, 1028 (1996) [Plasma Phys. Rep. **22**, 932 (1996)].

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Theory of stimulated nonresonant emission from relativistic beams

D. N. Klochkov, M. Yu. Pekar, and A. A. Rukhadze*)

Institute of General Physics, Russian Academy of Sciences, 117942 Moscow, Russia

(Submitted 23 October 1998)
Zh. Éksp. Teor. Fiz. **115**, 2037–2050 (June 1999)

A study is made of the radiative Pierce instability of a relativistic electron beam in a waveguide stabilized by an infinitely strong magnetic field. Analytical and computational methods are used to determine the growth rate of the instability, as well as the efficiency for conversion of the beam energy into electromagnetic field energy as a function of the beam current, how relativistic the beam is, and the geometry of the system. The physical nature of the instability is clarified and the mechanisms for its saturation are discussed. © *1999 American Institute of Physics.* [S1063-7761(99)01006-9]

1. INTRODUCTION

It is well known that a relativistic electron beam can radiate coherent electromagnetic waves extremely efficiently as a result of a resonant interaction with a wave.[1] Recently it has been established[2] that a relativistic electron beam propagating rectilinearly in a waveguide is capable of radiating efficiently even when the phase velocity of the wave exceeds the beam velocity. This nonresonant collective mechanism for the interaction of the beam electrons with an *E*-mode electromagnetic wave is essentially a radiative Pierce instability.

The radiative Pierce instability has been realized for nonrelativistic electron beams in monotron generators.[3] Short systems, in which the transverse dimensions of the cavity exceed the longitudinal dimensions, were used.

The theory of the nonrelativistic monotron generator was first developed by Muller and Rostas,[4] who analyzed the properties of the generator with the aid of an equivalent circuit of lumped elements. It was found that the condition for the appearance of oscillations has the following form for the electron drift angle:

$$\theta \equiv \frac{\omega L}{u} = \left(2n + \frac{1}{2}\right)\pi,$$

where L is the length of the system and n is an integer. This condition corresponded to a loading characteristic with negative conductivity. This can be said in a different way: a positive feedback develops, which leads to self oscillations. In the 1940's Pierce[5] studied another instability, which also relied on positive feedback (between the output and input electrodes through an external circuit). Later the instability was interpreted in a different way: when the beam current exceeds a critical level, a beam wave develops in the system and propagates counter to the beam to create feedback (Ref. 1, Sections 49 and 50). It would have been natural to regard the two apparently different instabilities as one with two regimes. The first, potential regime is possible when the Pierce parameter $\omega_b \gamma^{-3/2}/k_\perp u$ exceeds unity. The second, radiative (wave) regime occurs at cavity frequencies above the critical frequency, when electromagnetic waves propagating counter to the beam can exist in the system. In a plasma cavity, instabilities can develop on a plasma wave.

In this paper we use analytical and numerical methods to study the linear and nonlinear stages of the nonresonant, stimulated emission of electromagnetic radiation by a rectilinear relativistic electron beam in a smooth cavity. A detailed investigation is made of the cylindrically symmetric case, i.e., when only azimuthally symmetric modes are excited in a cavity with a circular cross section. The mechanism by which the beam electrons interact with the electromagnetic field, and which leads to the development of the instability, is studied in the limit of small wave amplitude.

Two mechanisms for saturation of the instability are identified on the basis of numerical simulation. Phase portraits of the beam electrons are obtained which can be analyzed to reveal the physical nature of both saturation mechanisms for the instability. For the steady-state instability, we obtain the amplitude of the field in the cavity, as wellas the conversion coefficient for the energy of the beam electrons into field energy, as functions of the system geometry and the density and relativistic character of the electrons. The spectrum of the electromagnetic radiation is investigated.

We conclude by examining the dynamic instability of the beam particles in the field of the excited wave.

2. BASIC EQUATIONS

Let us consider a smooth, cylindrical metallic waveguide. The ends ofthe cavity are covered with metal foil or a mesh to create mirror boundary conditions for the electromagnetic field while being transparent to the beam electrons. The cavity length is L and its radius is R. The system lies in a uniform magnetic field that is directed along the cavity axisand is strong enough that the transverse motion of the beam electrons can be neglected. Maxwell's equations for the nonzero components of the field in a cylindrical coordinate system take the form

$$\frac{1}{r}\frac{\partial}{\partial r}(rE_r) + \frac{\partial E_z}{\partial z} = 4\pi\rho, \qquad -\frac{\partial B_\varphi}{\partial z} = \frac{1}{c}\frac{\partial E_r}{\partial t},$$

$$\frac{\partial E_r}{\partial z}-\frac{\partial E_z}{\partial r}=-\frac{1}{c}\frac{\partial B_\varphi}{\partial t}. \tag{1}$$

In this case it is also convenient to express the field in terms of asingle component of the Hertz polarization potential ψ:

$$E_z=\left(\partial_z^2-\frac{1}{c^2}\partial_t^2\right)\psi,\quad E_r=\partial_t\partial_r\psi,\quad B_\varphi=-\frac{1}{c}\partial_t\partial_r\psi. \tag{2}$$

The initial equation and boundary conditions for the field are

$$\partial_z\left(\Delta-\frac{1}{c^2}\partial_t^2\right)\psi=4\pi\rho,\quad E_z|_{r=R}=E_r|_{z=0,L}=0. \tag{3}$$

The perturbed charge density of the particles is then given by

$$\rho=en_b(r)\left[\int\delta(z-z(t,z_0))\,dz_0-1\right], \tag{4}$$

where $z=z(t,z_0)$ is the solution of the characteristic system of Vlasov equations:

$$\frac{dz}{dt}=v,\quad \frac{dv}{dt}=\frac{e}{m}\tilde{\gamma}^{-3}E_z. \tag{5}$$

(n_b is the unperturbed beam electron density and $\tilde{\gamma}=(1-v^2/c^2)^{-1/2}$ is the Lorentz factor.)

The beam charge is assumed to be neutralized. When an electron enters the cavity its unperturbed velocity is u. The beam leaves the cavity without hindrance, carrying away the acquired perturbations.

3. LINEAR THEORY

The radiative instability of a straight, neutralized relativistic electron beam in the linear approximation was examined on the basis of a solution to the dispersion equation.[2] The conditions for development of the instability and its growth rate were obtained a functions of the system geometry and beam current. It has been shown that the instability is a Raman instability, is self-oscillatory, and has no current threshold. In order to reveal the mechanism for transfer of the directed energy of the beam to the electromagnetic field, let us consider the work done by the longitudinal component E_z of the radiation field on an electron as it passes through the cavity,

$$A=e\int_0^L E_z(t[z],z)\,dz. \tag{6}$$

The component E_z is expressed, with the aid of Eq. (2), in terms of the potential ψ, which, to first order in the field amplitude for a waveguide with a beam that is uniform over the cavity cross section, is given by (Ref. 1, Section 8)

$$\psi=\phi_s(r)\sum_{\nu=1}^{4}A_\nu\cos(\omega t-k_\nu z), \tag{7}$$

where $\phi_s(r)$ is the transverse structure of the field. The longitudinal wave number k_ν corresponding to a wave with amplitude A_ν is related to the frequency ω by the dispersion relation (Ref. 1, Section 8)

$$k_\perp^2+\left(k_\nu^2-\frac{\omega^2}{c^2}\right)\left[1-\frac{\omega_b^2\gamma^{-3}}{(\omega-k_\nu u)^2}\right]=0. \tag{8}$$

If we average the work A of the field over the phases of the electrons or, equivalently, over the time t_0 the electrons spend in the cavity,

$$\langle A\rangle=\frac{1}{\tau}\int_0^\tau A(t_0)\,dt_0,\quad \omega^{-1}\ll\tau\ll(\delta\omega)^{-1}, \tag{9}$$

then summing over the phases of the unperturbed electrons makes no contribution to the radiation. In order to obtain nonzero coherent radiation, it is necessary to include the reaction of the radiation field on the beam electrons. In this case the beam comes into phase and a stimulated coherent radiation effect occurs that corresponds to a nonzero contribution to the work by the radiation field. Here there is no phasingof the radiation field accompanying the beam electrons.

When the beam is modulated, the velocity and trajectory of an electron are slightly perturbed, so that

$$v=u+\tilde{v},\quad t[z]=t_0+\frac{z}{u}+\tilde{t}. \tag{10}$$

Here $\tilde{v}$ and $\tilde{t}$ are the solutions of the linearizedequations of the characteristic system of Vlasov equations,

$$\frac{d\tilde{t}}{dz}=-\frac{\tilde{v}}{u^2},\quad \frac{d\tilde{v}}{dz}=\frac{e}{mu}\gamma^{-3}E_z. \tag{11}$$

The solution of Eqs. (11) with the potential (7) has the form

$$\begin{aligned}\tilde{v}&=\frac{e}{m}\gamma^{-3}\phi_s(r)\sum_{\nu=1}^{4}\frac{E_\nu}{\omega-k_\nu u}\sin\left[\omega t_0+\left(\frac{\omega}{u}-k_\nu\right)z\right],\\ \tilde{t}&=\frac{e}{mu}\gamma^{-3}\phi_s(r)\sum_{\nu=1}^{4}\frac{E_\nu}{(\omega-k_\nu u)^2}\cos\left[\omega t_0+\left(\frac{\omega}{u}-k_\nu\right)z\right].\end{aligned} \tag{12}$$

Here $E_\nu=(\omega^2/c^2-k_\nu^2)A_\nu$. In this case, the work averaged over the phases of the electrons is

$$\langle A\rangle=\frac{e^2\omega}{2mu}\gamma^{-3}\phi_s^2(r)X\sum_{i<j}a_{ij}\left[\frac{1}{(\omega-k_ju)^2}-\frac{1}{(\omega-k_iu)^2}\right], \tag{13}$$

where $X=E_2^2$, while

$$a_{ij}=\frac{E_iE_j}{E_2^2}\,\frac{\cos[(k_i-k_j)L]-1}{k_j-k_i}. \tag{14}$$

The average work done on the beam electrons per unit time is

$$\mathscr{A}=\int_{S_\perp}\langle A\rangle n_b(r)u\,dS_\perp. \tag{15}$$

The integral is taken over the transverse cross section of the cavity.

In order to calculate a specific value of the work $\mathscr{A}$, it is necessary (for finding the ratios A_i/A_2) to use the boundary conditions, in particular the conditions that there be no perturbations in the charge densities and beam current in the $z=0$ plane:

$$\sum_{\nu=1}^{4}(k_\nu^2-a^2)A_\nu=0,\quad \sum_{\nu=1}^{4}k_\nu^3A_\nu=0, \tag{16}$$

as well as two boundary conditions on the radiation field,

$$\sum_{\nu=1}^{4}k_\nu A_\nu=0,\quad \sum_{\nu=1}^{4}k_\nu A_\nu e^{ik_\nu L}=0. \tag{16a}$$

Here $a^2=\omega^2/c^2-k_\perp^2$. The first three of Eqs. (16) and (16a) yield the ratio A_ν/A_2 and the last, a. Here the wave numbers k_ν satisfy Eq. (8), which for a small Pierce parameter,

$$\chi\equiv\frac{\omega_b^2\gamma^{-3}}{k_\perp^2u^2}\ll 1, \tag{17}$$

has the solution[2]

$$k_{1,2}=\pm a\pm\frac{\beta_{1,2}}{2a}\omega_b^2,\quad \beta_{1,2}=\frac{k_\perp^2\gamma^{-3}}{(\omega\mp au)^2},$$

$$k_{3,4}=\frac{\omega}{u}\pm\alpha\omega_b,\quad \alpha=\frac{\omega}{u}\frac{\gamma^{-5/2}}{\sqrt{\omega^2-a^2u^2}},\quad a\neq 0. \tag{18}$$

Condition (17) ensures a Raman instability development,[1] and also assumes that there is no Pierce potential instability which would disrupt the current in the beam and not contribute to the radiation field.

Given the above remarks, we obtain the following expression for the work

$$\mathscr{A}=(-1)^{n+1}\frac{\omega_b}{2\pi}\|\phi_s\|\frac{\omega^2u\gamma^{-1/2}}{(\omega^2-a^2u^2)^{3/2}}\sin(\alpha\omega_bL)\sin\left(\frac{\omega L}{u}\right)X, \tag{19}$$

where

$$\|\phi_s^2\|=\int_{S_\perp}\phi_s^2(r)\,dS_\perp .$$

Here we have used the fact that $a=\pi n/L$ holds in the leadingorder. The main contribution to the work $\mathscr{A}$ is from the crossed beam-electromagnetic terms. This means that when the beam is modulated by the electromagnetic part of the radiated wave, most of the work is done by the beam waves. And, conversely, when the trajectory of an electron is perturbed by the beam wave, the maximum contribution to the work of the radiation field is from electromagnetic waves. Thus, the two oscillatory systems are coupled: the beam and the electromagnetic field in the cavity. The result of this interaction is a shift $\delta\omega$ in the frequency, whose imaginary part is nothing other than the growth rate of the instability. Here the effect is obviously collective, since the development of the instability requires that beam plasma waves be excited in the system.

For calculating the instability growth rate we write down the balance equation for the energy in the cavity,

$$\frac{d}{dt}\langle W\rangle=-\mathscr{A}, \tag{20}$$

where the time averaged energy $\langle W\rangle$ includes both the energy of the field and the kinetic energy of the beam electrons in the cavity, i.e.,

$$\langle W\rangle=\langle W_{el}\rangle+\langle W_e\rangle.$$

For small values of the Pierce parameter, $\chi\ll 1$, the energy contained in the cavity will be determined mainly by the energy of the electromagnetic field. In this case, Eq. (20) takes the form

$$\frac{dX}{dt}=2(-1)^n\omega_b\frac{k_\perp^2c^2\gamma^{-1/2}}{(\omega^2-a^2u^2)^{3/2}}\frac{u}{L}\times\sin(\alpha\omega_bL)\sin\left(\frac{\omega L}{u}\right)X. \tag{21}$$

For a cylindrical waveguide

$$\phi_s(r)=J_0(k_{\perp s}r),$$

where $k_{\perp s}=\mu_{0,s}/R$. ($\mu_{0,s}$ is the root of the Bessel function $J_0(x)$).

X is quadratic in the field, so that

$$\frac{dX}{dt}=2\,\delta\omega X. \tag{22}$$

Comparing Eqs. (21) and (22), we obtain the instability growth rate

$$\delta\omega=(-1)^n\omega_b\frac{k_\perp^2c^2\gamma^{-1/2}}{(\omega^2-a^2u^2)^{3/2}}\frac{u}{L}\sin(\alpha\omega_bL)\sin\left(\frac{\omega L}{u}\right) \tag{23}$$

at the frequency

$$\omega\equiv\omega_{s,n}=c\sqrt{\left(\frac{\mu_{s,0}}{R}\right)^2+\left(\frac{\pi n}{L}\right)^2}. \tag{24}$$

The condition for development of the instability,

$$(-1)^n\sin\left(\frac{\omega L}{u}\right)\sin(\alpha\omega_bL)>0, \tag{25}$$

determines the cavity and beam parameters for which electromagnetic waves can be amplified in the system.

To conclude this section, we note that Eqs. (23)–(25) are completely identical with those derived in Ref. 2 on the basis of an analysis of the dispersion relation, confirming the correctness of the proposed mechanism for the radiative Pierce instability as nonresonant, stimulated emission by a rectilinear relativistic electron beam in a cavity.

4. NONLINEAR THEORY OF THE INSTABILITY

The nonlinear system of Eqs. (1), (4), and (5) can be studied only by numerical simulation of the instability. According to the linear theory, several modes with similar frequencies are excited at once in thecavity. The absence of a distinct frequency, as well as wavelength, makes it impossible to separate the field into slowly and rapidly varying

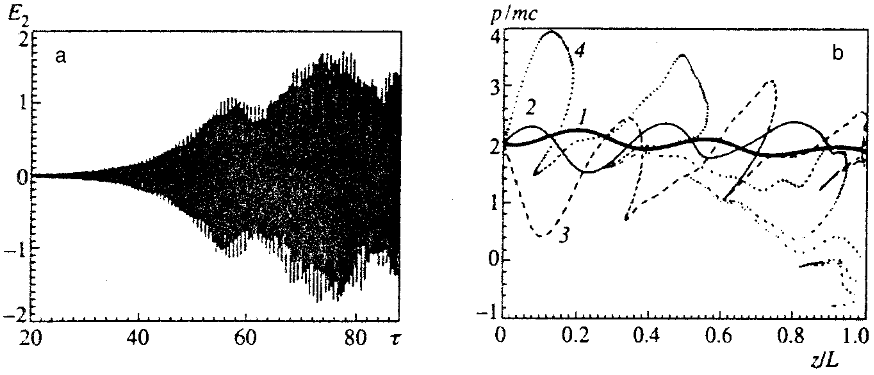

FIG. 1. (a) Dynamics of the wave amplitude in a short cavity, $\xi=6$ for $\gamma=2$ and $\chi=0.5$. (b) Phase planes of beam electronsin the cavity at different times $\tau=40$ (*1*), 45 (*2*), 60 (*3*), and 65 (*4*).

components with subsequent averaging over time or position. Thus, we found it necessary to use the methods described in Refs. 6 and 7. In our case, the beam was simulated by a particle method.[6] For better smoothing and elimination of sawtooth fluctuations in the density and charge, the core of the particle was chosen to have a triangular shape; the particle size coincided with the cell size (which can be regarded as a modification of the PIC method). From 10 to 40 particles belonged to asingle grid cell. An explicit scheme with step size changes was used to solve the Maxwell equation.[6] For studying long systems with $L/R\sim30$ an explicit-implicit Crank-Nicholson scheme with factorization with respect to the coordinate z was used. In the transverse direction, the field was expanded in the eigenfunctions of the waveguide. In order to eliminate grid fluctuations various methods of numerical filtering were used for the field components, as well as for the current density and charge of the beam.[7] The following dimensionless quantities were used in the numerical simulation:

$$\tau=\frac{u}{L}t,\quad p=\frac{p_z}{mc},\quad \varepsilon=\frac{eL}{mc^2\gamma^3}E_z. \tag{26}$$

In the numerical calculation, we considered an infinitely thin beam with a cylindrical geometry and used a soft regime for its entry into the cavity. The Pierce parameter χ was defined as the ratio of the working current I to the limiting Pierce current I_n, where

$$I_n=\frac{mc^3}{e}\frac{u^3}{c^3}\frac{\gamma^3}{2\ln(R/r_0)}. \tag{27}$$

In all the calculations we chose $r_0/R=0.4$. Here the instability saturated after 10–10000 transit times, depending on the beam current and the cavity geometric parameter $\xi=L/R$. In most cases the first transverse mode $s=1$ was excited. For very short systems $\xi\leqslant1$ and for some values of ξ in long systems, higher transverse modes were excited. For short systems, with $\xi<12$, the instability regime was mainly single-mode; for certain values of ξ two longitudinal harmonics corresponding to the selection rule (25) were excited. In long systems ($\xi>12$), a multimode regime sets in. In fact, for large ξ ($\xi\gg1$), the relation

$$\xi=\frac{1}{2}\frac{c}{u}\frac{\pi}{\mu_{s,0}}\frac{n_2^2-n_1^2}{\Delta\theta}, \tag{28}$$

holds, where n_i is the longitudinal mode number. Two modes with the same growth rates can be excited if the difference between their drift angles obeys $\Delta\theta<\pi$. Since Eq. (25) implies that either even or odd harmonics can be excited simultaneously in long systems, we find that for $\gamma=2$, when $\xi\geqslant3$, a two-mode regime exists and for $\xi\geqslant12$, a three-mode regime.

We have distinguished two mechanisms for saturation of the instability. The first mechanism occurs in short ($\xi<12$) systems with single-mode regimes. It is analogous to the Landau mechanism.[8] The four-wave process

$$\omega_{1,n}+\omega_{1,n}\rightarrow\omega_{1,n-1}+\omega_{1,n+1}, \tag{29}$$

which results in pumping of energy from the wave back into the beam in accordance with the condition (25), occurs in the system. In this case, for large field amplitudes in the cavity, the oscillations of the beam electrons relative to the equilibrium position cease to be harmonic.

Figure 1 shows the results of a calculation for $\xi=6$, $\gamma=2$ and $\chi=0.05$ (which roughly corresponds to a current $I\approx2.4$ kA). Regular oscillations in the field amplitude with a modulation frequency on the order of $\delta\omega$ in a steady-state saturation regime can be seen clearly. The position of the beam electrons in the phase planes is represented by Fig. 1b. For time $\tau=45$, when the instability is still linear, the modulation of the beam is purely harmonic. As the instability develops, nonlinear distortions appear, which cause breaking: the electrons begin to overtake each other and the beam stratifies in velocity. Breaking occurs whem the instability passes into a stationary regime. As the instability becomes saturated, the changes in the beam electron density acquire the character of a deep modulation. This corresponds to an increase in the ratio of the amplitudes of the beam waves to the electromagnetic wave and to a shift in the longitudinal wave numbers k_ν toward larger values. The beam remains cold, despite the rather long time after the radiation amplitude reaches its stationary level.

In long systems ($\xi\geqslant12$), when the instability is multimode from the outset, the mechanism responsible for saturation is the randomization of the beam particles in the field of many waves. As a result of the randomization of the particle trajectory, the modulation of the beam becomes uniform, i.e., the phases of the electromagnetic field relative to the electrons are distributed uniformly in the interval $[0;2\pi]$. Thus,

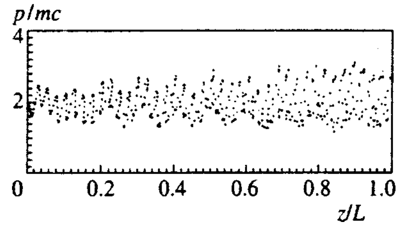

FIG. 2. Phase plane of beam electrons in a long cavity ($\xi=18$) at the time saturation sets in for $\gamma=2$ and $\chi=0.05$.

the contribution to the stimulated emission goes to zero. As the numerical studies showed, saturation of the instability sets in simultaneously with the chaos of the beam particles. Figure 2 shows the calculations for $\xi=18$, $\gamma=2$, and $\chi=0.05$. The phase plane of the beam electrons is shown for the time the instability saturates, $\tau=40$. By the middle of the cavity, the beam is completely randomized. Although the beam is highly chaotic, it is still modulated at the initial level. At the cavity exit the effective temperature of the beam is on the order of 0.1–0.3 MeV, which is not surprising, since the field strength in thecavity reaches 100–300 kV/cm. For this operating regime of the generator, the field in the cavity has a broad spectrum of longitudinal harmonics; higher harmonics are also excited in the transverse structure. Thus, for example, excitation of the third transverse mode was observed. Beam bunching was absent for both instability saturation mechanisms.

In order to estimate the degree of randomness in the oscillations of the system at the time when the stability is only beginning to reach saturation, two test particles were launched into the cavity separated by a rather small distance in the phase plane, with velocities roughly equal to the beam electron velocity u. The maximum distance by which the particles could separate in the phase plane as they pass through the cavity was chosen as a measure of the randomization of the beam. Figure 3 is a plot of the maximum phase separation between the particles,

$$D=\sqrt{\left(\frac{z}{L}\right)^2+\left(\frac{p_z}{mc}\right)^2},$$

as a function of the controlling parameter, which is the geometric factor ξ, for $\gamma=2$ and $\chi=0.05$. This graph shows clearly the existence of two saturation regimes for the instability which appear in regions consistent with the above estimates.

As the beam current is raised, and the Pierce parameter χ approaches unity (or greater), a Pierce potential instability develops in the system. Figure 4 shows a phase pattern of the beam at the time the instability saturates, when $\tau=10$ holds, for $\xi=18$, $\chi=0.95$, and $\gamma=2$. A virtual cathode is observed to form at the cavity entrance and the beam electrons are partially reflected from it. In the meantime, the randomization of the beam corresponds to the presence in the cavity of a radiative instability whose development is somewhat suppressed by the potential instability. As χ is increased further, the growth rate of the aperiodic Pierce instability increases more rapidly than that of the radiative instability. As a result, thepotential branch suppresses the radiative branch. Thus, the aperiodic and radiative Pierce instabilities can be regarded as two regimes of a single instability which are realized for different values of χ.

A straight relativistic electron beam is, therefore, capable of stimulated emission of electromagnetic waves in a smooth cavity. Naturally, the question of the efficiency with which the energy of the electron beam is transferred to the radiation field arises. Since the radiation is trapped inside the resonant cavity, we take the conversion efficiency of the beam electron energy to electromagnetic radiation energy to be the ratio of the radiation flux to the incident beam energy flux,

$$\eta=\frac{\langle|\mathbf{s}|\rangle}{mc^2nu\gamma}, \tag{30}$$

where $\langle|\mathbf{s}|\rangle$ is the magnitude of the Poynting vector of the electromagnetic wave near the right hand boundary of the waveguide, averaged over a long time interval $t\gg 2\pi/\omega$.

For short systems ($\xi<1$) the growth rate is small and is $\propto\omega_b^2$, so using a Pierce source for pulsed currents with low ξ is inefficient. The quantity $\eta(\xi)$ attains its maximum in long systems, with $1<\xi<8$ (Fig. 5). When the cavity length is

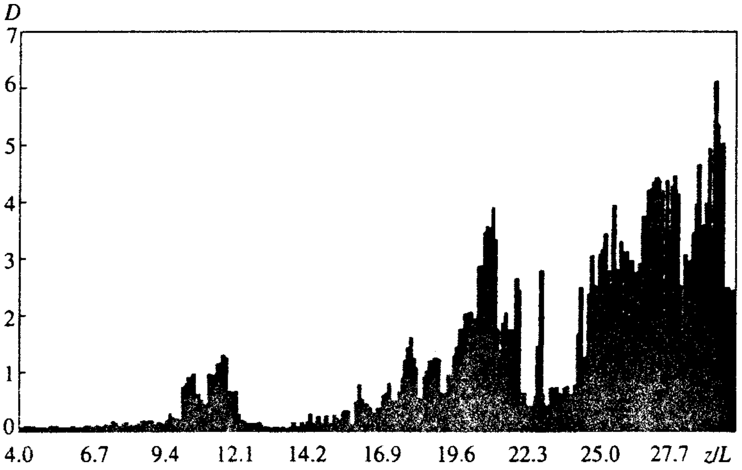

FIG. 3. The maximum distance between two test particles in the phase plane as a function of the geometric parameter ξ at the time the instability enters the nonlinear stage.

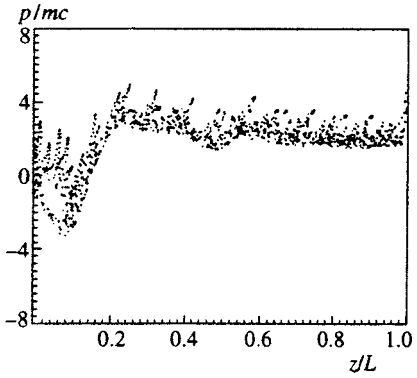

FIG. 4. Phase plane of the beam electrons at the time of saturation for $\xi=25$, $\chi=0.95$, and $\gamma=2$.

increased further, the conversion efficiency falls off rapidly, in agreement with the linear theory.[2] Therefore, ξ of order five is the optimum.

Numerical simulation yielded the dependence of the conversion coefficient η on the beam current for different values of the geometric parameter ξ. As the current is increased, η decreases because of the heating of the beam. Heating sets in faster in long cavities, and this leads to more rapid drop in theconversion efficiency as the beam current is raised.

The dependence of η on the relativistic character (γ) of the beam is consistent with the linear theory.[2] As γ increases, the energy conversion efficiency η initially increases quadratically with γ but then saturates for $\gamma>5$ (Fig. 6).

5. DYNAMIC INSTABILITY OF THE BEAM PARTICLE MOTION

As the instability reaches the nonlinear stage, the field amplitude continues to rise slowly. The regular amplitude modulations in the field are replaced by random oscillations. This is because, as the instability develops further, the frequency spectrum of the oscillations broadens(Fig. 7) owing to nonlinear many-wave processes which cause a redistribution of the energy in the radiation spectrum.

$$\omega_{1,n}+\omega_{1,n}\to\omega_{1,n-m}+\omega_{1,n+m}, \tag{31}$$

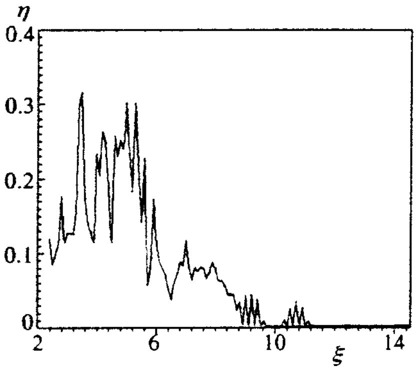

FIG. 5. Energy conversion coefficient η as a function of the geometric parameter of the cavity for beam parameters $\gamma=2$ and $\chi=0.05$.

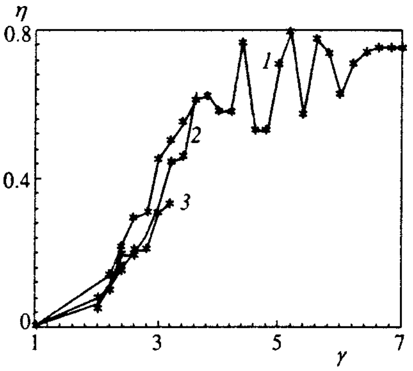

FIG. 6. Energy conversion coefficient η as a function of the relativistic characteristic γ of the beam for $\chi=0.05$ and $\chi=4$ (*1*), 5 (*2*), and 6 (*3*).

in terms of the integer $m<n$.

A monotron is an open system, since it exchanges energy with the surroundings by means of an electron beam. The development of an instability assumes the existence of strong feedback. Naturally, in such a system there is a dy-

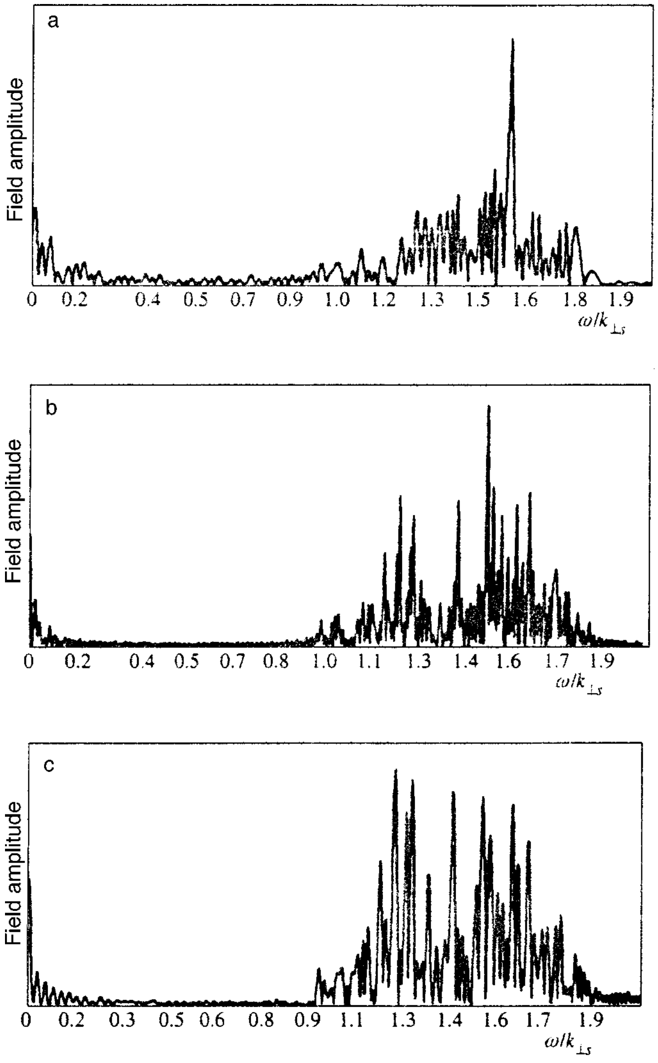

FIG. 7. Fourier spectrum of the electromagnetic oscillations at different times: (a) at the time the instability saturates, (b) in the nonlinear stage and (c) advanced nonlinear stage.

namic instability of the motion—an exponential spread in the particle trajectories,[9] which was observed in the nonlinear stage of the numerical calculations for long systems with $\xi \geqslant 8$. The development of the instability is, indeed, accompanied by a broadening of the spectrum of the oscillations. It is convenient to take the *K*-entropy, which is defined by the expression[9]

$$K=\lim_{t\to\infty}\lim_{D(0)\to 0}\frac{1}{t}\ln\frac{D(t)}{D(0)}, \tag{32}$$

as the characteristic of the dynamic chaos. Since the time the particles spend in the cavity is limited and is on the order of L/u, the *K*-entropy can be calculated approximately using the formula

$$K=\frac{1}{T}\ln\frac{D_{\max}(T)}{D(0)}, \tag{33}$$

where T is the transit time of the particles. The *K*-entropy obtained in this way for different values of the geometric parameter was a decreasing function of ξ. The *K*-entropy is low for $\xi>8$ because randomization of the beam is responsible for saturation of the instability at these values of the geometric parameter. Thus, the level of chaos which develops in the linear stage remains the same even in the advanced nonlinear stage. As noted above, for $\xi<8$ the mechanism responsible for stabilizing the instability is not associated with heating of the beam, so stronger amplification of the electromagnetic waves and, therefore, stronger subsequent heating of the beam are possible.

During the numerical calculations, for some values of ξ an intermittence between the ordered and chaotic regimes was observed in time. Besides the regime instability, stratification of the beam into two components was observed: cold, in which the particle motion was ordered, and hot, in which the particles were subject to a dynamic instability. Here the transit time through the cavity for the cold particles was considerably shorter than the residence time in the cavity for the hot particles.

6. CONCLUSION

Based on the above results, we can answer the question of why monotron generators were inefficient. All these devices operated in a parameter range where weak generation occurred, specifically, with $\gamma\simeq 1$ (nonrelativistic beam) and $\xi=L/R\ll 1$ (ultrashort systems). Numerical simulation showed that for optimum generation, the device efficiency can be on the order of 20–30%.

There is yet another parameter which can be used to regulate the generator efficiency—the radius r_b of a thin cylindrical beam. Infact, the electrons with transverse coordinate r_b are modulated by a radiation field proportional to $\phi_s(r_b)$. Since the entry into the steady state is determined by the modulation depth of the beam, the saturation amplitude will depend on the beam radius r_b.

Finally, we note that the existence of two different instability regimes opens up the possibility of creating both wide and narrow band generators. The fact that the regimes depend only on the geometric parameter ξ of the cavity makes an attempt to construct a tuneable generator tempting.

We thank V. A. Cherepenin for discussing the problems associated with numerical integration of Maxwell equations and for valuable comments.

*)E-mail: rukh@fpl.gpi.ru

---

[1] M. V. Kuzelev and A. A. Rukhadze, *Electrodynamics of Dense Electron Beams in Plasmas* [in Russian], Nauka, Moscow (1990).

[2] D. N. Klochkov and A. A. Rukhadze, Fizika Plazmy **23**, 646 (1997) [Plasma Phys. Rep. **23**, 598 (1997)]; D. N. Klochkov and M. Yu. Pekar, Fizika Plazmy **23**, 650(1997) [Plasma Phys. Rep. **23**, 602 (1997)]; D. N. Klochkov, M. Yu. Pekar, and A. A. Rukhadze, Radiotekhnika i Élektronika **44** (1999).

[3] F. Biquard, P. Grivet, and A. Septier, IEEE Trans., Dec., Im- 17, No. 4, 354 (1968).

[4] J. J. Muller and E. Rostas, Helv. Phys. Acta **13**, 435 (1940).

[5] J. R. Pierce, J. Appl. Phys. **15**, 721 (1944).

[6] Yu. A. Berezin and V. A. Vshivkov, *Particle Methods in Low-Density Plasma Dynamics* [in Russian], Nauka, Novosibirsk (1992).

[7] V. P. Tarakanov, in *Users Manual for Code KARAT*, Berkley Research. Assoc., Inc., Springfield, VA, USA (1992).

[8] L. D. Landau and E. M. Litshitz, *Fluid Mechanics*, Pergamon, Oxford (1987) [Russian original, Nauka, Moscow (1986)].

[9] Yu. L. Klimontovich, *Turbulent Motion and Structural Chaos* [in Russian], Nauka, Moscow (1990).

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Observation of ions with energies above 100 keV produced by the interaction of a 60-fs laser pulse with clusters

S. Dobosz, M. Schmidt, M. Perdrix, P. Meynadier, O. Gobert, D. Normand, K. Ellert, and T. Blenski

CEA-DSM/DRECAM/SPAM, CE Saclay, Bât. 524, 91191 Gif-sur-Yvette Cedex, France

A. Ya. Faenov, A. I. Magunov, T. A. Pikuz, and I. Yu. Skobelev*⁾

Multiply-Charged Ion Spectral Data Center, All-Russian Institute of Physicotechnical and Radio Engineering Measurements, 141570 Mendeleevo, Moscow Region, Russia

N. E. Andreev

High Energy Density Research Center, Joint Institute of High Temperatures, Russian Academy of Sciences, 127412 Moscow, Russia

(Submitted 18 August 1998)
Zh. Éksp. Teor. Fiz. **115**, 2051–2066 (June 1999)

The x-ray spectra of a plasma generated by heating CO_2 and Ar clusters with high-intensity femtosecond laser pulses with $q_{\text{las}} \simeq 10^{18}\,\text{W/cm}^2$ are investigated. Spatially resolved x-ray spectra of a cluster plasma are obtained for the first time. Photoionization absorption is observed to influence the spectral line profiles. The recorded features of the x-ray emission spectra definitely indicate the existence of a large relative number of excited ions ($\simeq 10^{-2}-10^{-3}$) with energies of $0.1-1$ MeV in such a plasma. Possible mechanisms underlying the acceleration of ions to high energies are discussed. It is shown that the experimental results can be attributed to the influence of ponderomotive forces in standing waves generated by the reflection of laser radiation from the clusters.
[S1063-7761(99)01106-3]

1. INTRODUCTION

The fabrication of lasers that generate subpicosecond pulses at a power of several terawatts has unleashed new possibilities in research on the interaction of laser radiation with matter. In addition to the study of fundamental problems in the physics of high energy densities, various applied problems are of major interest as well, for example, the development of efficient sources of both incoherent and coherent x-rays utilizing a laser plasma, or the acceleration of charged particles. Together with the solid-state targets conventionally used in experiments involving picosecond and nanosecond laser pulses, targets consisting of agas jet expanding in vacuum have come into widespread use for ultrashort pulses. The practical advantages of these targets are ease of control, simplicity of replacement, and good repeatability. Unlike solid-state targets with a multitude of plasma-generating processes, the plasma of a low-density gas target is generated for the most part by tunneling or above-barrier ionization in a strong optical field.

A new class of targets has emerged in recent years: so-called cluster targets, i.e., gas jets containing clusters that consist of a significant number (10^3-10^7 of molecules or atoms having near-solid densities at diameters of 10–100 nm. Such targets, on the one hand,have the practical advantages of gas targets and, on the other, can be used to investigate the kinds of processes typical of high material densities. A great many important phenomena have been discovered by means of cluster targets, including the self-focusing of a laser beam in a plasma,[1–4] the formation of multiply-charged ions with completely empty inner shells (hollow ions),[1–4] and the generation of higher harmonics[5] and ions with energies above 1 Mev.[6,7]

It is important to note that the details of the interaction of ultrashort laser pulses with clusters are not at all clear at this point. There are several models at the present time, each with a different description of the ionization of cluster targets.[1,8,9] The data available from mass spectrometer measurements are too meager to explain the mechanism underlying the generation of fast ions in such a plasma. Their occurrence can be attributed to the Coulomb explosion of a molecule or cluster and to hydrodynamic (collective) processes.

In principle, the presence of fast ions is detectable not only by mass spectrometer techniques, but by x-ray spectral methods as well. The instrumentation for the latter is simpler, and they can be used to systematically amass experimental data on plasma processes. In fast-ion detection applications, the sensitivity threshold of the x-ray spectral method is lower than in mass spectrometer measurements. However, if the lasing efficiency is high enough to produce a reasonably large relative number of fast particles, their influence can be detected in the plasma emission spectrum.

In this paper we show that a large relative population of ions ($10^{-2}-10^{-3}$) with energies greater than 100 keV is generated in the plasma when femtosecond laser pulses with a peak intensity of $10^{18}\,\text{W/cm}^2$ interact with gas targets con-

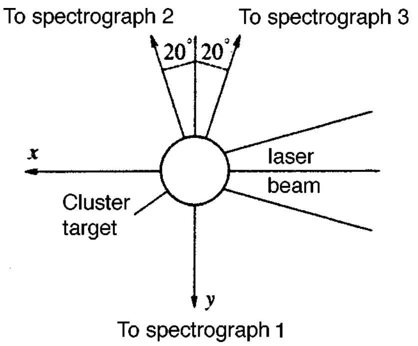

FIG. 1. Experimental layout.

taining CO_2 and Ar clusters. The presence of these ions shows up in the x-rayspectral lines.

2. EXPERIMENTAL SETUP AND RESULTS

Experiments have been carried out with the laser system at the Saclay Research Center (Center d'Études de Saclay, Commissariat à l'Énergie Atomique, Direction de Sciences de la Matière) in France. The plasmawas heated by a titanium–sapphire laser ($\lambda_{las} = 0.8\,\mu m$) with a pulse duration of 60 fs and energy of 70 mJ. Focusing of the radiation by a parabolic mirror produced a radiant flux density up to $10^{18}\,W/cm^2$. A cluster target was formed by the adiabatic expansion in vacuum of a comparatively small burst of gas (CO_2 or Ar) issuing from a gas valve through a pulsed nozzle of diameter 0.3 mm. The gas pressure in the valve was 10–40 atm.

The plasma-emitted x-rays were detected simultaneously by three spectrographs with spherically concave mica crystals.[10] The relative positions of the laser beam, the gas target, and the spectrographs are shown schematically in Fig. 1. The radii of curvature of the crystals were $R=100$ mm (spectrographs 2 and 3) and $R=150$ mm (spectrograph 1). Spectrographs 2 and 3 provided a spectral resolution of $\lambda/\Delta\lambda = 7000$ and spatial resolution $\delta x \sim 25\,\mu m$. These parameters were somewhat better for spectrograph 1: $\lambda/\Delta\lambda = 10\,000$ and $\delta x \sim 15\,\mu m$.

For the CO_2 plasma the spectral intervals 17–17.8 Å and 15.8–16 Å were detected in the first reflection order from the crystal; they contain the $1snp\ ^1P_1 - 1s^2$ lines ($n \geqslant 4$) of He-like O VII and the $3p-1s$ line of H-like O VIII, respectively. The emission spectra of the Ar plasma were detected in the fifth andfourth reflection orders from the crystal. The corresponding spectral intervals 3.35–3.5 Å and 3.9–4.2 Å contained the $1s\,3p - 1s^2$ lineof He-like Ar XVII and its Li-like dielectronic satellites and the $1s\,2p\ ^{1,3}P_1 - 1s^2$ lines of Ar XVII with satellite structures due to transitions in Li-, Be-, and B-like argon ions. Examples of the recorded spectra of the CO_2 and Ar plasmas are shown in Fig. 2.

Figure 2 reveals the sharp distinction between the emission spectra of the CO_2 and Ar plasmas in connection with the strong asymmetry of the spectral line profiles of the oxygen ions. This asymmetry in theprofiles of the oxygen ion lines, which is also observed in the lines of He-like O VII and the $3p-1s$ line of H-like O VIII, cannot be identified either with a thermal Doppler mechanism or with a Stark mechanism of spectral line broadening in the plasma. In the next section we propose a simple model, which provides a consistent explanation of the observed line profiles of the oxygen (asymmetric) and argon (symmetric) ions. The principal attributes of this model are, first, line broadening due to macroscopic motion (expansion) of the plasma and, second, the onset of asymmetry of the line profiles of the oxygen ions as a result of the photoionization of unheated clusters and the presence of H-like and He-like carbon ions in the CO_2 plasma.

3. SPECTRAL LINE PROFILES IN AN EXPANDING DENSE PLASMA

We assume that the interaction of a femtosecond laser pulse with a gas jet generates a plasma in a spatial region of characteristic length l_{las} in the vicinity of the focal spot. The generated plasma then expands into the surrounding medium, which contains unheated (or slightly warmed) clusters. At time $t \gg \tau_{las}$ (τ_{las} is the duration of the laser pulse) we have a mixture of unheated clusters and outwardly dispersing ions and electrons. This complex object has a characteristic length $L_{mix} \gg l_{las}$, and the velocity distribution of the ions in it is highly nonuniform. Specifically, high-velocity ions are situated closer to the boundary of the object, whereas slower ions are localized close to its center. In other words, ions with different velocities are spatially separated in the given object. If we assume that ions are acceleratedonly at the time of incidence of the laser pulse, then at $t \gg \tau_{las}$ an ion with velocity v is situated at a distance $y = vt$ from the center of the plasma (i.e., from the focus of the laserbeam). In this regard we consider the following one-dimensional model, which,on the one hand, reflects all the principal features of the object and, onthe other, can be used to derive simple expressions for its spectral characteristics.

Let the y axis be directed toward the recording spectrograph (spectrograph 1 in Fig. 1). We denote by $N_i(y)$ the density of ions in the excited state i situated at a distance y from the centerof the plasma (i.e., from the focus of the laser pulse), and by $v(y)$ the velocity of these ions; in keeping with the foregoing discussion, $v(y)$ is a linear function: $v(y) = by$, where $b = \text{const}$. Let us consider the spectral line associated with the radiative transition $i \to k$ with probability A_{ik} and frequency ω_{ik}, and let its profile at $y=0$ (i.e., for ions at rest on the average) be characterized by the symmetric function $S_{ik}(|\omega - \omega_{ik}|)$. Inasmuch as the spectrograph records the total radiation from all ions, the observed line profile in an optically thin plasma is described by the integral

$$F_{ik}^{\text{obs 1}}(\omega - \omega_{ik}) = \frac{A_{ik}\int_{-\infty}^{\infty} N_i(y) S_{ik}(|\omega - \omega_{ik}(1+by/c)|)\,dy}{A_{ik}\int_{-\infty}^{\infty} N_i(y)\,dy}, \tag{1}$$

where the factor $(1+by/c)$ in the argument of S_{ik} is associated with the Doppler shift in frequency.

It is readily apparent from Eq. (1) that if the function $N_i(y)$ is symmetric, the observed line profile $F_{ik}^{\text{obs}}(\omega - \omega_{ik})$ will be symmetric about the frequency ω_{ik}. Consequently, the expansion of an optically thin plasma can impart asym-

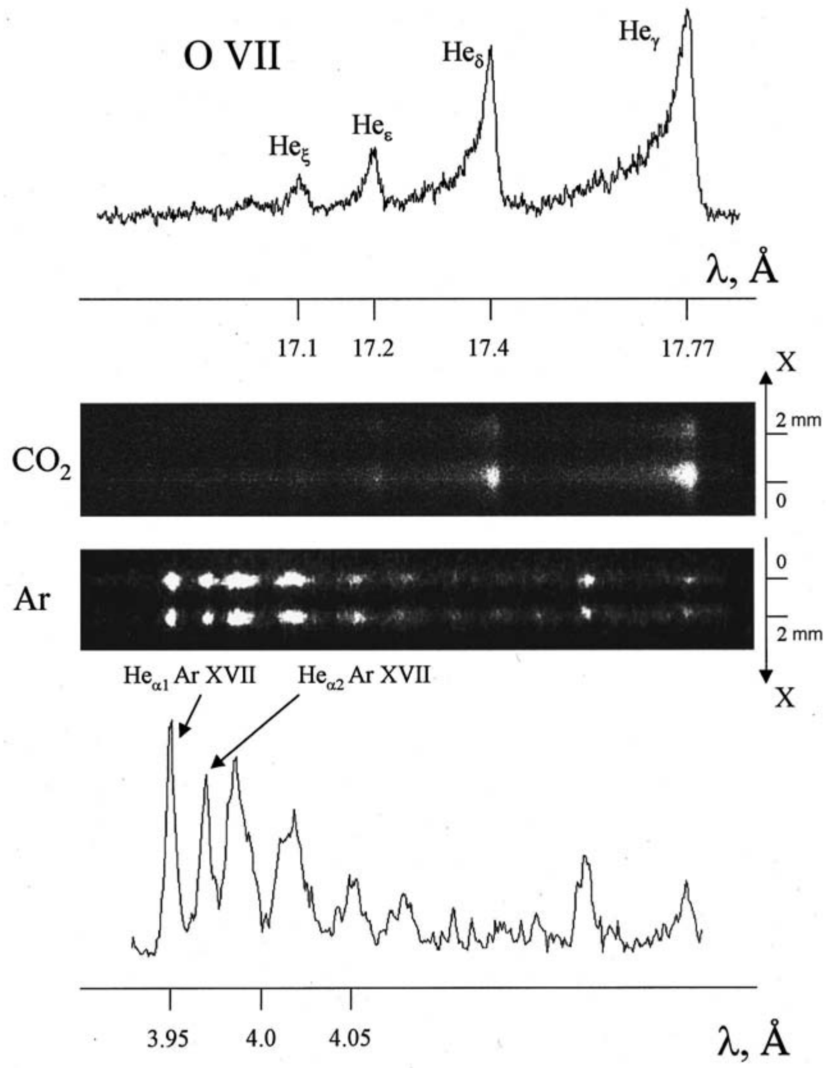

FIG. 2. Emission spectra of CO_2 (upper) and Ar (lower) plasmas produced by the interaction of a femtosecond laser pulse with cluster targets. The spectra were obtained with one-dimensional spatial resolution along the *x*-axis, which defines the direction of propagation of the laser beam (see Fig. 1).

metry to the observed line profiles only when the ion distribution $N_i(y)$ itself is asymmetric. But then the asymmetry must be detected differently by spectrographs aimed in different directions toward the plasma. If, for example, the second spectrograph is situated in the region $y<0$ (spectrographs 2 and 3 in Fig. 1), it will observe a spectrum $F_{ik}^{\mathrm{obs}\,2}(\omega-\omega_{ik})$ given by

$$F_{ik}^{\mathrm{obs}\,2}(\omega-\omega_{ik})=F_{ik}^{\mathrm{obs}\,1}(-\omega+\omega_{ik}), \tag{2}$$

which implies that the line profile detected by the second spectrograph is obtained from the line profile detected by the first via symmetric reflection about the frequency ω_{ik}, i.e., if, for example, the long-wavelength wing of the line is suppressed in one spectrum, the opposite, short-wavelength wing will be suppressed in the other.

Equations (1) and (2), which have been obtained for an optically thin plasma, cannot account for the asymmetric profiles of the lines emitted by the CO_2 plasma because in the first place, there is no reason to believe that the expansion is not symmetric about the axis of the laser beam, and even more importantly, the spectra detected by spectrographs aimed at the plasma in opposite directions have the same form, in direct conflict with Eq. (2).

We now inquire how the line profile equation (1) is modified for a plasma with absorption. We assume an absorption mechanism such that in a narrow frequency band $\Delta\omega/\omega\ll 1$ the absorption coefficient $k(y)$ does not depend on ω (properties of this kindare found, for example, in absorption by inverse bremsstrahlung). The spectral line profile is then given by

$$F_{ik}^{\mathrm{obs}}(\omega-\omega_{ik})=\frac{\int_{-\infty}^{\infty}\exp\left\{-\int_{y}^{\infty}k(x)\,dx\right\}N_i(y)A_{ik}S_{ik}\left(\omega-\omega_{ik}\left(1+\frac{by}{c}\right)\right)dy}{\displaystyle\int_{-\infty}^{\infty}\exp\left\{-\int_{y}^{\infty}k(x)\,dx\right\}N_i(y)A_{ik}\,dy}. \tag{3}$$

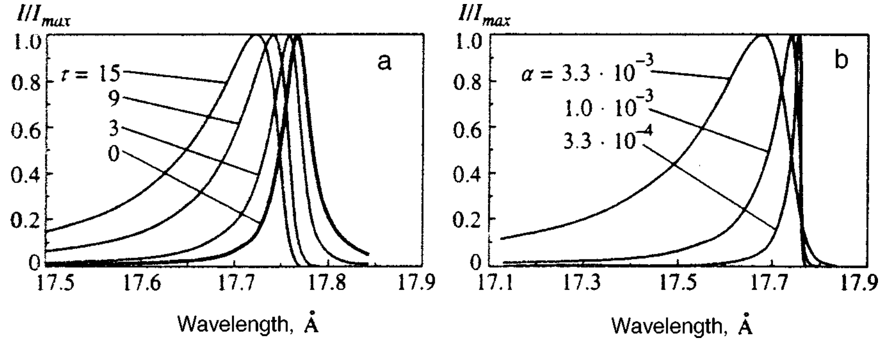

FIG. 3. a) Profile of the $1s\,4p\ {}^1P_1-1s^2\ {}^1S_0$ line of the He-like O VII ion, calculated from Eq. (6) for various optical depths of the plasma τ and values of the parameter $\alpha=10^{-3}$. b) Profile of the $1s\,4p\ {}^1P_1-1s^2\ {}^1S_0$ line of the He-like O VII ion, calculated from Eq. (6) for an optical depth of the plasma $\tau=9$ and various values of the parameter α.

It is evident from Eq. (3) that the factor $\exp\{-\int_y^\infty k(x)dx\}$ in the integrand imparts asymmetry to the observed line profile even when the functions $k(y)$, $N_i(y)$, and $S_{ik}(\omega-\omega_{ik})$ themselves are symmetric. Indeed, allowing for the fact that the main contribution to the integral is from the region where the argument of the function S_{ik} vanishes, from Eq. (3) we obtain the estimate

$$\frac{F_{ik}^{\mathrm{obs}}(-\omega+\omega_k)}{F_{ik}^{\mathrm{obs}}(\omega-\omega_{ik})}\approx\exp\left\{-\int_{-\frac{(\omega-\omega_{ik})c}{\omega_{ik}b}}^{+\frac{(\omega-\omega_{ik})c}{\omega_{ik}b}}k(y)\,dy\right\}<1$$

$$\text{for}\quad \omega>\omega_{ik}. \tag{4}$$

Equation (4) implies that the long-wavelength wing of the line inan optically thick plasma is suppressed in comparison with the short-wavelength wing. This result is physically obvious: long-wavelength photons are emitted mainly by ions that move away from the spectrographand are therefore farther from it. To enter the spectrograph, these photons must pass through a large volume of the plasma and, hence, undergo greater absorption. We also note that Eq. (4) still holds for a spectrograph aimed toward the plasma from the opposite direction, owing to the overall symmetry of the problem for symmetric distribution functions $N_i(y)$ and $k(y)$.

Further simplification of Eq. (3) requires specification of the functions $N_i(y)$, $k(y)$, and $S_{ik}(\omega-\omega_{ik})$. We note above all that the width of the distribution $S_{ik}(\omega-\omega_{ik})$, being primarily associated with thermal Doppler broadening, is much less than the width of the observed line profile under our experimental conditions (even at an ion temperatures $T_i=1$ keV the Doppler width of the oxygen ion lines is $\Delta\omega_{ik}^D/\omega_{ik}\simeq 6\times10^{-4}$, in contrast with the observed widths $\Delta\omega_{ik}^{\mathrm{obs}}/\omega_{ik}\simeq 3\times10^{-3}$. This difference permits us to regard $S_{ik}(\omega-\omega_{ik})$ as a δ-function. The absorption coefficient $k(y)$ and the populationof excited levels $N_i(y)$ have different functional forms in general,because $k(y)$ is determined primarily by the density of unheated clusters andions in the ground state $N_0(y)$, and even though the quantities $N_i(y)$ are proportional to $N_0(y)$, the proportionality factor is a complex function of both the density and the temperature of the plasma. In our simplemodel we assume that this coefficient is constant, whereupon the form of the functions $k(y)$ and $N_i(y)$ is determined by the distribution of the plasma density, which we model as $N_0/(1+(y/y_0)^2)$:

$$N_i(y)=\frac{a_{ik}N_0}{1+(y/y_0)^2},$$

$$k(y)=\frac{b_{ik}N_0}{1+(y/y_0)^2}. \tag{5}$$

In this case the observed line profiles are described by the equation

$$F_{ik}^{\mathrm{obs}}(\omega-\omega_{ik})$$

$$=M\frac{\exp\{-\tau_{ik}[1/2-(1/\pi)\tan^{-1}((\omega/\omega_{ik}-1)/\alpha)]\}}{1+[(\omega/\omega_{ik}-1)/\alpha]^2}, \tag{6}$$

where $\alpha=by_0/c$, $\tau_{ik}=b_{ik}N_0y_0\pi$ is the optical depth of the plasma, and

$$M^{-1}=\int_{-\infty}^{\infty}\frac{\exp\{-\tau_{ik}(1/2-1/\pi)\tan^{-1}(y/y_0)\}}{1+(y/y_0)^2}dy$$

is a normalization factor independent of ω. It is evident from Eq. (6) that for our simple model the line profile depends on onlytwo parameters: α and τ_{ik}. The parameter α occurs in Eq. (6) only in combination with the frequency mismatch and characterizes the width of the line emitted by an optically thin,expanding plasma. We emphasize that the values of the parameter α are identical for all spectral lines. The parameter τ_{ik} is the optical depth of the plasma at frequency ω_{ik}, and characterizes the asymmetry of the observed line profile. Its valuescan differ for different spectral lines, but if a wideband absorptionmechanism prevails (free–free or bound–free transitions), these differences are extremely small.

Line profiles plotted from Eq. (6) for the $1s\,4p\ {}^1P_1-1s^2$ line of the O VII ion for various values of the parameters α and τ are shown in Fig. 3. It is evident from the figure that, first, appreciable deformation of the line profiles sets in by $\tau\simeq3$, second, absorption suppresses not only the long-wavelength wing, but also the center of the line, shifting the observed maximum of the line to shorter wavelengths, and third, absorption significantly increases the observed linewidth. Comparing the model calculations of Fig. 3 with the experimental spectra in Fig. 2, we see that the proposed simple model yields a qualitatively correct description ofboth the argon plasma spectrum (at $\tau=0$) and the CO_2 plasma spectrum(at $\tau\geqslant3$). The only remaining dilemma is why the CO_2 plasma in our experimental situation was optically thick, whereas the Ar plasma was optically thin, even though the average parameters of the plasma (density and size) were approximately the same in the two cases. To resolve the issue, we analyze possible x-ray absorption mechanisms in the plasma.

X rays can be absorbed in the plasma by free–free (inverse bremsstrahlung), bound–free (photoionization), and bound–bound (photoexcitation) transitions. To estimate the efficiencies of the various absorption mechanisms, we need information about such parameters of the plasma as its electron and ion densities, temperature, charge composition, and size. In the CO_2 plasma, the cluster sizes range from 80–100 nm; they are separated by 300–400 nm, and the particle density number is 8×10^{21} cm^{-3}. This means that the average density of carbon ions in the plasma generated in the focal spot (with a characteristic length $l_{\rm las}\simeq20$ nm) is $N_{\rm C}\simeq(0.64-3)\times10^{20}\,{\rm cm}^{-3}$, and the density of oxygen ions is $N_{\rm O}=2N_{\rm C}\simeq(1.3-6)\times10^{20}\,{\rm cm}^{-3}$. It follows from the x-ray spectral measurements that H-like and He-like O VIII and O VII ions occur inequal amounts in the plasma. It is natural to assume that carbon ions will be predominantly H-like. We therefore infer an estimate of the average electron density of the plasma: $N_e\simeq(1.2-5.5)\times10^{21}\,{\rm cm}^{-3}$. The electron temperature of the plasma can be estimated, on the one hand, from the ratio of H-like and He-like O VIII and O VII ions, and on the other, from the intensities of the resonance line of the He-like Ar XVII ion and its dielectronic satellites: $T_e\simeq200-700$ eV. During expansion the density of the plasma decreases, and at the instant that it expands to a length $L_{\rm mix}\simeq300-400\,\mu$m the densities of electrons and multiply-charged ions become three orders of magnitude lower than their initial values.

We now consider absorption associated with bound–bound transitions, i.e., resonant photoabsorption or self-absorption of the spectral lines. The cross section of this process is greater than the cross sections of free–free or bound–free transitions, and for lines of the resonanceseries the laser plasma is normally optically thick, i.e., $\tau^{bb}\gg1$. Two important conclusions can be drawn from the resonant character of this process.

First of all, the efficiency of the process declines when large velocity gradients are present in the expanding plasma. In fact, as a result of the Doppler shift associated with the translational motion of the plasma photons emitted in the spatial region y can be absorbed only in parts of the plasma close to this point with a length $\Delta y\simeq(c/b)(\Delta\omega_{ik}^D/\omega_{ik})$, where $\Delta\omega_{ik}^D$ is the intrinsic linewidth, i.e., the width of the function $S_{ik}(\omega-\omega_{ik})$. For large velocity gradients, i.e., large values of b, Δy is much less than the length of the plasma L, so that the optical depth τ^{bb} decreases by a factor $\sim L/\Delta y$. The values of τ_{ik}^{bb} can be estimated for this case from theequation

$$\tau_{ik}^{bb}=N_k\lambda_{ik}^3\frac{g_i}{g_k}\frac{A_{ik}}{b}. \tag{7}$$

We infer from Eq. (7) that the optical depth τ^{bb} for the middle of the $1s\,4p\,{}^1P_1-1s^2$ line of O VII at the instant of plasma generation is

$$\tau^{bb}(1s\,4p\,{}^1P_1-1s^2)\simeq30-100.$$

As the plasma expands, τ^{bb} rapidly decreases as a result of the decreasing ion density.

Second, photoabsorption can result in deformation of the spectral line profiles across essentially the full width of the function $S_{ik}(\omega-\omega_{ik})$ (see, e.g., Refs. 11 and 12) or, more precisely, in the range

$$|\omega-\omega_{ik}|\leqslant(1/2)\Delta\omega_{ik}^D\sqrt{\ln\tau_{ik}^{bb}},$$

which is considerably narrower than the total width of the observed line profile.

We now discuss inverse bremsstrahlung, i.e., absorption associated with free–free transitions. The optical depth $\tau^{ff}(\lambda)$ of theplasma due to this mechanism can be estimated to be[13]

$$\tau^{ff}\approx1.3\times10^{-49}\lambda^3\frac{Z^2N_iN_e}{\sqrt{T_e}}L, \tag{8}$$

where L is measured in cm, λ in Å, T_e in eV, and both N_i and N_e in cm^{-3}. For the lines of O VII and O VIII ions in the range 16–18 Å we obtain[7]

$$\tau_0^{ff}\simeq(0.06-3.3)\times10^{-4}\ll1,$$

and for the lines of Ar XVII ions ($\lambda\simeq3-4$ Å) we have

$$\tau_{\rm Ar}^{ff}\simeq(0.03-2.5)\times10^{-5}\ll1.$$

Consequently, inverse bremsstrahlung is weak both in the CO_2 plasma and in the Ar plasma.

Next we consider absorption associated with bound–free transitions, i.e., with photoionization. The main contributing factor is photoionizationfrom the ground state of an ion, because the populations of the excited levels are normally small. This phenomenon is a threshold process in the photon energy, since photons with energies above the ionization potential are absorbed. At energies well above threshold, the cross section of the process is approximately[13]

$$\sigma^{bf}=0.55\times10^{-16}\frac{1}{Z^2}\left(\frac{I}{\hbar\omega}\right)^{7/2}\ {\rm cm}^2, \tag{9}$$

where I and Z are the ionization potential and the spectroscopic symbol of the absorption-inducing ion. We note that although Eq. (9) has been derived for photoionization of the ground state of an H-like ion, it can be used for order-of-magnitude estimates of the photoionization cross sections of multielectron ions.

The threshold nature of photoionization absorption explains why this mechanism does not show up in the emission spectra of a chemically homogeneous plasma. Indeed, the spectral lines of the most conspicuous ions in the plasma can ionize only ions of lower multiplicity, which have a low density in the plasma, and despite the large cross sections σ^{bf}, the photoionization absorption cross section is not very large. This situation is encountered, for example, in our experiments with the Ar plasma. The spectral lines of the Ar X-

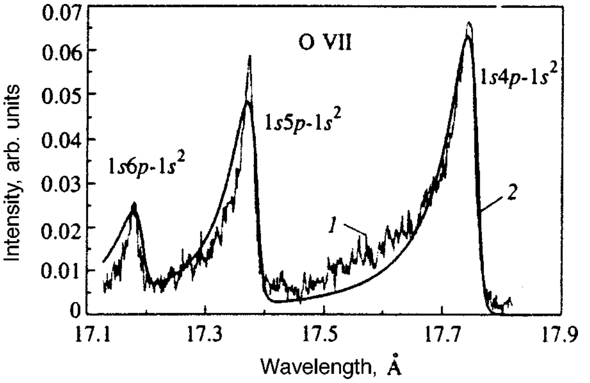

FIG. 4. Comparison of the model spectrum (6) (heavy curve) with the experimental result for the He-like lines of the H-like O VII ion.

VII and Ar XVIII ions can only lead to photoionization of the Li-like Ar XVI ion and ions of lower multiplicity. If the abundances of H-like and He-like ions are approximately the same (as in our experimental conditions), the abundance of even the Li-like Ar XVI ion will not exceed 0.1, and the optical depth of the plasma is

$$\tau^{bf}_{\mathrm{Ar}}\simeq 10^{-3}-10^{-2}\ll 1.$$

A different situation can arise in a plasma of complex chemical (mixtureof ions of different chemical elements) or ionization (mixture of a multiply-charged plasma and unheated clusters) composition. In this case the spectral lines associated with ions of one chemical element can photoionize not only ions of the other element, which can have a very high density in the plasma, but also neutral atoms contained in unheated clusters. For example, in our investigated CO_2 plasma the He-like O VII lines can photoionize the H-like C VI ion as well as C I and O I atoms. An estimate of the optical depth $\tau^{bf}_{CO_2}$ of the plasma associated with both processes can be obtained from Eq. (9): $\tau^{bf}_{CO_2}\simeq 0.83-5.3$. Of course, absorption by unheated clusters is also possible in the case of the Ar plasma, but the estimate $\tau^{bf}_{\mathrm{Ar}}<1$ is still valid by virtue of the somewhat smaller photoionization cross sections. Consequently, photoionization absorption should lead to modification of the emission spectra of the CO_2 plasma and should not affect the spectra of the Ar plasma, as we have indeed observed in our experiments.

As mentioned, in our model (6) the profile of each spectral line depends only on two parameters. The parameter α is identical for all lines, and according to (9) the values of τ^{bf}_{ik} for different lines are related by the equation

$$\tau^{bf}_{i_1k_1}\omega^{7/2}_{i_1k_1}=\tau^{bf}_{i_2k_2}\omega^{7/2}_{i_2k_2}.$$

Altogether, therefore, we have two free parameters to describe a whole group of spectral lines: $\alpha=v_0/c$, which is associated with the ''average'' plasma expansion rate $[v_0=by_0)$, and the optical depthof the plasma at the frequency of one of the spectral lines. We have attempted to describe qualitatively the emission spectra of He-like and H-like oxygen ions in the CO_2 plasma. The results are shown in Figs. 4 and 5. It is evident from these figures that the simple model (6) accurately describes, not just qualitatively buteven quantitatively, the profiles of all observed lines of oxygen ions for values of the free parameters $\alpha=10^{-3}$ and $\tau^{bf}(1s\,4p-1s^2)=9$. (We note that the model (6) describes only the line profiles; the relative line intensities used in plotting the theoretical spectrumare taken from experiment.)

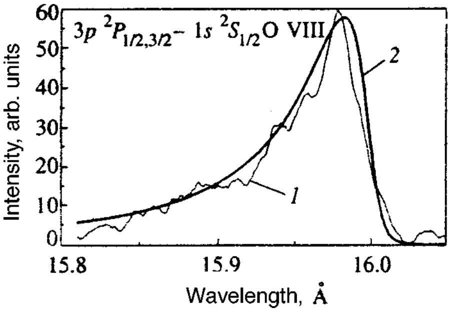

FIG. 5. Comparison of the model spectrum (6) (heavy curve) with the experimental result for the Ly_β line of the H-like O VIII ion.

Since the broadening of the spectral lines in our experiments has been attributed to translational motion of the plasma, the analysis of the line profiles can be used, generally speaking, to measure the distribution of excited ions with respect to their dispersal energies. This procedure is simplest (and most accurate) for the Ar plasma, for which photoionization absorption does not occur, and the spectral line profile is directly proportional to the dispersal velocity distribution of excited ions:

$$F^{\mathrm{obs}}_{ik}(\omega-\omega_{ik})\simeq N_i(v/c). \qquad (10)$$

The distribution $N_i(E)$ obtained from (10) (with the normalization $N_i(0)=1$) is shown in Fig. 6 (curve *3*). This distribution can be plotted only for energies $E\leqslant 100$ keV, because for larger mismatches the line profile cannot be distinguished from the noise.

In the case of CO_2, direct measurements of $N_i(E)$ are rendered impossible by the presence of photoionization absorption. Model functions of the type (5) with values of the parameters α and τ such as to ensure good agreement with experiment are useful only for estimating the distribution $N_i(E)$. Such a model distribution is shown in Fig. 6 (curve *1*). Interestingly enough, in the range $E\leqslant 100$ keV this distri-

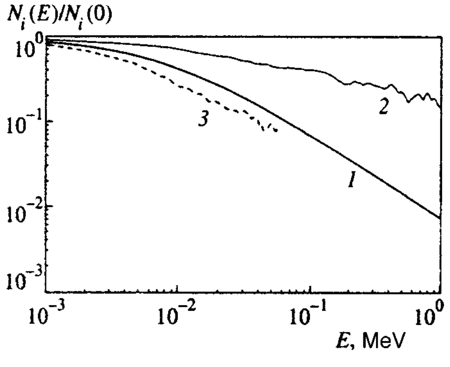

FIG. 6. Distribution of excited ions with respect to their translational kinetic energy (expansion): (*3*) direct measurement for an Ar plasma; (*1*) approximate profile for a CO_2 plasma; (*2*) distribution obtained for a CO_2 plasma without regard for photoionization absorption.

bution scarcely differs from the measurement result for the Ar plasma. Again we emphasize that in the presence of photoionization absorption (i.e., for the CO_2 plasma) Eq. (10) would yield the totally incorrect result represented by curve *2* in Fig. 6.

The main conclusions drawn from the foregoing discussion are as follows.

1) The short-wavelength wings of the spectral lines for O VII and O VIII ions are attributable to translational motion of the ions, and the lack of long-wavelength wings is associated with photoionization absorption.

2) The detection of a radiative intensity well above the noise level inthe vicinity of frequency mismatches $|(\omega-\omega_{ik})/\omega_{ik}|=1.2\times10^{-2}$ proves unequivocally the existence in the plasma of an appreciable number (beyond that characterized by the tail of a Maxwellian distribution) of ions moving at velocities $v\approx3.6\times10^{8}$ cm/s, i.e., with energies of approximately 1 MeV.

Next we take a look at possible mechanisms underlying the production ofsuch fast ions in a femtosecond laser plasma.

4. GENERATION OF FAST IONS IN A FEMTOSECOND LASER PLASMA

Ions can be accelerated by various processes associated with the emergence of strong electric fields in the laser plasma. The possible mechanisms include the Coulomb explosion of molecules and clusters[7] and the ''hydrodynamic'' acceleration of ions in the dispersal of clusters under the influence of electron kinetic energy.[14] The ponderomotive acceleration of ions under the influence of forces created by high-frequency pressure directly as a result of strong laser radiation in the plasma near the critical density N_c has been discussed[15,16] for long laser pulses of nanosecond duration, when $\tau_{\rm las}$ greatly exceeds the acceleration time.

It is fairly simple to estimate the maximum ion energy for the various acceleration mechanisms. For example, in the Coulomb explosion of a molecule (when as a result of rapid ionization, ions with charge Z_i are situated at the characteristic interatomic spacing of the molecule r_a) the Coulomb potential of the ionic core is $\varphi_{\rm mol}\simeq\Sigma Z_i e/r_a$. Consequently, the kinetic energy that can be acquired by anion with charge $Z|e|$ as a result of the expansion of the ionic core is

$$E_{\rm ion}=Z|e|\varphi_{\rm mol}\simeq Z\sum Z_i e^2/r_a. \tag{11}$$

Because the characteristic interatomic spacing of the molecule is of the same order as the radius of the neutral atom and spans at least several Bohr radii, we infer from Eq. (11) that the maximum energy of oxygen ions formed by the explosion of a CO_2 molecule cannot be too high: $E_{\rm ion}<4$ keV.

Generally speaking, the Coulomb energy of a large cluster can be significantly greater. For example, for a density of molecules $N_{CO_2}\simeq8\times10^{21}\,\text{cm}^{-3}$ the maximum Coulomb energy of an ion $Z|e|$ $Q_{\rm cl}$ in a cluster of radius $r_{\rm cl}\simeq40$ is

$$E_{\rm ion}\simeq Z|e|Q_{\rm cl}/r_{\rm cl}; \tag{12}$$

for $Q_{\rm cl}=Q_{\rm cl}^{\rm max}=22e(4\pi/3)r_{\rm cl}^3N_{CO_2}$ (charge of the cluster for fully ionized oxygen and carbon atoms) this expression yields the estimate $E_{\rm ion}\simeq1.4Z$ MeV. In reality, of course, the cluster charge cannot be greater than $Q_{\rm cl}\simeq E_e r_{\rm cl}/|e|$, i.e., the value dictated by the kinetic energy E_e of electrons capable of escaping the cluster. The maximum ion energy associated with Coulomb ''repulsion'' in the cluster is determined by the electron kinetic energy:

$$E_{\rm ion}\simeq ZE_e. \tag{13}$$

We note that this is precisely the kinetic energy acquired by the bulk of the ions with charge Z in ambipolar hydrodynamic plasmaexpansion (at the speed of sound) under the influence of the thermal pressure of electrons (with $T_e=E_e$). According to Eq. (13), the emergenceof a sizable number of ions with energy $\simeq1$ meV at $Z\simeq6-8$ as a result of ''Coulomb explosion'' of a cluster requires a very ''hot'' component of electrons with energies $\simeq140$ keV. The number of such electrons should be of the order of 10% of all electrons resulting from the complete ionization of all atoms in the cluster, the probability of which is exceedingly low.

Somewhat more detailed estimates of the maximum energy and number of ions can be obtained if the kinetics of ion motion for various electron energy distributions is taken into account.[14] In the case of a Maxwellian electron distribution function, the maximum energy of fast ions in the expansion in vacuum of a finite plasmoid of characteristic radius R is

$$E_{\rm ion}\simeq ZT_e\tau_m^2, \tag{14}$$

where $\tau_m=\sqrt{2}\ln(R/r_{De})$ $r_{De}=v_{Te}/\omega_{pe}=\sqrt{Te/4\pi n_e^2 n_e}$ is the Debyeradius of the electrons. The number of ions $N_{\rm ion}\simeq N_0\exp(-\sqrt{2}\,\tau_m)$ is exponentially small ($\tau_m\gg1$) in this case. At an electron temperature $\simeq1$ keV these relations show that on the order of 0.2% of the total number of ions in the cluster (for $R\simeq r_{\rm cl}\simeq40$ nm and $Z=8$) cannot acquire an energy greater than 200 keV.

The maximum ion acceleration energy depends heavily on the form of the electron distribution function. For example, in the case of a distribution fortified with fast electrons, $f_e\simeq\exp(-v/\sqrt{T_{\rm eff}/m})$, for $v\gg\sqrt{T_{\rm eff}/m}$ we have $\tau_m\simeq2^{5/2}/3[\ln(R/r_{De})]^{3/2}$; according to Eq. (14), for $Z=8$ and $T_{\rm eff}=1$ keV we obtain $E_{\rm ion}\simeq1.44$ MeV, and the relative number of ions with such energies is approximately 0.6×10^{-3}.

During the motion of the clusters, therefore, a small number of ions canbe accelerated to megawatt energies only for non-Maxwellian electron distributions fortified with fast particles. The feasibility of such distributions and the determination of their characteristics require further investigation.

The intriguing possibility of the ponderomotive acceleration of ions to high energies is afforded by the irradiation of a plasma with an extremely strong field of ultrashort laser pulses. The acceleration of ions by femtosecond laser pulses can be highly nonstationary in this case, so that the maximum energy of the accelerated ions depends not only on the radiative intensity, but also on the duration of the laser pulse.

We now consider the model of a plasma generated by the irradiation of a cluster gas, when the density of the plasma (formed from uncondensed gas) between clusters is subcritical, and the cluster plasma density greatly exceeds the critical value. During the laser pulse, the ponderomotive force in standing waves generated by reflection fromclusters will affect ions of the intercluster plasma by way of the charge separation field. During this time an ion with charge Z_i and mass M_i acquires a velocity

$$v_i \simeq \frac{Z_i}{M_i} F_p \tau_{\rm las}, \tag{15}$$

where $F_p = (m_e c^2/4l_E) a_0^2 (1 + a_0^2/2)^{-1/2}$ is the ponderomotiveforce on an electron in a standing wave of the laser field with characteristic scale $l_E \sim c/2\omega_p$, whose intensity I is related to the dimensionless amplitude a_0 by

$$a_0 = eE_{\rm las}/\omega_0 m_e c = 8.6\times 10^{-10} \lambda_L\,[\mu\mathrm{m}]\, I^{1/2}\,[\mathrm{W\cdot cm^{-2}}]$$

(L is the wavelength and ω_0 is the laser frequency). We note that in a standing wave the strength of the field $E_{\rm las}$ can exceed the amplitude of the laser field incident on the plasma. Equation (15) holds for laser pulse durations shorter than the acceleration time, $\tau_{\rm las} < l_E/v_i$, but longer than the period of the electron plasma waves. According to Ref. 15, the maximum energy of the accelerated ions is

$$\frac{M_i v_i^2}{2} = \frac{Z_i^2 m_e^2 c^2}{4M_i} \frac{a_0^4}{1 + a_0^2/2} (\omega_0 \tau_{\rm las})^2, \tag{16}$$

where the transient acceleration condition has the form

$$\tau_{\rm las} < \frac{2}{\omega_0} \left(\frac{M_i}{Z_i m_e} \right)^{1/2} \frac{(1 + a_0^2)^{1/4}}{a_0} \equiv \tau^*. \tag{17}$$

For laser pulses of longer duration, such that the opposite of condition (17) holds ($\tau_{\rm las} > \tau^*$), the maximum ion energy is determined by the total ponderomotive potential difference:

$$\frac{M_i v_i^2}{2} = \frac{Z_i m_e c^2}{4} \frac{a_0^2}{(1 + a_0^2/2)^{1/2}}. \tag{18}$$

If the duration of the laser pulse corresponds to the transition from transient to steady-state acceleration ($\tau_{\rm las} \simeq \tau^*$), the two equations (16) and (18) yield essentially the same energy. Under the conditions of our experiment ($Z_i = 8$, $\tau_{\rm las} \simeq 60$) for $a_0 \simeq 1$ and $\tau_{\rm las} \simeq \tau^*$, according to Eqs. (16) and (18), the maximum energy of the accelerated ions is $\simeq 1$ MeV, which is consistent with the x-ray spectral measurements. We also note that for an uncondensed (intercluster) gas of density $n_a \simeq 3\times 10^{19}\,\mathrm{cm}^{-3}$ ($P = 1$ atm), the number ofions in the intercluster plasma is only slightly less than the number of ions in the clusters ($n_a (l/r_{\rm cl})^3 / n_{\rm cl} \simeq 0.4$). The number of electrons is now sufficient to maintain a charge separation field such that a ponderomotive force is applied to the ions. The number of accelerated ions, which is determined by their density inthe region where a standing wave exists, will also be sufficiently large for the given conditions, with the distance between clusters being of the order of half the laser wavelength.

5. CONCLUSION

The reported experimental investigations show that x-ray spectral methods can be used for the systematic acquisition of experimental data about the properties of the plasma formed when clusters are heated by high-intensity femtosecond laser pulses. We have obtained spatially resolved x-ray spectra of CO_2 and Ar cluster plasmas for the firsttime, and we have observed that photoionization absorption influences the profilesof the spectral lines in a plasma of complex chemical composition. The recorded features of the x-ray emission spectra indicate unequivocally the presencein the cluster plasma of a large relative number of excited ions ($\simeq 10^{-2} - 10^{-3}$) with energies of 0.1–1 MeV. A comparison of mechanisms capable of accelerating ions to high energies shows that the experimental results can be attributed to ponderomotive forces generatedin standing waves resulting from the reflection of laser radiation from the clusters.

The authors are grateful to Prof. A. N. Starostin for stimulating discussions. This work received partial financial support from the Russian Fund for Fundamental Research (Grants 98-02-16263 and 96-02-16111) and INTAS (Grant No. 2090-98).

*)E-mail: skobelev@orc.ru

---

[1] B. D. Thompson, A. McPherson, K. Boyer, and C. K. Rhodes, J. Phys. B At. Mol. Opt. Phys. **27**, 4391 (1994).

[2] A. B. Borisov, A. McPherson, B. D. Thompson, K. Boyer, and C. K. Rhodes, J. Phys. B At. Mol. Opt. Phys. **28**, 2143 (1995).

[3] A. B. Borisov, J. W. Longworth, A. McPherson, K. Boyer, and C. K. Rhodes, J. Phys. B At. Mol. Opt. Phys. **29**, 247 (1996).

[4] A. B. Borisov, A. McPherson, K. Boyer, and C. K. Rhodes, J. Phys. B At. Mol. Opt. Phys. **29**, L113 (1996).

[5] T. Donnelly, T. Ditmire, N. Neuman, M. P. Perry, and R. W. Falcone, Phys. Rev. Lett. **76**, 14 (1996).

[6] T. Ditmire, J. W. G. Tisch, E. Springate, M. B. Mason, N. Hay, J. P. Marangos, and M. H. R. Hutchinson, Phys. Rev. Lett. **78**, 2732 (1997).

[7] M. Lezius, S. Dobosz, D. Normand, and M. Schmidt, Phys. Rev. Lett. **80**, 261 (1998).

[8] C. Rose-Petruck, K. Schaefer, and C. P. J. Barry, in *Applications of Laser Plasma Radiation II*, M. C. Richardson and G. A. Kryla (Eds.); Proc. SPIE **2523**, 272 (1995).

[9] E. M. Snyder, S. A. Buzza, and A. W. Castleman Jr., Phys. Rev. Lett. **77**, 3347 (1996).

[10] I. Yu. Skobelev, A. Ya. Faenov, B. A. Bryunetkin, V. M. Dyakin, T. A. Pikuz, S. A. Pikuz, T. A. Shelkovenko, and V. M. Romanova, Zh. Éksp. Teor. Fiz. **108**, 1263 (1995) [JETP **81**, 692 (1995)].

[11] V. A. Boĭko, A. V. Vinogradov, S. A. Pikuz, I. Yu. Skobelev, and A. Ya. Faenov, in *Progress in Science and Technology, Radio Engineering Series* [in Russian], N. G. Basov (Ed.), VINITI, Moscow (1980), p. 27.

[12] B. A. Boĭko, S. A. Vinogradov, S. A. Pikuz, I. Yu. Nosach, G. V. Peregudov, A. L. Petrov, S. A. Pikuz, I. Yu. Skobelev, A. Ya. Faenov, V. A. Chirkov, and K. A. Shilov, Kvantovaya Élektron. **8**, 28 (1981) [Sov. J. Quantum Electron. **11**, 13 (1981)].

[13] I. I. Sobelman, *Atomic Spectra and Radiative Transitions*, Springer-Verlag, Berlin–New York (1979).

[14] A. V. Gurevich and L. P. Pitaevskiĭ, in *Problems in Plasma Theory* [in Russian], No. 10, A. B. Mikhaĭlovskiĭ (Ed.), Atomizdat, Moscow (1980), p. 3.

[15] V. P. Silin, JETP Lett. **21**, 152 (1975).

[16] N. E. Andreev, Yu. A. Zakharenkov, N. N. Zorev, V. T. Tikhonchuk, and A. S. Shikanov, Zh. Éksp. Teor. Fiz. **76**, 976(1979) [Sov. Phys. JETP **49**, 492 (1979)].

Translated by James S. Wood

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Role of stochastic fluctuations in the charge on macroscopic particles in dusty plasmas

O. S. Vaulina, A. P. Nefedov, O. F. Petrov,*) and S. A. Khrapak

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

(Submitted 11 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 2067–2079 (June 1999)

The currents which charge a macroscopic particle placed in a plasma consist of discrete charges; hence, the charge can undergo random fluctuations about its equilibrium value. These random fluctuations can be described by a simple model which, if the mechanisms for charging of macroscopic particles are known, makes it possible to determine the dependence of the temporal and amplitude characteristics of the fluctuations on the plasma parameters. This model can be used to study the effect of charge fluctuations on the dynamics of the macroscopic particles. The case of so-called plasma-dust crystals (i.e., highly ordered structures which develop because of strong interactions among macroscopic particles) in laboratory gaseous discharge plasmas is considered as an example. The molecular dynamics method shows that, under certain conditions, random fluctuations in the charge can effectively heat a system of macroscopic particles, thereby impeding the ordering process. [S1063-7761(99)01206-8]

1. INTRODUCTION

A macroscopic particle placed in a gaseous discharge plasma acquires an electrical charge by absorbing electrons and ions from the plasma. The equilibrium potential (charge) is determined by the condition that the current to the particle equals zero. It is negative because the electrons are more mobile than the ions and can be very high (on the order of 10^3-10^5 elementary charges for micron-sized particles). Such a large number of charges means that systems of macroscopic particles are often highly nonideal, i.e., the potential energy of the interaction between the macroscopic particles greatly exceeds their average kinetic energy. In this case, the highly ordered structures of macroscopic particles known as plasma-dust crystals can develop.[1] The first reports of the creation of crystals from dust particles in low pressure rf discharge plasmas in inert gases appeared in 1994.[2–5] Ordered structures were observed somewhat later in other types of plasmas, as well: in a thermal atmospheric pressure plasma[6] and in stratified glow discharges.[7] (A detailed review of these experiments can also be found in Ref. 8.) Systems consisting of a plasma and charged macroscopic particles are also widespread in space and in devices for plasma materials processing. This explains the great interest in systems of this type at this time.

Dusty plasmas are often described theoretically in the approximation of a constant charge on the macroscopic particles, since the times of interest are typically considerably longer than the time to establish an equilibrium charge. In this case, a dusty plasma can be regarded as a multispecies plasma in which the macroscopic particles represent ions with anomalously high charges and masses. This approach, however, does not take an important feature of these systems into account. In fact, the charge on the macroscopic particles is determined by the local plasma parameters (potential, ion and electron densities and temperatures, currents, etc.). Therefore, it has to be regarded as one of the set of dynamic variables determining the behavior of the plasma. Thus, for example, perturbations in the charge on macroscopic particles during propagation of electrical oscillations in a dusty plasma may give rise to new mechanisms for dissipation and instability.[9]

Even in an isotropic, spatially uniform, unperturbed plasma, however, the charge on the macroscopic particles will undergo random fluctuations about its equilibrium value. This is because ions and electrons are absorbed by the surfaces of macroscopic particles at random times and in random sequences. In this paper we examine precisely this type of fluctuation in the charges on macroscopic particles in a plasma and propose a model that can be used to obtain their quantitative characteristics. The model is also applicable to studies of the effect of random charge fluctuations on the dynamic behavior of a system of macroscopic particles. As an example of such a system, we examine the formation of dust particles in the electrode region of a gaseous discharge, where the force of gravity acting on the macroscopic particles is balanced by an electric field.

Before proceeding to a description of the stochastic fluctuation model,let us summarize the earlier work. Morphill *et al.*,[10] first pointedout the need to include random charge fluctuations, assuming that the amplitude of the fluctuations is given by $\Delta Z \equiv \sqrt{\langle \delta Z^2 \rangle} = \sqrt{\langle |Z| \rangle}$, where $\langle Z \rangle$ is the equilibrium charge on a macroscopic particle in units of the electronic charge. Cui and Goree[11] used a numerical simulation to obtain the temporal sequence of the values of the charge on macroscopic particles in the absence of emission processes. (The orbital approximation and Maxwellian distributions of the electrons and ions were used to determine the current to a macroscopic particle.) Their main result[11] isthat $\Delta Z = 0.5\sqrt{\langle |Z| \rangle}$ for a wide range of parameters of the plasma and macroscopic particle, provided that $\langle|Z|\rangle \gg 1$. Finally, Matsukas and Russell[12] presented an analytical model developed for the same conditions which yielded a charge distribution of the macroscopic particles (and, therefore, the characteristic amplitude of the fluctuations). Our paper differs from these[10–12] in that we are primarily interested in the dynamical, rather than the static, properties of the random fluctuations in charge, since it is they which determine the effect of the fluctuations on the dynamics of a system of macroscopic particles.

2. MODEL OF RANDOM FLUCTUATIONS IN THE CHARGE ON MACROSCOPIC PARTICLES

Let us consider an isolated macroscopic particle which carries charge Z. For simplicity we assume that all the negative charges in the system are carried by electrons and the positive ions have unit charge. The equation describing the random fluctuations in the charge can be written in a form analogous to the Langevin equation of motion for a free Brownian particle in one dimension, with the sole difference that the motion takes place ina one-dimensional charge space, rather than in velocity space. For small deviations $\delta Z(t)$ of the charge from its equilibrium value $(|\langle Z\rangle| \gg |\delta Z(t)|)$, we can write

$$\frac{d\,\delta Z}{dt}+\beta\,\delta Z=F(t), \tag{1}$$

where $\beta=-\partial I/\partial Z|_{z=\langle Z\rangle}$, I is the current to the macroscopic particle (in s^{-1}), and

$$F(t)=\sum_j \delta(t-t_j)(\pm 1)_j$$

is a random term which represents the source of the fluctuations and describes the absorption of a single plasma electron (minus sign) or ion (plus sign) at a random time t_j. Under the above assumptions, the current to the macroscopic particle can be written in the form $I=I^+-I^-$, where I^+ is the current determined by the absorption of plasma ions and I^- is the current determined by the absorption of electrons. The term $\beta\,\delta Z$ $(\beta>0)$ acts as a restoring force which tends to return the charge to its equilibrium value $\langle Z\rangle$ determined by the condition $I=0$. It is easy to verify that the term $F(t)$ obeys the relations

$$\langle F(t)\rangle=0, \quad \langle F(t)F(t')\rangle=\frac{1}{t_0}\delta(t-t'), \tag{2}$$

where $1/t_0$ characterizes the rate at which ions and electrons are absorbed, so that $t_0^{-1}=I^+ + I^- = 2I^-$. Integrating Eq. (1) using Eq. (2) allows usto obtain the temporal autocorrelation function of the charge fluctuations in the form

$$\langle \delta Z(t)\,\delta Z(t')\rangle=\frac{1}{2t_0\beta}\exp(-\beta|t-t'|). \tag{3}$$

Equation (3) is an important result and we shall use it for studying the effect of random charge fluctuations on the dynamic behavior of the macroscopic particles. It contains information on the amplitude $\Delta Z=(2t_0\beta)^{-1/2}$, as well as on the correlation time of the fluctuations, $\tau_c=1/\beta$. We now determine the values of ΔZ and τ_c for the practically important case of the charging of macroscopic particles in a laboratory gaseous discharge plasma typical of the experimental conditions in Refs. 2–5 and 7. (Other charging mechanisms, including the thermionic emission of electrons characteristic of thermal plasmas[6] and photoemission charging of macroscopic particles, which plays a major role in space,[13] are examined elsewhere.[14]) To do this, we use the standard equations for the electron and ion fluxes onto the surface of a negatively charged macroscopic particle:[15]

TABLE I. Values of γ, α, and β (see text) for different plasma parameters m_i and T_i/T_e.

m_i, amu	T_i/T_e	γ	α	β, s^{-1}
1	0.05	1.698	0.61	$1.1\cdot 10^5$
1	1	2.501	0.56	$4.0\cdot 10^4$
40	0.05	2.989	0.50	$2.5\cdot 10^4$
40	1	3.952	0.46	$8.7\cdot 10^3$

$$I^-=I_e=\sqrt{8\pi}\,a^2 n_e v_{Te}\exp\left(\frac{Ze^2}{aT_e}\right), \tag{4}$$

$$I^+=I_i=\sqrt{8\pi}\,a^2 n_i v_{Ti}\left(1-\frac{Ze^2}{aT_i}\right), \tag{5}$$

where a is the particle radius, $n_{e(i)}$, $m_{e(i)}$, and $T_{e(i)}$ are the density, mass, and temperature of the electrons (ions), respectively, and e is the electronic charge. We have also used the notation $v_{Te(i)}=(T_{e(i)}/m_{e(i)})^{1/2}$ for the thermal velocity of the electrons (ions). If the density of macroscopic particles is nottoo high, so that $n_e \approx n_i$, then the equilibrium charge can be written in the form

$$\langle Z\rangle=-\gamma\left(\frac{T_i}{T_e},\frac{m_i}{m_e}\right)\frac{aT_e}{e^2}.$$

Here $\gamma>0$ is a coefficient on the order of unity and defined by the equation $\exp(-\gamma)=(1+\gamma T_e/T_i)v_{Ti}/v_{Te}$. Using Eqs. (4) and (5), we can readily obtain

$$\beta=-\left.\frac{\partial I}{\partial Z}\right|_{z=\langle z\rangle}=\sqrt{8\pi}\,a^2 n_e v_{Te}\frac{e^2}{aT_e}\left[\sqrt{\frac{m_e}{m_i}\frac{T_e}{T_i}}+\exp(-\gamma)\right], \tag{6}$$

$$\Delta Z=\sqrt{\frac{1+\gamma\theta}{\gamma(1+\theta+\gamma\theta)}}\,\sqrt{\langle|Z|\rangle}=\alpha\sqrt{\langle|Z|\rangle}, \tag{7}$$

where $\theta=T_e/T_i$. The same expression for ΔZ has been obtained in a slightly different way in Ref. 12. A numerical calculation[11] has shown that $\alpha\approx 0.5$ for a wide range of parameters of the plasma and of the macroscopic particles α [determined using Eq. (7)] for the plasma parameters used in the calculations of Ref. 11 and values of β [Eq. (6)] for a macroscopic particle of 1 μm diameter under the following conditions: $n_e=5\cdot 10^8$ cm^{-3} and $T_e=4$ eV. It is evident that α depends weakly on the plasma parameters and that it is

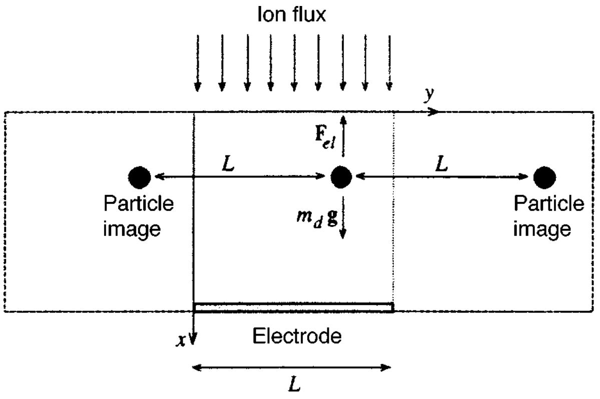

FIG. 1. An isolated macroscopic particle contained near an electrode as a result of the balance between the force of gravity $m_d g$ and the electrical force $F_{el}=ZeE$. The ions have the Bohm velocity $v_0=\sqrt{T_e/m_i}$. The two images of the macroscopic particle are a consequence of using periodic boundary conditions in the y direction. The electric field increases from the boundary of the electrode sheath, where $E(x=0)\approx 0$, to $E=E_s$ at the electrode surface.

close to 0.5; at the same time, the frequency of the fluctuations, β, varies by more than an order of magnitude as the plasma parameters are changed.

When there is a directed flux of ions, whose velocity v_0 greatly exceeds the ion thermal velocity v_{Ti}, as happens, for example, in the electrode sheaths of discharges, the following expression should be used instead of Eq. (5):

$$I^+=I_i=\pi a^2 n_i v_0\left(1-\frac{2Ze^2}{m_i v_0^2}\right).$$

Then, instead of Eq. (6) we obtain

$$\beta=-\left.\frac{\partial I}{\partial Z}\right|_{z=\langle z\rangle}=\sqrt{8\pi}\,a^2 n_e v_{Te}\frac{e^2}{aT_e}\left[\frac{v_0}{v_{Te}}\frac{2T_e}{m_i v_0^2}+\exp(-\gamma)\right]. \tag{8}$$

Equation (7) is still valid if we redefine θ as the ratio of the electron temperature to the kinetic energy of the directed ion motion, i.e., $\theta=2T_e/m_i v_0^2$.

3. ESTIMATE OF THE EFFECT OF RANDOM CHARGE FLUCTUATIONS ON THE DYNAMICS OF MACROSCOPIC PARTICLES

We now proceed to an examination of the effect of random charge fluctuations on the dynamic processes in a system of macroscopic particles. As a first, simple example, let us consider a macroscopic particle contained in the electrode sheath of a discharge, where the force of gravity $F_g=m_d g$ is balanced by the electrostatic force $F_{el}=ZeE$. Here m_d is the mass of the macroscopic particle and E is the electric field directed toward and increasing toward the macroscopic particle (Fig. 1). This situation corresponds to the experimental conditions in Refs. 2–5, where negatively charged macroscopic particles located near the lower electrode (cathode) form ordered structures. Because of random fluctuations in its charge, the macroscopic particle will undergo random oscillations in the direction of the field. It is easy to confirm that during these oscillations, the electric field will perform positive work on the macroscopic particle, increasing its kinetic energy. Without dissipation, the energy will increase without bound. Friction on the neutral gas in a weakly ionized plasma means that the energy of the random oscillations of the macroscopic particle will reach some steady-state value. Assuming that the electric field is linear near the equilibrium position x_0, the force acting on the macroscopic particle, including the charge fluctuations, is $F_{el}+m_d g=-m_d\omega^2\delta x+eE_0\delta Z$, where δx is a small displacement, $E_0=E(x_0)$ is the electric field at the equilibrium position $(m_d g+\langle Z\rangle eE_0=0)$, and

$$\omega^2=-\frac{\langle Z\rangle e}{m_d}\left.\frac{\partial E}{\partial x}\right|_{x_0}$$

is a constant characterizing the rigidity with which the macroscopic particle is confined (stiffness coefficient) in the potential well formed by the action of the electrostatic and gravitational forces (recall that $\langle Z\rangle<0$). The results of the previous section can be used to determine the average kinetic energy of an isolated macroscopic particle in the direction of the field:

$$K_x=\frac{T_n}{2}+\frac{m_d g^2\alpha^2}{2\langle|Z|\rangle}\frac{\beta}{\nu[\omega^2+\beta(\beta+\nu)]}\approx\frac{T_n}{2}+\frac{m_d g^2}{2\langle|Z|\rangle\xi}, \tag{9}$$

where the kinetic energy is introduced through the formula $K_x=m_d\langle\delta\dot{x}^2\rangle/2$, ν is the coefficientof friction on the neutral gas (in s^{-1}), T_n is the neutral gas temperature, $\xi=\nu\beta/\alpha^2$, and the last estimate was made assuming that $\omega\ll\beta$ and $\nu\ll\beta$. (As a rule, this assumption is fully justified.) The first term on the right of Eq. (9) corresponds tothe thermal energy per degree of freedom of the macroscopic particle (in equilibrium with the surrounding gas). The additional kinetic energy owing to the fluctuations is given by the second term on the right of Eq. (9). For simplicity, in the following we omit the Brownian motion and study only the added term owing to the random charge fluctuations. Since the motion of the macroscopic particles caused by the random charge fluctuations is chaotic, we can also consider an effective temperature of the macroscopic particle, $K_x=T_{dx,\mathrm{eff}}/2$. Note that in this case of an isolated particle,the fluctuations can ''heat'' a macroscopic particle only in the direction ofthe field and do not change the kinetic energy of the macroscopic particle in directions perpendicular to the field, so that $K_y=T_n/2$. Here the inequality $T_{dx,\mathrm{eff}}>T_{dy}=T_n$ holds.

We have derived Eq. (9) assuming that the macroscopic particle is isolated, which corresponds to a weak interactions among the macroscopic particles. As a rule, however, a system of macroscopic particles under these conditions is highly nonideal. The effect of the interaction on our results will be examined below in a numerical simulation of a system. Here we point out another effect. It has been noted[11] that in a system of interacting charged particles, fluctuations in the charges give rise to fluctuations in the interparticle interaction potential, which should cause chaotic motion of the macroscopic particles in addition to the thermal motion. We shall estimate the magnitude of this effect in terms of an approach[17] used to study the effect of the spatial dependence of the charge on the macroscopic particle dynamics. Specifically, we separate the force, owing to the other charges, which acts on a test charge into constant and fluctuating components. The fluctuating component is determined in our case by stochastic fluctuations in the charges on the macroscopic particles and not by the change in the charges as the macroscopic particles move. Following Ref. 17, we obtain the following estimate for the additional kinetic energy beyond the thermal energy (neglecting screening):

$$K_{ad} \sim \frac{\langle Z\rangle^2 \langle \delta Z^2\rangle e^4}{m_d \nu l^4}\tau_c \sim \frac{\langle |Z|\rangle^3 e^4}{m_d l^4}\frac{1}{\xi},$$

where l is the average distance between the macroscopic particles and τ_c is the correlation time for the random force, which is given by the correlation time of the fluctuations in our case. Thus, because of the charge fluctuations, the effective temperature of a system of macroscopic particles is always somewhat higher than the thermal level. The reason for this inequality is that the system of dust particles is open, so thatit can exchange charge and energy with the gaseous discharge plasma.

In the case of a system of interacting macroscopic particles located in the electrode sheath of a gaseous discharge, both of the effects described above occur. The ratio of the contributions from the two is

$$\frac{K_x}{K_{ad}} \sim \left(\frac{m_d g}{\langle Z^2\rangle e^2/l^2}\right)^2. \tag{10}$$

Therefore, their relative roles are determined by the square of the ratio of the force of gravity to the Coulomb interparticle interaction force. For the typical parameters of an experiment with a dusty plasma, $a=1\ \mu$m, density $\rho=5\ \mathrm{g/cm^3}$ of the material in the particles ($m_d \approx 2\cdot 10^{-11}$ g), $\langle Z\rangle = a\phi_s \approx 3\cdot 10^3$ ($\phi_s \approx 5$ V, the floating potential), and $l \approx 500\ \mu$m, we obtain $K_x/K_{ad} \approx 600$. For larger macroscopic particles this ratio will be still higher, since $K_x/K_{ad} \sim a^2 l^4$. Therefore, under typical experimental conditions,[2–5] the first effect, of magnitude given by Eq. (9), will make the larger contribution to the disequilibrium.

4. NUMERICAL SIMULATION

In order to account for the interaction among the macroscopic particles, we have done a molecular dynamics simulation of a system of macroscopic particles with fluctuating charges located in the electrode sheath of a discharge. The computational region consisted of a square with sides of length L equal to 50 times the average interparticle separation. Periodic boundary conditions were imposed in the y direction perpendicular to the direction of the external force F_{ext}. (See Fig. 1.) The number of particles in the simulation was varied between 50 and 300. The two-dimensional equation of motion was solved for each macroscopic particle, including the pairwise interaction F_{int} between the particles, the frictional force, the totalexternal force F_{ext} in the gravitational and electric fields, and therandom fluctuations in the charges of the macroscopic particles:

$$m_d \frac{d^2\mathbf{r}_k}{dt^2} = \sum_j F_{\mathrm{int}}(r)\big|_{r=|\mathbf{r}_k-\mathbf{r}_j|}\frac{\mathbf{r}_k-\mathbf{r}_j}{|\mathbf{r}_k-\mathbf{r}_j|} - m_d \nu \frac{d\mathbf{r}_k}{dt} + \mathbf{F}_{\mathrm{ext}}.$$

Here the pairwise interaction between the macroscopic particles is given by

$$F_{\mathrm{int}}(r) = -eZ(t)\frac{\partial \phi_D}{\partial r},$$

where

$$\phi_D = \frac{eZ(t)}{r}\exp\left(-\frac{r}{\lambda}\right)$$

is the screened Coulomb potential with a screening length λ determined by the appropriate Debye radius. (For a homogeneous, isotropic plasma, $\lambda^{-2} = \lambda_{De}^{-2} + \lambda_{Di}^{-2} \approx \lambda_{Di}^{-2}$.) Note that the interaction force depends on time, as noted at the end of the previous section. Here the time dependence shows up only through the time variation in the charge. This can be assumed, if the characteristic time of the charge fluctuations, τ_c, is much longer than the reciprocal of the plasma frequency (ω_{pi} for a homogeneous, isotropic plasma), so that the plasma electrons and ions are able to fine-tune themselves in the charge fluctuations.

In the numerical simulations the electric field in the electrode sheath was specified as a linear function of the coordinate x, so that $E(x) = E_s x/L$. Here the external force depends on both position andtime, and acts only in the x direction: $F_{\mathrm{ext}}(t,x) = m_d g + eZ(t)E(x)$.

The random fluctuations in the macroscopic particle charges are assumed to be uncorrelated. Here the random fluctuations were modelled by a gaussian random variable, so that at each integration step the increment in the charge was given by

$$Z_{i+1} = \langle Z\rangle + [(Z_i - \langle Z\rangle) + \delta Z_t \zeta](1-\beta\Delta t),$$

where $Z_i = Z(t_i)$, $t_{i+1} = t_i + \Delta t$, $\delta Z_t = \Delta Z\sqrt{2\beta\Delta t}$, $\zeta = \sin(2\pi\chi_1)\sqrt{2\ln(1/\chi_2)}$, and χ_1 and χ_2 are random numbers distributed uniformly on the segment $[0,1]$. The integration step Δt was $\Delta t = \tau_c/20$ in the simulations. (Note that $\tau_c \ll \nu^{-1}$ in our calculations.) For simplicity, in the calculations we have neglected the Brownian force owing to the surrounding gas, so that the motion of the macroscopic particles was determined solely by the charge fluctuations.

5. RESULTS OF THE NUMERICAL SIMULATIONS

We have found that the macroscopic particles form layers (up to six layers in our calculations) in the potential well created for them by theexternal force (the sum of the electrostatic and gravitational forces) in the x direction. (See Fig. 2a.) The layers are perpendicular to the direction of the external force and the number of layers is determined by the system parameters: the number of particles, screening length λ, and potential well characteristic ω. For constant ω, an increase in the number of particles or in the screening length increases the number of layers, since their mutual repulsion increases. Layer formation also occurred in the experiments of Refs. 2–5, where ~ 10 layers were observed. We have also found that layer formation is discrete in nature; in par-

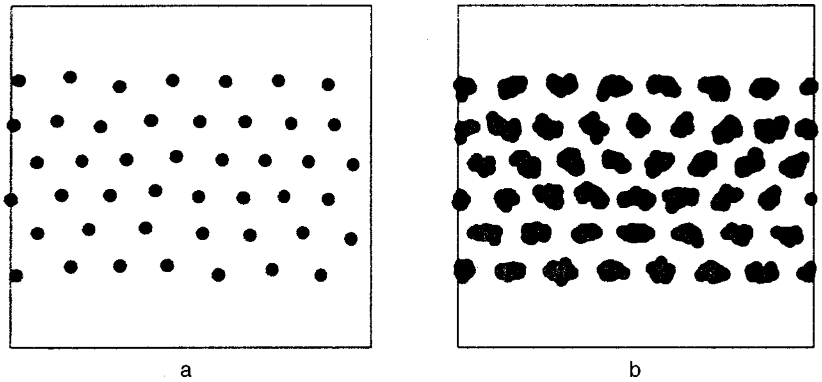

FIG. 2. A portion of the computational region. Macroscopic particles with radii $a=25$ μm form six layers ($\xi\approx 5\cdot 10^4$ s^{-2}): (a) instantaneous positions of the macroscopic particles, (b) a sequence of positions of the macroscopic particles, separated by a time interval ν^{-1}, with an exposure time of 20 ν^{-1}.

ticular, for certain critical values of ω and λ, adding a new particle increases the number of layers by unity. On the whole, the process of layer formation in our calculations is similar to layer formation in an earlier three-dimensional simulation of dusty plasmas in a one-dimensional external force field.[18]

The main results of the simulation were obtained for two types of macroscopic particles with different radii, $a=5$ μm and 25 μm. The density of the macroscopic particle material was taken to be $\rho=5$ g/cm^3. The friction rate was $\nu=1$ s^{-1} (for $a=25$ μm) and $\nu=5$ s^{-1} (for $a=5$ μm). These values correspond to a pressure of about 0.5 Torr in helium at room temperature. The equilibrium surface potential of the macroscopic particles, ϕ_s, was assumed equal to 5 V, which corresponds to a charge $\langle Z\rangle\approx 1.7\cdot 10^4$ on the macroscopic particles with $a=5$ μm and $\langle Z\rangle\approx 8.7\cdot 10^4$ for $a=25$ μm. The screening length was assumed to be $\lambda=450$ μm. For each type of macroscopic particle, we determined the dependence of the dynamic characteristics on the parameter ξ (actually on the parameter β/α^2, since ν must be fixed, in order to solve the equations of motion) in the approximate expression (9).

We found that the velocity distributions of the particles are anisotropic and Maxwellian, characterized by two temperatures (which corresponding to two different directions): $T_{d,x}=2K_x^s$ and $T_{d,y}=2K_y^s$, where the $K_{x(y)}^s=m_d\langle v_{x(y)}^2\rangle/2$ are determined using the data from the simulations (denoted by the index s). Note that K_y^s is everywhere less than K_x^s. This is because energy enters the system only in the directionof the external force (x). However, because of the interaction between the macroscopic particles, their kinetic energy in the direction perpendicular to the external force is nonzero. (Recall the estimate (10) and ourneglect of the Brownian motion in our calculations.) Figure 2a shows the instantaneous position of 25-μm-diam macroscopic particles which form six layers. The average distance between the macroscopic particles (determined from the position of the first maximum of the pairwise correlation function) is 330 μm. The average total kinetic energy of the macroscopic particles is $K^s\equiv K_x^s+K_y^s=11$ eV ($K_x^s=7$ eV and $K_y^s=4$ eV). Figure 2b shows the average deviation of the particles of a time $t=20/\nu$. Note that the root mean square displacement of the particles along the y axis exceeds that along the x axis, although the difference in the kinetic energies obeys $K_x^s>K_y^s$. This is because there is a potential barrier for the particles in the x direction.

The dependence of the total kinetic energy K^s on the parameter ξ for macroscopic particles of both sizes and different numbers of layers is shown in Fig. 3. For comparison, the analytic result for an isolated particle is also shown in the graph. [The lines correspond to Eq. (9)]. It can be seen that the total energy is close to the analytic value for K_x. The deviation $\delta=(K_x-K^s)/K^s$ is greatest for six layers (a strongly interacting system) and is roughly 25%. It is possible to find the coupling parameter Γ, which is defined as the ratio of the Coulomb interaction energy between neighboring particles to their average kinetic energy: $\Gamma=Z^2e^2/lK^s$. For six layers ($l\approx 330$ μm), Γ varies from 13 ($\xi\approx 5\cdot 10^2$ s^{-2}) to $1.3\cdot 10^4$ ($\xi\approx 5\cdot 10^5$ s^{-2}) for the particles with a radius $a=5$ μm and from 140 ($\xi\approx 5\cdot 10^3$ s^{-2}) to $1.4\cdot 10^5$ ($\xi\approx 5\cdot 10^6$ s^{-2}) for the particles with $a=25$ μm. This suggests that Eq. (9) can also be used as an estimate for systems of strongly interacting macroscopic particles (if we treat K_x as the total kinetic energy of the macroscopic particles), although we derived it assuming an isolated macroscopic particle. Furthermore, these results provide a qualitative explanation for some of the experimentally observed phenomena. In particular, in many experiments melting of the plasma crystal has been observed as the neutral pressure is reduced.[1,19,20] From the standpoint of charge fluctuations, thisis easily explained. In fact, lowering the pressure causes a reduction inthe coefficient of friction, ν, and, therefore, in ξ. The kinetic energy of the macroscopic particles increases (Fig. 3), while the coupling parameter Γ decreases, which

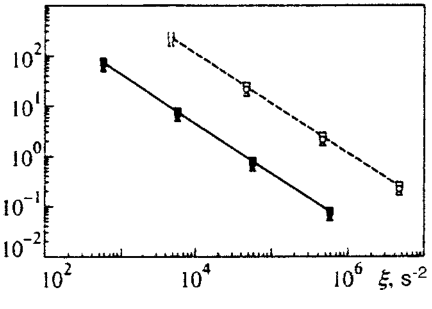

FIG. 3. The total kinetic energy of macroscopic particles $K^s=K_x^s+K_y^s$ as a function of ξ for macroscopic particle radii $a=25$ μm (open points) and $a=5$ μm (filled points): (squares) a single layer; (circles) three layers; (triangles) six layers. The smooth and dashed curves are calculations according to Eq. (9) for $a=5$ and 25 μm, respectively.

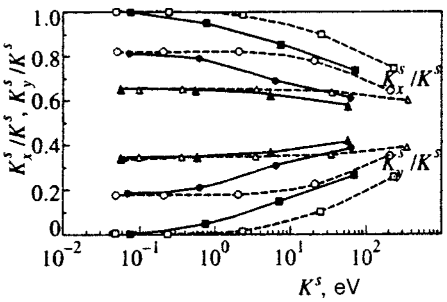

FIG. 4. The ratios K_x^s/K^s and K_y^s/K^s as functions of K^s, which illustrate the distribution of the kinetic energy over the degreesof freedom. The notation is as in Fig. 3.

may be the reason for the melting. Here the quantitative estimates are complicated because, as the pressure is varied, other parameters of the plasma (electron and ion densities and temperatures and, therefore, the charge on the macroscopic particles, screening length, etc.) may also change.

Figure 4 shows plots of the ratios K_x^s/K^s and K_y^s/K^s, which characterize the fraction of the energy K_x that is redistributed into the y direction, as a function of K^s. In our case, this redistribution is related exclusively to the interaction (collisions) between macroscopic particles. It is clear that, the stronger the interaction between the macroscopic particles is, the more strikingly the redistribution of the kinetic energy over the directions should showup. Thus, for example, in the case of six layers, the total energy K^s is essentially uniformly distributed between the x and y directions for large K^s. We believe that this result is independent of the nature ofthe random motion, i.e., the interaction redistributes a random motion which develops in one direction over the other directions. The simulations also showed that the energy redistribution coefficient does not depend on the ratio β/α^2, but is determined by the coefficient of friction, ν. As an illustration of this fact, Fig. 5 shows plots ofthe ratios K_x^s/K^s and K_y^s/K^s as functions of K^s for a single layer of particles with radii $a=5$ μm and different values of ν.

We now estimate this effect numerically for the typical conditions in the cathode region of a discharge. Take $T_e \approx 4$ eV, $m_e/m_i \approx 1.4 \cdot 10^{-4}$, and $n_e \approx n_i \approx 10^8$ cm^{-3}. The ions move toward the cathode at the Bohm velocity $v_0 \sim \sqrt{T_e/m_i} \sim v_{Te}\sqrt{m_e/m_i}$. Here the charging equation gives $\gamma = -e\phi_s/T_e \approx 3$ and, using Eq. (7), we have $\alpha \approx 0.51$, regardless of the size of the macroscopic particles and $\xi \approx 1.1 \cdot 10^6$ s^{-2}, also independently of the size of the macroscopic particles, since $\beta \propto a$ and $\nu \propto a^{-1}$. Now, for the kinetic energy of the macroscopic particles, Eq. (9) gives $K_x \approx 0.04$ eV for $a=5$ μm and $K_x \approx 1$ eV for $a=25$ μm. These energies exceed the thermal energy ($T_n \approx 0.025$ eV at room temperature). Therefore, random fluctuations can cause a substantial disequilibrium between the dust and neutral components. At the same time, the energies observed in experiments with plasma-dust formations have been even higher under certain conditions.[19–21] Thus, we cannot unambiguously assert that the proposed mechanism for the disequilibrium is the most important one in a gaseous discharge plasma. In particular, in an anisotropicregion such as the electrode sheath of a plasma, it is necessary to include the position dependence of the charges, which can also cause a disequilibrium,[17] as well as the random temporal fluctuations in the charge on the macroscopic particles.

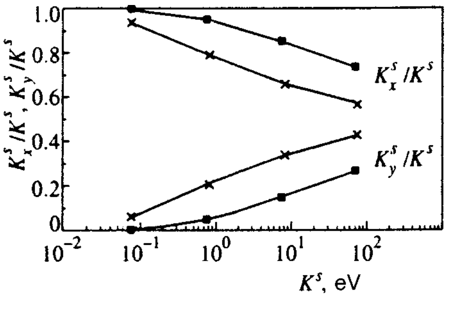

FIG. 5. The ratios K_x^s/K^s and K_y^s/K^s as functions of K^s for viscosities $\nu = 5$ s^{-1} (filled squares) and $\nu = 0.5$ s^{-1} (crosses). The macroscopic particles with $a=5$ μm form a single layer.

We conclude this section with a brief discussion of the simplifying assumptions we have explicitly or implicitly used in studying the effect of random charge fluctuations on the dynamic properties of a system of macroscopic particles. We have assumed that the equilibrium charge of the macroscopic particles is independent of the coordinate x. In general, this is not realistic, since near the electrode the electron density falls off more rapidly than the ion density, while the ions are accelerated by an electric field directed toward the electrode. The approximation we have been examining corresponds to a situation in which the equilibrium position of the macroscopic particles is close to the boundary of the electrode sheath, so that the difference between the ion and Bohm velocities can be neglected, while the electric field depends on the coordinate x much more strongly than does the equilibrium charge. Here we should add the approximation of a linear dependence for the electric field on position used in the numerical simulation. In general, the self-consistent problem must be solved, which allows the distributions $E(x)$ and $\langle Z(x) \rangle$, where the effect of the macroscopic particles themselves on the field distribution in the electrode region has to be taken into account. Then the fluctuations in the charges of different macroscopic particles were assumed to be uncorrelated. This is true, provided

$$\frac{Ze^2 n_p^{1/3}}{\min\{T_e, T_i\}} \ll 1,$$

so that the effective mean free path for interactions of the macroscopic particles with the plasma electrons and ions is much shorter than the average distance between macroscopic particles. Only two forces were considered to act on a macroscopic particle in the electrode sheath, electrostatic and gravitational forces. At the same time, it is known that for a particle with a small radius, the drag force in the ion flow may be greater than the force of gravity.[22] In this case, the equilibrium position will be determined by the balance between the ion drag and electrostatic forces. Finally, the interaction among the macroscopic particles was described by a screened Coulomb potential. At present, other possible mechanisms for the interaction between dust particles can be found in the literature, such as attraction owing to direct bombardment by the plasma electrons and ions,[1] an interaction in an ion flow associated with their focussing in the region between particles,[23,24] and a dipole interaction between macroscopic particles.[24] The effect of charge fluctuations in these cases can be analyzed in a fashion analogous to the above discussion, but lies beyond the scope of this paper.

6. CONCLUSION

The results obtained here mean that stochastic fluctuations in the charge on macroscopic particles in a plasma can have a significant effect on the dynamic properties of a system of macroscopic particles. Determining the magnitude of this effect in each specific case requires knowledge of the characteristic amplitudes and correlation times of the fluctuations. The model proposed here can be used to determine both characteristics of the fluctuations analytically, if expressions are known for the currents which charge a macroscopic particle.

As an example, we have considered the conditions typical of experiments in gaseous discharge plasmas, where the macroscopic particles are charged by absorbing electrons and ions from the plasma and form ordered structures near the lower electrode.[2–5,19,20] Analytical estimates and numerical simulations have shown that in this case, random fluctuations in the charges on macroscopic particles can serve as a source of kinetic energy beyond the thermal energy and thereby cause a disequilibrium between the system of macroscopic particles and the neutral gas. Random fluctuations therefore impose a limit on the minimum kinetic energy of the macroscopic particles and under certain conditions can facilitate the melting of a plasma crystal, greatly reducing the coupling parameter Γ below its equilibrium value. Numerical estimates for the typical experimental plasma parameters show that the kinetic energy of the macroscopic particles can be several eV. Although our approach involves a considerable simplification of the physical picture, it provides a simple and clear illustration of the importance of random fluctuations in the charge on macroscopic particles in dusty plasmas.

This work was supported by the Russian Fund for Fundamental Research (Grants No. 97-02-17565 and No. 98-02-16825).

*)E-mail: ipdustpl@redline.ru

---

[1] V. N. Tsytovich, Usp. Fiz. Nauk **167**, 57 (1997).
[2] J. Chu and Lin I, Phys. Rev. Lett. **72**, 4009 (1994).
[3] H. Thomas, G. Morfill, V. Demmel *et al.*, Phys. Rev. Lett. **73**, 652 (1994).
[4] A. Melzer, T. Trottenberg, and A. Piel, Phys. Lett. A **191**, 301 (1994).
[5] Y. Hayashi and K. Tachibana, Jpn. J. Appl. Phys., Part 1 **33**, 804 (1994).
[6] V. E. Fortov, A. P. Nefedov, O. F. Petrov *et al.*, JETP Lett. **63**, 187 (1996).
[7] V. E. Fortov, A. P. Nefedov, V. M. Torchinskiĭ *et al.* JETP Lett. **64**, 92 (1996).
[8] A. P. Nefedov, O. F. Petrov, and V. E. Fortov, Usp. Fiz. Nauk **167**, 1215 (1997).
[9] M. Jana, A. Sen, and P. Kaw, Phys. Rev. E **48**, 3930 (1993).
[10] G. Morfill, E. Grun, and T. Johnson, Planet. Space Sci. **28**, 1087 (1980).
[11] C. Cui and J. Goree, IEEE Trans. Plasma Sci. **22**, 151 (1994).
[12] T. Matsoukas and M. Russell, J. Appl. Phys. **77**, 4285 (1995).
[13] M. Rosenberg, D. Mendis, and D. Sheenan, Planet. Space Sci. **24**, 1422 (1996).
[14] S. Khrapak, A. Nefedov, O. Petrov, and O. Vaulina, submitted to Phys. Rev. E (1999).
[15] J. Goree, Plasma Sources Sci. Technol. **3**, 400 (1994).
[16] O. Vaulina and S. Khrapak, in *Materials from the Conf.on the Physics of Low-temperature Plasmas* [in Russian], Izd. Petrozavodskogo Universiteta, Petrozavodsk (1998), p. 675.
[17] V. V. Zhakhovskiĭ, V. I. Molotkov, A. P. Nefedov, V. M. Torchinskiĭ, A. G. Khrapak, and V. E. Fortov, JETP Lett. **66**, 419 (1997).
[18] H. Totsuji, T. Kishimoto, Y. Inoue *et al.*, Phys. Lett. A **221**, 215 (1996).
[19] H. Thomas and G. E. Morfill, Nature (London) **379**, 806 (1996).
[20] A. Melzer, A. Homann, and A. Piel, Phys. Rev. E **53**, 2757 (1996).
[21] J. B. Pieper and J. Goree, Phys. Rev. Lett. **77**, 3137 (1996).
[22] M. Barnes, J. Keller, J. Forster *et al.*, Phys. Rev. Lett. **68**, 313 (1992).
[23] S. V. Vladimirov and M. Nambu, Phys. Rev. E **52**, R2172 (1995).
[24] F. Melandso and J. Goree, Phys. Rev. E **52**, 5312 (1995).

Translated by D. H. McNeill

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Nuclear reactions triggered by laser-accelerated high-energy ions

V. Yu. Bychenkov*) and V. T. Tikhonchuk

P. N. Lebedev Institute of Physics, Russian Academy of Sciences, 117924 Moscow, Russia

S. V. Tolokonnikov

Russian Research Center ''Kurchatov Institute,'' 123182 Moscow, Russia

(Submitted 27 January 1999)
Zh. Éksp. Teor. Fiz. **115**, 2080–2090 (June 1999)

A technique is suggested for triggering nuclear reactions by accelerating ions with a powerful ultrashort laser pulse in a plasma. The underlying idea of the suggested compact ''reactor'' is utilization of high-energy ions accelerated by the charge-separation electrostatic field in the direction perpendicular to the laser beam axis in a gas-filled capillary. Accelerated ions with energies of several MeV penetrating the target from the inside surface of a channel give rise to nuclear reactions which can be used to create a compact source of fast neutrons and neutrons of intermediate energies for generating various (short- and long-lived, light and heavy) isotopes, for generating gamma radiation over a broad energy range, for making sources of light ion and induced radioactivity. The yield of the corresponding nuclear reactions as a function of the laser beam parameters has been investigated. The suggested technique for triggering nuclear reactions provides a practical tool for studies of nuclear transformation on the pico- and nanosecond scales, which cannot be achieved using other methods. © *1999 American Institute of Physics.* [S1063-7761(99)01306-2]

1. INTRODUCTION

The unique possibilities afforded by present-day techniques for generating ultrashort laser pulses open up avenuse to their practical application in nuclear physics. They have been demonstrated, for example, in experiments on the laser acceleration of electrons and ignition of nuclear reactions by accelerated electrons.[1–4] Although the idea of using powerful ultrashort laser pulses to accelerate ions and trigger nuclear reactions was put forward[5] as early as the year 1987 (see also the review by Luther–Davis *et al.*[6]), practical steps to its implementation have been undertaken only recently. Experimental data[7,8] indicate that ions with energies $\simeq 0.5$ MeV are generated as a result of a laser pulse acting on a gaseous target at wavelengths of 1.06 and 0.53 μm with an intensity of $\simeq 5\times 10^{18}$ W/cm^2 and pulsewidth of 400–600 fs. Theoretical models[9,10] have demonstrated that methods of ion acceleration by laser pulses can find novel applications, especially in the field of nuclear physics. An advantage of ions is that, unlike electrons, they can participate in strong interactions, and therefore have much larger cross sections for nuclear transformations. The topic of this paper is the development of a theoretical background for the technique of laser–ion ignition of nuclear reactions.

The high ion energies achieved in the focus of laser beams shows that existing generators of ultrashort optical pulses suffice for effective ignition of nuclear reactions. In recent experiments[11] fluxes of $\simeq 7\times 10^7$ neutrons/sr were detected when a laser pulse with an energy of 20 J, a duration of 1.3 ps, and an intensity of 10^{19} W/cm^2 at a wavelength of 1.054 μm was focused on a film made of deuterated polystyrene or frozen deuterium. In another experiment,[12] a laser pulse of energy 200 mJ and duration 160 fs at a wavelength of 0.79 μm interacted with a plasma produced by a laser prepulse in a deuterated polyethylene film, and the neutron yield averaged over many pulses was $\simeq 140$ neutrons per pulse. However, the experimental configuration[12] was far from optimal, and the number of reacting particles was small owing to the small target size.

In order to increase the efficiency of the laser energy transfer to ions, it is expedient to use a plasma which is transparent to the laser radiation and a sufficiently short and powerful laser pulse. The plasma should be surrounded by the material that will be used as a target. One configuration that satisfies these conditions includes a target in the form of a capillary filled with a gas. The ions from a gas are accelerated by a laser pulse and interact with walls containing a material for the nuclear reaction. The gas density should be such that the plasma density is below the critical value for the laser radiation, and the thickness of the capillary wall should be of order the range of accelerated ions. The laser pulse power should be matched to its duration so that the latter equals the ion acceleration time at the laser beam focal spot.

The technique of laser-triggered nuclear reactions could be implemented by focusing a powerful ($I\lambda^2 \gtrsim 10^{19}$ W/cm$^2\cdot\mu$m^2) laser radiation inside a cylindrical (with a radius of several tens of microns) channel in a target (capillary) filled with a gas at the required pressure. The choice of the gas and target material (capillary walls), i.e., the selection of reacting pairs, is determined by the specific nuclear reactions that are to be ignited. Gas ions accelerated in the direction perpendicular to the beam axis penetrate the target and interact with its nuclei in a layer with a thickness equal to the ion mean free path in the target. High energy ions are generated by the Coulomb explosion,[13] which means that the ions are accelerated by the electrostatic field due to the charge separation. This field results from the ejection of electrons from the region of intense laser field in the radial direction by the ponderomotive force and/or the thermal pressure of electrons over the entire length of the laser channel. Creating an extended laser channel raises a possibility that the number of accelerated particles can be large enough to make possible a considerable number of nuclear reactions. According to Ref. 9, at a length of 1 mm of a channel with a diameter of 10 μm filled with gas at a density of 10^{20}–10^{21} cm^{-3}, the number of radially accelerated ions will be 10^{13}–10^{14}. Then even a reaction with a relatively low efficiency ($\sim 10^{-5}$) will yield $\sim 10^8$–10^9 particles per laser pulse, which can be of practical significance. This paper presents for the first time estimates of the yield of nuclear reactions triggered by fast ions and its dependence on the laser pulse parameters. The threshold laser intensities required for triggering the corresponding reactions have also been determined.

2. PHYSICAL MODEL

An intense laser pulse propagating through a gas ionizes it almost instantaneously owing to the tunneling effect, and ions are imparted a radial acceleration directed perpendicular to the beam axis due to the Coulomb explosion. The acceleration is determined by the ponderomotive force and the electron pressure, which is proportional to the average electron energy (effective temperature T_{eff}) in the region of intense laser field. (This effective temperature is generated through the nonlinear interaction between the laser radiation and electrons, which causes nonadiabatic electron heating during the laser pulse.) The acceleration of an ion of mass M and charge Z in the field of a linearly polarized laser beam is described by the following equation[9]:

$$\frac{du_i}{dt} = -\frac{Z}{M} mc^2 \nabla_r \sqrt{1+\frac{a^2}{2}} - \frac{Z}{M} \nabla_r T_{\text{eff}}, \qquad (1)$$

where u_i is the ion velocity, m is the electron mass, c is the speed of light, $a = 0.85\times 10^{-9}\lambda\sqrt{I}$ is the dimensionless vector-potential of the laser field, the intensity I is measured in W/cm^2 and the laser wavelength λ in microns.

At present there is no quantitatively accurate theory that would relate T_{eff} to the laser light intensity, and direct measurements of the electron energy spectrum in the subpicosecond range of pulse durations have also been impossible so far. Wilks *et al.*[14] suggested that at laser field intensities close to the relativistic value, $I\lambda^2 \gtrsim 10^{18}$ W/cm$^2\cdot\mu$m^2, the electron temperature is of the same order of magnitude as the energy of electron oscillations in the laser field, $T_{\text{eff}} \approx mc^2(\sqrt{1+a^2/2}-1)$. This is in qualitative agreement with numerical calculations.[12,15] We therefore assume that the contribution of the electron pressure to the accelerating force in Eq. (1) is comparable to the ponderomotive contribution.

An important parameter of ion acceleration by the laser field is the laser pulse duration τ. At a very short pulses ($\tau < d/u_i$, where d is the laser beam diameter), ions located in the focal spot do not have enough time to acquire the maximum possible energy, and if the laser pulse is too long, $\tau \gg d/u_i$, its energy is wasted, since ions leave the region of interaction long before the pulse termination. Thus, the condition of equality between τ and the ion transit time through the interaction region, d/u_i, determines the optimal laser pulse duration:

$$\tau \gtrsim 0.1 d[\mu\text{m}](\sqrt{1+a^2/2}-1)^{-1/2}\,\text{ps}. \qquad (2)$$

Here we use the approximate relation $A \simeq 2Z$, where A is the atomic number of accelerated ions.

By equating the force of electron pressure in Eq. (1) to the ponderomotive force, we obtain an approximate equation for the characteristic ion energy due to acceleration by the laser field:

$$\epsilon \approx Z(\sqrt{1+a^2/2}-1)\,\text{MeV}. \qquad (3)$$

In particular, for ions accelerated by a laser pulse of intensity $I\lambda^2 \simeq 10^{21}$ W/cm$^2\cdot\mu$m^2 and duration $\tau \simeq 100$ fs, with a channel diameter $d \simeq 10$ μm, Eq. (3) an estimate yields $\epsilon \approx 20Z$ MeV. Note that these parameters are close to those of the existing petawatt laser.[1] Ions of such energies can trigger nuclear reactions over a wide range. However, considerably lower ion energies suffice for many nuclear reactions, i.e.,lower laser field intensities are required. For example, the DT reaction is most efficient for particles colliding at a relative energy of $\simeq 0.1$ MeV, which can be achieved using a terawatt laser generating picosecond pulses (for example, $I\lambda^2 \simeq 10^{18}$ W/cm$^2\cdot\mu$m^2, $\tau \simeq 3$ ps, and $d \simeq 10$ μm). In general, we can assert that ion energies of several MeV are quite feasible in existing systems generating ultrashort optical pulses and are sufficient for triggering various nuclear reactions.

Fast ions passing through matter lose their energy to ionization and excitation of atoms (deceleration losses). The stopping length l_i is fairly adequately described by the Bethe formula:

$$l_i(\epsilon) = \int_0^{\epsilon} \frac{d\epsilon}{|d\epsilon/dr|}, \quad \frac{d\epsilon}{dr} \simeq -\frac{M^*}{m\epsilon} 2\pi e^4 Z^2 Z_a n_a \Lambda, \qquad (4)$$

where n_a is the density of target atoms, Z_a is their charge, M^* is the reduced mass of colliding particles, $\lambda = \ln(4m\epsilon/M^*J)$, and J is the average ionization energy of target atoms (for hydrogen $J = 14.9$ eV and for heavy atoms $J \simeq 10Z_a$ eV). At this point, we disregard relativistic effects for ions, corrections due to the binding energy of electrons in K and L shells of atoms, and the correction due to the so-called density effect. This approximation is sufficiently accurate in the range of ion energies up to 100 MeV.

Accelerated ions react with target nuclei in a layer of thickness equal to their stopping length, Eq. (4). The total number N of reactions can be estimated by the formula $N \approx N_i^0 n_a \sigma l_i$, where σ is the characteristic reaction cross section. Thus the number of reactions is smaller than the number N_i^0 of accelerated ions by a factor equal to the ratio of the ion mean free path associated with the reaction, $(\sigma n_a)^{-1}$, and the total stopping length l_i. With due account of Eqs. (3) and (4), we obtain

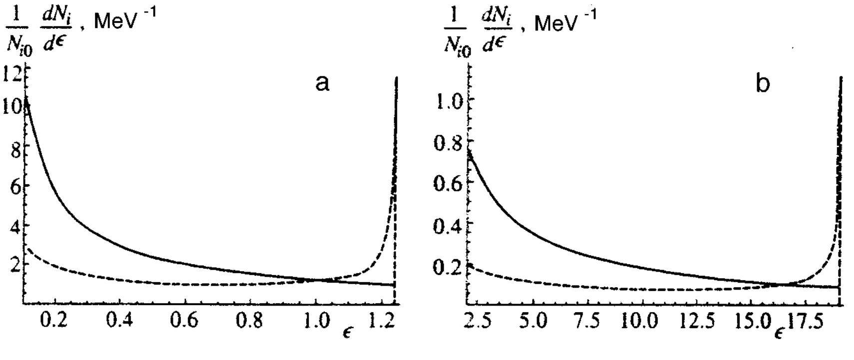

FIG. 1. Spectrum of protons $dN_i/d\epsilon N_i^0$ accelerated by a laser pulses of intensities (a) $I_0\lambda^2=10^{19}$ W/cm$^2\cdot\mu$m^2 and (b) 10^{21} W/cm$^2\cdot\mu$m^2. The solid and dashed lines correspond to Gaussian and super-Gaussian laser intensity profiles, $I_0\exp[-(2r/d)^2]$ and $I_0\exp[-(2r/d)^4]$, respectively.

$$N\approx N_i^0\frac{4\sigma}{AZ_a\Lambda}\left(\sqrt{1+\frac{a^2}{2}}-1\right)^2, \tag{5}$$

where the reaction cross section is measured in millibarns. Note that the number of nuclear reactions is independent of the charge of accelerated ions and the target material density. The latter circumstance means that a gaseous medium can be also used as a target. In practice, such a gas target can be realized using a thin-wall capillary with a wall thickness that has a little decelerating effect on fast ions. In this case the reaction time $\simeq l_i/u_i$, however, is longer, just as the size $l_i\propto n_a^{-1}$ of the reaction region is larger, hence larger targets are required. The latter condition has not been fulfilled in the experiment conducted by Pretzler *et al.*[12] This can be one of the reasons why the number of reactions detected in that experiment was low.

3. SPECTRUM OF ACCELERATED IONS

Equations (2) and (5) are convenient for approximate estimates of the efficiency of nuclear reactions and their optimization. They allow one to estimate the order of magnitude of the nuclear reaction yield, given the laser pulse basic parameters. At the same time, these estimates are based on a specific characteristic energy of fast ions (3), whereas in reality there is a certain energy distribution of ions, $dN_i/d\epsilon$, whose shape depends on the temporal pulse profile and its spatial distribution.[9] Many reaction cross sections have resonances, so the reaction yield can strongly depend on the shape of the particle energy distribution. An expression for the number of reactions which is more accurate than Eq. (5) contains the cross section averaged over the ion energy spectrum, the so-called overlap integral:

$$N=n_a\int_0^\infty d\epsilon\frac{dN_i}{d\epsilon}\int_0^\epsilon d\epsilon'\,\sigma(\epsilon')\left|\frac{d\epsilon'}{dr}\right|^{-1}, \tag{6}$$

where the stopping power $d\epsilon/dr$ is given by Eq. (4).

Under the conditions when accelerated particles can leave the focal region during the laser pulse duration, it follows from Eq. (1) that a universal momentum distribution function of ions independent of the pulse duration is established:

$$\frac{dN_i}{d\epsilon}\simeq\frac{2\pi r^*(\epsilon)n_i^0(2Zmc^2+\epsilon)}{Z^2m^2c^4|da^2/dr|_{r=r^*(\epsilon)}}, \tag{7}$$

where n_i^0 is the initial ion density, the energy-dependent radius r^* is found by solving the equation

$$\epsilon=\Phi(r^*)\equiv2Zmc^2(\sqrt{1+a(r^*)^2/2}-1),$$

and the spectrum is cut off at higher energies in accordance with the formula $\epsilon\leqslant\max[\Phi(r)]$. In the case of a Gaussian radial distribution of the laser beam intensity, $I=I_0\times\exp(-4r^2/d^2)$, Eq. (7) yields a decreasing energy distribution:

$$\frac{dN_i}{d\epsilon}\simeq2\,N_i^0\frac{\epsilon+Z}{\epsilon(\epsilon+2Z)}\Theta(\epsilon_{\max}-\epsilon),$$
$$\epsilon_{\max}=Z(\sqrt{1+0.4I_0\lambda^2}-1). \tag{8}$$

Here $N_i^0=\pi n_i^0d^2/4$, the parameters ϵ, I_0, and λ are measured in MeV, 10^{18} W/cm^2, and microns, respectively, Θ is the unit step-function, and the lower bound of the energy range is determined by the condition $\epsilon>\Phi(R)$, where R is the radius of the plasma generated during fast tunneling ionization by the laser pulse preceding acceleration of ions. Since $R\gg d$ actually holds (Ref. 9), this energy is negligible in comparison with the maximum ion energy $\epsilon_{\max}$. In comparison with the dependence ϵ^{-1} over the ranges $\epsilon\ll Z$ and $\epsilon\gg Z$, Eq. (8) shows flattening of the fast-ion spectrum in the energy range $\epsilon\sim Z$, which enhances the reaction yield in this range.

Note that the ion energy distribution and hence the reaction yield depends sensitively on the spatial distribution of the laser field intensity, which in general can significantly differ from the intensity distribution in the incident laser beam owing to its interaction with the plasma. By selecting an appropriate laser intensity profile, one can maximize the triggering efficiency of a nuclear reaction. In particular, for most nuclear reactions, an energy of several MeV is required along with fairly large numbers of ions with these energies. Therefore, given the limited energies of pulses generated by present-day lasers, it is advisable to shape the laser beam profile to redistribute energy to the faster accelerated ions. One example is a super-Gaussian laser pulse profile, $I=I_0\exp[-(2r/d)^4]$, for which the analogue of the spectrum (8) is

$$\frac{dN_i}{d\epsilon}\simeq N_i^0\frac{\epsilon+Z}{\epsilon(\epsilon+2Z)}\left[\ln\frac{\epsilon_{\max}(\epsilon_{\max}+2Z)}{\epsilon(\epsilon+2Z)}\right]^{-1/2}\Theta(\epsilon_{\max}-\epsilon). \tag{9}$$

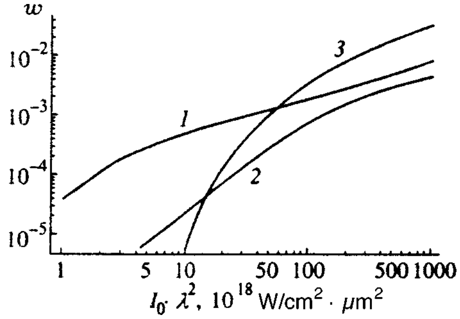

FIG. 2. Neutron yield in reactions (*1*) $T(d,n)^4He$, (*2*) $D(d,n)^3He$, and (*3*) T $(p,n)^3He$ versus laser intensity.

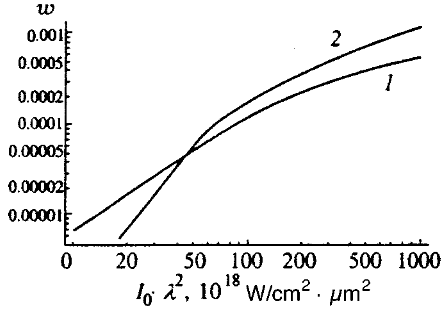

FIG. 3. Yields (in units of 10^{18} W/cm^2) of reactions (*1*) $^6Li(d,\alpha)^4He$ and (*2*) $^7Li(p,\alpha)^4He$ generating fast α particles as functions of the laser power density.

This is to be compared with the spectrum (8) corresponding to the Gaussian distribution in Fig. 1 for (a) relativistically strong, $a\sim 1$, and (b) ultrarelativistic, $a \gg 1$, laser pulses accelerating protons. The plots in Fig. 1 clearly show that the ion spectrum (9) generated by a super-Gaussian laser beam is richer in ions with energies $\epsilon \lesssim \epsilon_{\max}$. The number of slow ions is accordingly suppressed. The latter tendency is well illustrated by going over to a parabolic laser intensity profile, $I = I_0(1-4r^2/d^2)$, which results in a spectrum increasing with the energy and proportional to $(\epsilon + Z)\Theta(\epsilon_{\max}-\epsilon)$.

4. YIELDS OF NUCLEAR REACTIONS

Now let us discuss yields $w = N/N_i^0$ of reactions which are of interest for possible applications. For the sake of definiteness, our analysis will be based on formulas for a Gaussian distribution of laser intensity and data on the cross sections of nuclear reactions from the EXFOR data file.[18] Most importantly, the laser technique of ion acceleration provides a microsource of short pulses of neutrons of different energies ranging between hundreds of kiloelectrovolts and tens of MeV. Figure 2 shows the yield of fast neutrons ($\geqslant 14.1$ MeV) and neutrons of moderate energies (from several hundreds of keV to several MeV) as functions of the laser intensity I_0 in the reactions $T(d,n)^4He$ and $T(p,n)^3He$, respectively. Figure 2 also shows the yield of the reaction D $(d,n)^3He$, for which Norreys *et al.*[11] reported efficient production of neutrons with a characteristic energy of $\simeq 2.45$ MeV.

These authors[11] estimated the total number of accelerated ions to be $\sim 10^{14}$. According to curve *2* in Fig. 2, the characteristic yield of the $D(d,n)^3He$ reaction under the conditions of the experiment[11] is $\sim 10^{-5}$. This means that for the parameters under discussion, one should expect generation of $\sim 10^9$ neutrons per laser pulse, which is in qualitative agreement with the experimental data.

Figure 2 shows that the DT reaction is triggered at relatively small (comparing to modern experimental facilities) laser radiation intensities. At $I_0\lambda^2 = 10^{18}$ W/cm$^2 \cdot \mu$m^2, the reaction yield is already 4×10^{-5}. The proton charge exchange reaction, $T(p,n)^3He$, has a clearly defined threshold at 10^{19} W/cm$^2 \cdot \mu$m^2, but it is characterized by a higher yield of $\gtrsim 0.01$ for laser pulses of higher intensities.

Some fusion reactions triggered by laser pulses in this manner can be used to accelerate light ions. Light ions are accelerated as a result of the coversion of the energy generated in the fusion reaction into the high kinetic energy of reaction products. One example is the reaction $^6Li(d,\alpha)^4He$, which allows one to accelerate α-particles to an energy of 22.4 MeV; another is the reaction $^7Li(p,\alpha)^4He$, for which the maximum energy of generated α-particles is 17.3 MeV. Figure 3 plots yields of these reactions as functions of the laser beam intensity. They have a threshold at $\sim 10^{19}$ W/cm$^2 \cdot \mu$m^2 and lower yields than the thermonuclear reactions discussed above.

In practice, reactions in which artificial isotopes are produced are of great interest. Curves *1* and *2* in Fig. 4 show the efficiency of production of β^+- and β^--active light isotopes [reactions $^{14}N(p,n)^{14}O$ and $^{26}Mg(d,\alpha)^{24}Na$], and curves *3* and *4* show the production efficiency of heavier isotopes [reactions $^{124}Te(p,n)^{124}I$ and $^{95}Mo(p,n)^{95}Tc$]. These last two reactions, which create β^+-active isotopes, are interesting for medical applications, in particular, positron tomography. The long-lived isotope ^{24}Na is often used as a source of secondary gamma rays. The yields of these reactions are relatively high ($w \gtrsim 10^{-5}$), although they require higher laser field intensities: $I_0\lambda^2 \gtrsim 10^{19}$–$10^{20}$ W/cm$^2 \cdot \mu$m^2.

At a sufficiently high laser field intensity, $I_0\lambda^2 \gtrsim 10^{20}$ W/cm$^2 \cdot \mu$m^2, which is required to drive an incident particle over the fission barrier (7–8 MeV) and provide sufficient penetrability of the Coulomb barrier $[\approx 1.4ZZ_a/(A^{1/3}+A_a^{1/3})$ MeV$\sim$20 MeV], laser-accelerated ions trigger fission reactions in targets made of heavy elements. Figure 5 illustrates this effect on the examples of uranium-238 and

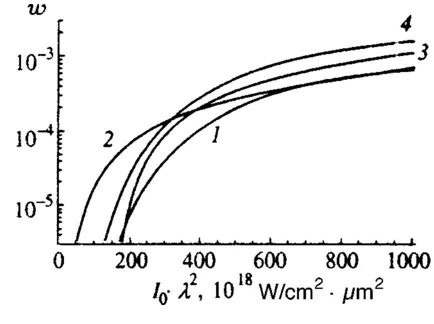

FIG. 4. Yields (in units of 10^{18} W/cm^2) of isotopes generated in reactions $^{14}N(p,n)^{14}O$ and $^{26}Mg(d,\alpha)^{24}Na$ (curves *1* and *2*, respectively), and $^{124}Te(p,n)^{124}I$ and $^{95}Mo(p,n)^{95}Tc$ (curves *3* and *4*, respectively) as functions of the laser power density.

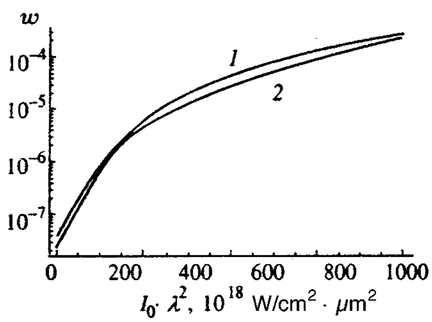

FIG. 5. Yields of fission reactions involving ^{238}U and ^{232}Th triggered by protons (curves *1* and *2*, respectively) as functions of the laser power density.

thorium-232 fission triggered by protons. Unlike other proton sources, the laser acceleration technique allows one to generate induced activity in an extremely short time. Specifically, a proton of energy $\sim$10 MeV traverses the stopping length (4) in a uranium salt, for example, in a time of $\sim$10 ps, whereas alternative proton sources generate pulses about one nanosecond long or longer. Therefore the laser acceleration technique offers a unique opportunity for fundamental research in the field of nuclear physics in a new, subnanosecond range of times. In particular, it becomes possible to study a decay of radioactive fission products on a time scale inaccessible to other triggering techniques. The suggested technique opens prospects for determination of lifetimes of ultrashort-lived isotopes (mostly $\beta^{\pm}$-active) which are far from the stability band, and for investigating a strongly nonequilibrium system of isotopes relaxing to the secular equilibrium. Although the yields are relatively low and the thresholds are high, an ignition of fission reactions in laboratory facilities is a unique experiment because of the high energy yield, $\gtrsim$200 MeV per one fission event. Such a reaction can lead to a thermal explosion, and these experiments will, apparently, require radiation safety measures because of accumulation of radioactive isotopes and neutrons emitted in fission reactions.

Finally, let us consider examples of generation of hard gamma rays and generation of photons with a wide spectral range, including the softer component, by allowing laser-accelerated protons to hit a target. Thus, the proton-capture reaction T$(p,\gamma)^4$He (curve *1* in Fig. 6) is a source of hard gamma rays in which the released energy $\gtrsim$19.8 MeV is carried by three or four photons, while in a similar reaction ^{11}B$(p,\gamma)^{12}$C (curve *2* in Fig. 6), gamma rays are emitted over a broad spectrum with a maximum total energy $\epsilon_{\gamma\,\max}$ determined by the maximum proton energy $\epsilon_{p\,\max}$, namely

$$\epsilon_{\gamma\,\max}\approx 16\ \text{MeV}+\epsilon_{p\,\max}\approx(15+\sqrt{1+0.4I_0\lambda^2})\ \text{MeV},$$

where the laser intensity I_0 is measured in units of 10^{18} W/cm^2 and the laser wavelength in microns. In targets made of heavier materials, gamma rays with a softer spectrum and lower total energy are generated. Because these reactions are electromagnetic processes, the yields of these reactions are lower (since the electromagnetic constant $e^2/\hbar c=1/137$ is small) than those of fusion reactions, which are controlled by the strong interaction. This circumstance, however, does not rule out the possibility of using these reactions in research on the gamma-laser problem.

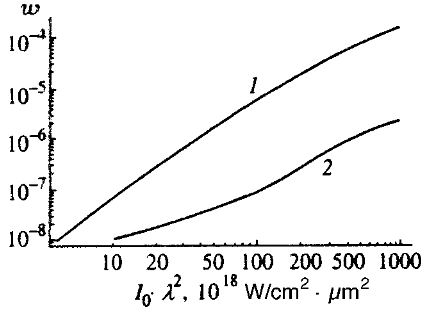

FIG. 6. Yields of reactions generating hard gamma rays [T$(p,\gamma)^4$He, curve *1*] and gamma rays of a broad-range spectrum [^{11}B$(p,\gamma)^{12}$C, curve *2*] as functions of the laser power density.

5. CONCLUSION

In this paper, we have suggested a scheme for a compact nuclear ''reactor'' based on laser acceleration of ions. This method of nuclear reaction ignition can be implemented in practice using available powerful lasers (10–10^3 TW) generating ultrashort pulses. They can become powerful tools of nuclear research. The effect of relativistic self-focusing of a laser beam can substantially increase the length of its caustic, thereby increasing the total number of accelerated ions. For this reason it is desirable to meet the conditions for self-focusing. It is also desirable to reduce the laser wavelength and at the same time increase the pressure in the channel, since this will allow one to generate more accelerated ions, and hence a higher yield of nuclear reactions. Note also that, even in the absence of self-focusing a capillary can trap a laser pulse over a fairly large distance owing to its waveguide properties. This was demonstrated in recent experiments on propagation of laser pulses in hollow capillaries,[16] where the beam was focused over a length of up to 10 mm. Investigations on the problem of laser hole boring in a plasma as a part of the fast ignitor program[17] will probably show that it is possible to dispense with prefabricated channels if the conditions for burning sufficiently long cavities with a high plasma density in them can be determined. The relativistic effect of plasma transparency can also be a favorable factor. Results of computer simulation[15] of formation of a plasma channel with a length of up to 20 μm and a density of about double the critical value in a solid-state target using a laser relativistic beam with an intensity $I\lambda^2\sim10^{20}$ W/cm$^2\cdot\mu$m^2 also provide evidence in favor of the hole-boring technique.

We have quantitatively estimated the yields of various reactions which can have numerous applications and demonstrated the possibility of creating compact sources of fast neutrons and neutrons of moderate energies, production of various isotopes, generation of hard gamma rays and photons with a wide energy spectrum, creation of sources of light ions and artificial radioactivity. The proposed scheme of laser ignition of nuclear reactions provides a practicable tool for fundamental research in nuclear reactions on time scales of several tens of picoseconds, which cannot be studied by other methods. We have demonstrated a possibility of generating gamma-radiation that can be used as a base for a concerted effort in designing gamma-lasers.

If a solid-state target is used, the suggested technique allows one to create an almost point-like source of activity with a size of several tens or hundreds of microns, which is a unique achievement. This makes the method of ion-laser ignition of nuclear reactions a promising tool for investigation of a wide range of phenomena in nuclear physics and will permit researchers to fabricate radioactive microsamples with prescribed properties. One possible practical application is, in our opinion, the production of isotopes for use in medicine and biology. It is known that the production cost of some isotopes is very high, so it is desirable to develop alternative techniques. Since the required activity of such samples is not high, the needed quantity of isotopes can be accumulated in one or several laser pulses.

Free access to the EXFOR data base (Experimental Nuclear Reaction Data File)[18] has been essential for completion of our research. The work was supported by the Russian Fund for Fundamental Research (Grant No. 99-02-17267).

*)E-mail: bychenk@sci.lebedev.ru

---

[1] G. A. Mourou, C. P. J. Barty, and M. D. Perry, Phys. Today **51**(1), 22 (1998).

[2] C. Bula, K. T. McDonald, E. J. Prebys *et al.*, Phys. Rev. Lett. **76**, 3116 (1996); D. L. Burke, R. C. Field, G. Horton-Smith *et al.*, Phys. Rev. Lett. **79**, 1626 (1997).

[3] P. L. Shkolnikov and A. E. Kaplan, J. Nonlinear Opt. Phys. Mater. **6**, 161 (1997).

[4] D. A. Gryaznykh, Ya. Z. Kandiev, and V. A. Lykov, JETP Lett. **67**, 257 (1998).

[5] M. D. Perry, E. M. Campbell, J. T. Hunt *et al.*, LLNL Report, *Ultra-High Brightness* ($>10^{21}$ W/cm^2) *Laser Facility* (1987).

[6] B. Luther-Davis, E. G. Gamaly, Y. Wang *et al.*, Laser Phys. **1**, 325 (1991).

[7] G. S. Sarkisov, V. Yu. Bychenkov, V. T. Tikhonchuk *et al.*, JETP Lett. **66**, 828 (1997).

[8] J. Fuchs, G. Malka, J. C. Adam *et al.*, Phys. Rev. Lett. **80**, 1658 (1998).

[9] G. S. Sarkisov, V. Yu. Bychenkov, V. N. Novikov *et al.*, Phys. Rev. E **59**(6), 7042 (1999).

[10] G. S. Sarkisov, V. Yu. Bychenkov, and V. T. Tikhonchuk, JETP Lett. **69**, 20 (1999).

[11] P. A. Norreys, A. P. Fews, F. N. Beg *et al.*, Plasma Phys. Controlled Fusion **40**, 175 (1998).

[12] G. Pretzler, A. Saemann, A. Pukhov *et al.*, Phys. Rev. E **58**, 1165 (1998).

[13] N. H. Burnett and G. D. Enright, IEEE J. Quantum Electron. **26**, 1797 (1990).

[14] S. C. Wilks, W. L. Kruer, M. Tabak, and A. B. Langdon, Phys. Rev. Lett. **69**, 1383 (1992); G. Malka and J. L. Miguel, Phys. Rev. Lett. **77**, 75 (1996).

[15] A. Pukhov and J. Meyer-ter-Vehn, Phys. Rev. Lett. **79**, 2686 (1997).

[16] M. Borghesi, A. J. Mackinnon, R. Gaillard *et al.*, Phys. Rev. E **57**, 4899 (1998).

[17] M. Tabak, J. Hammer, M. E. Glinsky *et al.*, Phys. Plasmas **1**, 1626 (1994).

[18] Experimental Nuclear Reaction Data File, http://www.nndc.bnl.gov/nndc/exfor/.

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Rarefaction wave and gravitational equilibrium in a two-phase liquid–vapor medium

N. A. Inogamov*) and S. I. Anisimov†)

L. D. Landau Institute for Theoretical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

B. Retfeld‡)

Technical University, Braunschweig, Germany

(Submitted 5 March 1998)
Zh. Éksp. Teor. Fiz. **115**, 2091–2105 (June 1999)

The problems studied in this paper involve the action of laser radiation or a particle beam on a condensed material. Such an interaction produces a hot corona, and the recoil momentum accelerates the cold matter. In the coordinate frame tied to the accelerated target, the acceleration is equivalent to the acceleration of gravity. For this reason, the density distribution ρ is hydrostatic in the zeroth approximation. In this paper the structure of such a flow is studied for a two-phase equation of state. It is shown that instead of a power-law density profile, which obtains for a constant specific-heat ratio, a complicated distribution containing a region with a sharp variation of ρ arises. Similar characteristics of the density profile arise with isochoric heating of matter by an ultrashort laser pulse and the subsequent expansion of the heated layer. The formation of a rarefaction wave and the interaction of oppositely propagating rarefaction waves in a two-phase medium are studied. It is very important to take account of the two-phase nature of the material, since conditions ($p_a \sim 1$ Mbar) are often realized under which the foil material comes after expansion into the two-phase region of the phase diagram.

[S1063-7761(99)01406-7]

1. INTRODUCTION

Let us consider the effect of a powerful laser or ion beam on an initially solid layer. The region heated as a result of theabsorption of beam energy has a definite mass thickness. If the thickness of the initial layer is greater than this value, then the flow separates into hot and cold zones.[1–4] We shall call them the coronal and foil zones, respectively. Layers ranging in thickness from 1 to 100 μm are used in experiments on acceleration of foils. An ablation front—a region where the entropy s of the material varies sharply—separates the corona and the foil. The surface $x=x_a$, where the function $-d\ln s/dx$ reachesits maximum value, gives the position of the front. Ordinarily, $-H(d\ln s/dx)_{\max}\gg 1$ holds, where H is the foil thickness and the x axis is directed toward the vacuum edge, which is the backside of the foil.

We shall give the typical values of the experimental parameters for which the effects due to the two-phase nature of the matter will be important. Aluminum in a shock wave melts at 1 Mbar.[4] Ablation pressures 1–3 Mbar are reached with incident laser radiation intensities $I=10^{12}-2\cdot 10^{13}$ W/cm^2 at wavelength $\lambda=200$ nm and $I=3\cdot 10^{12}-4\cdot 10^{13}$ W/cm^2 at $\lambda=500$ nm.[2,5]The characteristic pulse durations are $\tau=1-10$ ns. These effects are also important for subpicosecond pulses interacting with matter, if the pulse intensity is such that heating up to temperatures much higher than the melting temperature occurs. For femtosecond pulses the thickness of the heated layers is much less ($H\sim 0.1$ μm). For heating of a material by a charged-particle beam the target size and pulse duration are somewhat greater than for laser heating. There is an extensive literature devoted to these questions (see Refs. 6 and 7).

Hot gas in the corona produces pressure p_a which accelerates the foil. The ablation front ''absorbs'' the foil material with a velocity v_a. It is very important that this velocity is ordinarily low compared with the sound velocity c_s in the foil. For this reason, for a pulse withduration $\tau\sim H/v_a$ the time τ is sufficient to establish an atmosphere (hydrostatic equilibrium). It is established in time $\sim H/c_s$. Neglecting the velocity v_a compared with c_s, we arrive at a simple model in which the ablation boundary of the foil is a Lagrangian layer, where a pressure p_a is maintained. After the pressure p_a is switched on a shock wave punches through the initially uniform foil. If the pressure p_a is constant, then the shock wave is stationary. A uniform entropic background remains behind such a wave.

At some moment in time the shock wave reaches the backside of the layer.The flow arising here is described by the solution of the problem of the decay of a discontinuity.[8–10] As a result of the decomposition, a rarefaction wave passes into the interior volume of the foil. As mentioned above, an atmospheric distribution is established asymptotically in a time of order H/c_s. Thus, the problem of a rarefaction wave (Sec. 3) and the structure of the atmosphere (Sec. 2) are of interest in connection with the problem of the effect of laser radiation and particle beams on matter. The problem of the interaction of rarefaction waves is also related to these problems (Sec. 4). We shall consider the atmosphere problem first. We confine ourselves to a thermodynamic (i.e., equilibrium) analysis. It is valid in the ''dirty'' situation

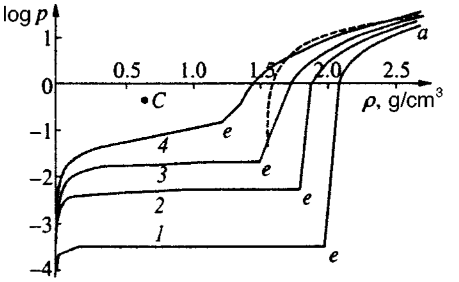

FIG. 1. Curves *1*, *2*, *3*, and *4* — Al isentropes with $s=3.05$, 3.48, 3.76, and 4.12 J/g·K.

(short kinetic times) and for not too small values of H and τ. From the analysis follows that the switch to two-phases causes the density profile $\rho(x)$ to vary sharply compared to the single-phase case.

2. TWO-PHASE HYDROSTATICS

2.1. *Thermodynamic description: phase equilibrium curve and isentropes.* Isochoric heating of thin layers by an ultrashort laser pulse or by ashock wave arising when pressure is switched on produces a uniform distribution of the entropy *s*. In this connection, it is of interest to analyze an isentropic two-phase atmosphere. We shall consider as anexample the isentropes of aluminum (Fig. 1). They have been calculated from wide-range equations of state.[4,11,12] The kinks *e* in the isentropes are due to the intersections of the binodal bounding the two-phase region on the liquid side. The pressure is given in GPa and density in g/cm^3. The critical parameters are $T_c=8000$ K, $p_c=0.45$ GPa, $\rho_c=0.64$ g/cm^3, and $s_c=4.83$ J/g·K (Ref.4, p. 345). All four isentropes lie below the critical point. The latter is marked by the letter *C* in Fig. 1. The point *C* together with the points *e* form the curve of boiling.

For simplicity, in all variants the density of solid aluminum $\rho_{ss}=2.7$ g/cm^3 is taken as the initial density. The initial pressure and temperature on the isentropes *1*, *2*, *3*, and *4* were, respectively, 19.8, 23.3, 31.3, and 36.5 GPa and 3960, 5990, 7610, and 10190 K, and the values of *s* on them were, respectively, 3.05, 3.48, 3.76, and 4.12 J/g·K.1) These isentropes were used to calculate the hydrostatic equilibrium (Sec. 2.2) and to calculate the rarefaction waves (Secs. 3 and 4). In the case of the atmosphere, the layer is heated isochorically, the pressure in the layer becomes p_a, and then this pressure is maintained on one of the boundaries of the layer. For rarefaction waves the material is heated isochorically and then expands.

2.2. *Structure of the atmosphere.* We write the equation of hydrostatics $p_x=-\rho g$ in the form $(\partial p/\partial\rho)_s d\rho/dX=-\rho$, where $X=gx$. Hence

$$-\int_{\rho_{ss}}^{\rho}[c(\rho')]^2\frac{d\rho'}{\rho'}=X, \tag{2.1}$$

where $c(\rho)$ is the sound velocity on a given isentrope. In Eq. (2.1) the ablation front X_a is chosen as the origin for the X axis$(X_a=0)$. The maximum density $\rho_{ss}=\rho_a$ is reached on the ablation front. The equation (2.1) implicitly determines the profile $\rho(X)$ and together with it the distributions of all other thermodynamic variables. Let us find the function $s(\rho)$ along the isentrope and integrate (2.1). Theresults of the numerical integration are displayed in Fig. 2. The curves *1–4* refer to the isentropes *1–4* presented in Fig. 1. The density is given in g/cm^3, and the quantity $X=gx$ (the ''height'') in (km/s)^2 is plotted along the ordinate. The atmosphere with a higher pressure p_a extends to greater heights $X=gx$. The small difference between the distributions at $X=0$ is due to the approximation error. The segments $a-e$ of the curves contain ≈ 10 approximation points, each curve *1–4* containing 50–100 such points. Here *a* is the surface on which the pressure p_a is maintained, and *e* is the vaporization surface (compare the correspondence between the points *a* and *e* in Figs. 1 and 2).

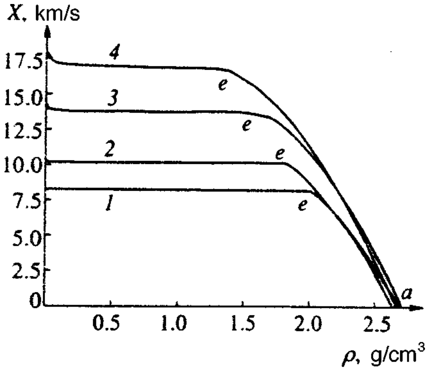

FIG. 2. Profiles $\rho(X)$ in accelerated foils for different values of the near-criticality parameter p_c/p_e.

The most important feature of two-phase liquid–vapor systems is the existence of a narrow zone where the density changes sharply. The function $\rho(X)$ varies smoothly on the segments $a-e$ and very rapidly (but continuously) decreases to values much less than ρ_e in a narrow layer bounding the surface *e* at the top (here ρ_e is the value of ρ at the point *e*).

The vaporized part of the atmosphere lies above the surface *e*. The fraction of the liquid phase there decreases rapidly with *X*. A transition occurs from a bubble to a droplet state (vapor). The two-phase layer is very thin compared with X_e-X_a, where X_e is the coordinate of the point *e*. This means that its spatial scale $(1/|d\ln\rho(X_e+0)/dX|)$ is small. Formally, the two-phase layer extends to infinity. It remains two-phase at all heights, since once the isentrope enters the two-phase region it remains there as ρ decreases further. The pressure in this layer near the surface *e* decreases slowly (the sections of the isentropes are approximately parallel to the abscissa in Fig. 1). In this connection we call attention to Refs. 13 and 14, where the internal structure of two-phase matter and the influence of the two phases on the motion of the medium in the near-critical region are investigated.

Along the binodal toward the critical point, the density ρ_e decreases and the entropy *s* increases. For $s\approx s_c$ and above the profiles of the variables become quite smooth (no sharp gradients along *X*). For $s>s_c$ ($s_c=4.83$ J/g·K; Refs. 4, 11, 12) the isentropes, on expansion, enter the two-phase region through the curve of condensation (the branch of the curve of phase equilibrium that lies to the left of the point *C*

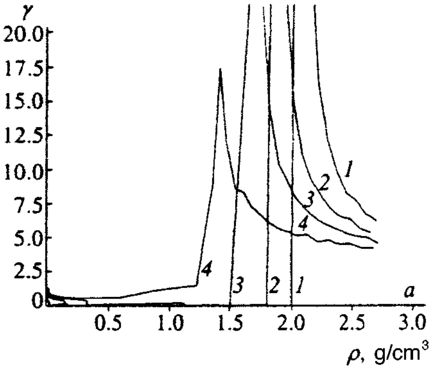

FIG. 3. $\gamma(\rho)$ on the isentropes *1–4*.

in Fig. 1). We note that all adiabats terminate in the two-phase region.

It is important to note that the pressure p_e at which vaporizationstarts is low. The ratio p_a/p_e ranges from 2.5 to 5 orders of magnitude forthe curves presented in Fig. 1. Therefore the mass m_e evaporatedfrom the backside of the foil (''tail'') is small, $m_e/m_{\text{total}}=p_e/p_a$. The tail grows in mass and thickness with s (compare the curves *1–4* in Fig. 2).

For a large decrease of pressure on the segment $a-e$ the density changes very little ($\rho_e/\rho_a=0.5-0.7$). The appearance of the above-described density profile is due to these two circumstances. Qualitatively, it resembles the density jump at the edge of a liquid layer in a gravitational field. Indeed, if the liquid is bounded above by a vacuum (we neglect vaporization), then at its edge the pressure vanishes and the density abruptly changes from a finite value to zero.

It is obvious that if the point a is close to the point e ($p_a \simeq p_e$, $p_a > p_e$, $s_a = s_e$), then the mass and thickness of the liquid layer should be of the same order of magnitude as those of the vaporized layer.

The strong decrease in p with a small change in ρ results because in the single-phase condensed state the values of the derivative $\gamma=(\partial \ln p/\partial \ln \rho)_s$ are large compared to 1. Plots of $\gamma(\rho)$ and $\gamma(X)$ are presented in Figs. 3 and 4, respectively. The values of γ grow as $\rho \rightarrow \rho_e$. For example, $\gamma(\rho_e+0)\approx 200$ on the curve *1* in Figs. 3 and 4. Conversely, in the two-phase region these functions near the point ρ_e are small compared to 1 in the case of the isentropes *1*, *2*, and *3*.

2.3. *Analytic models of the atmosphere.* We shall describe the profile $\rho(x)$ using simple analytic models. First, we assume $\gamma(\rho)\equiv$ const (polytropic atmosphere). Then, it is easy to see that a power-law dependence follows from the equation of hydrostatics. For example, the distribution ρ has the form

$$\rho/\rho_{ss}=[1-(\gamma-1)X/c_{sa}^2]^{1/(\gamma-1)}=(1-2X/Nc_{sa}^2)^{N/2}, \tag{2.2}$$

where c_{sa} is the sound velocity at the bottom of the atmosphere and $N=2/(\gamma-1)$. The index sa shows that the sound speed is determined by the ablation pressure.

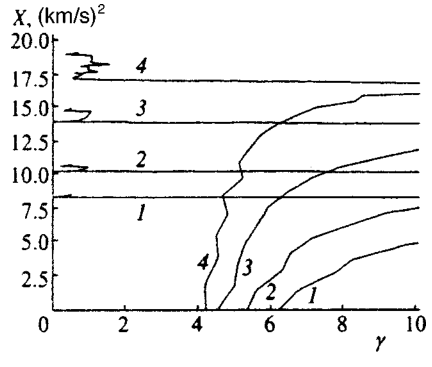

FIG. 4. $\gamma(X)$. The numbers *1–4* refer to the isentropes *1–4* in Fig. 1.

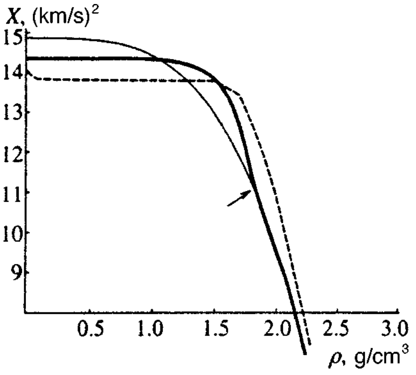

FIG. 5. Comparison of the profiles obtained with analytic and numerical fits of the isentrope. The single-layer (light solid curve) and two-layer (heavy solid curve) polytropes and the numerical fit (thin dashes) are presented for the example of isentrope 3. The arrow marks the point where the solutions are matched.

In the second model we approximate the real isentrope by two polytropes.In the layer $X_a<X<X_i$ the function $\gamma(X)\equiv\gamma_a$, whileabove this layer, for $X>X_t$, we assume that $\gamma(X)\equiv\gamma_t$, where $\gamma_t>\gamma_a$. Such a step function gives a better approximation of the real isentrope on which, as one can see from Figs. 3 and 4, in the single-phase condensed region the exponent γ increases with decreasing ρ. At the matching point X_t the densities and pressures calculated using formulas of the type (2.2) are matched.

In Fig. 5 the models 1 and 2 are compared with the numerical calculations. Here the light solid line refers to model 1, the heavy line refers to model 2, and the dashed line refers to the integral (2.1). The arrow marks the matching point of the solutions. In this example $s=3.76$ J/g·K. The values taken from the tabulated isentrope ($c_{sa}=7.2$ km/s, $\gamma_a=4.5$, $X_t=11$ (km/s)^2, $\rho_t=1.837$ g/cm^3, $p_t=5.527$ GPa, $\gamma_t=10$, $c_t=5.4$ km/s) were substituted into Eq. (2.2). The power-law distributions (2.2) for $\gamma>1$ vanish at a finite height $X_v=c^2/(\gamma-1)$. For large γ they decrease rapidly near the vacuum edge X_v. Nonetheless, this decrease follows a power law, $\rho\propto(X_v-X)^{1/(\gamma-1)}$ and does not cut off at the point X_e (compare the dashed and solid curves in Fig. 5 in the region ofrapid falloff).

We shall now consider a ''truncated'' polytrope. This models the zone of very large γ near and to the right of the point e (compare Figs. 1, 3, and 4). This is a particular case of a two-step polytrope (model 2). The finite value of γ is replaced at the point with finite ρ_t and p_t by the value $\gamma=\infty$ (incompressible liquid). In this case Eq. (2.2) describes the segment $a-t$ of the profile. At the matching point t a plateau—the layer $t-e$ of thickness $\Delta X=p_t/\rho_t$ where the density ρ_t is constant and the pressure decreases linearly

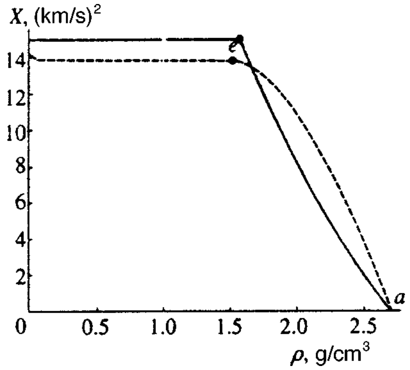

FIG. 6. Foil profiles for model 3 (2.3)–(2.4) (solid curve) and for a real isentrope (dashed curve). In this example $s = 3.76$ J/g·K (isentrope *3*).

from p_t to zero—is matched to it continuously as a function of ρ and p. The density vanishes abruptly at the point e.

The form of this profile is essentially identical to the profile shown by the heavy line in Fig. 5. It is only necessary to substitute a step for the smoothly decreasing function in the region above the point t marked by the arrow. Below the point t the isentrope and therefore the density profile remain unchanged.

Let us consider model 3. This model is based on a linear approximation of the pressure on the isentrope

$$p = p_a(\rho - \rho_e)/(\rho_{ss} - \rho_e). \tag{2.3}$$

In this case ρ undergoes a jump rather than dropping off gradually to zero, just as in polytropic models with finite γ. The pressure p (2.3) vanishes at a finite density ρ. The function (2.3), referring to the isentrope 3, is shown in Fig. 1 (broken curve; $p_a = 31.3$ GPa, $\rho_e = 1.56$ g/cm^3).

The profile ρ following from the isentrope (2.3) has the form

$$\rho(X) = \rho_{ss} \exp\left(-\frac{\rho_{ss} - \rho_e}{p_a} X \right), \tag{2.4}$$

for $X_a < X < X_v$ and $X_v - X_a = [p_a/(\rho_{ss} - \rho_e)]\ln(\rho_{ss}/\rho_e)$. For $X > X_v$ the density $\rho(X)$ is zero. A density jump occurs at the point X_v.

A comparison of the model 3 with the numerical calculations is shown in Fig. 6. The solid curve was calculated using Eq. (2.4) and the broken curve shows the numerical calculation of the isentrope 3.

Let us now examine the structure of the tail, i.e., the distribution of variables in the vaporized matter. Let the exponent γ be small ($\gamma \ll 1$, $\gamma \approx 0$, see Figs. 3 and 4) to the left of the point e (see Fig. 1). Let us consider two layers. The lower layer is an adiabatic with a large value of γ or an isentrope (2.3) linear in ρ (we shall require its segment $a-t$ that does not reach $p = 0$). We choose as the constant γ in the bottom layer an average value of the function $\gamma(\rho)$ over the single-phase region, and we choose for ρ_t a density somewhat greater than ρ_e. The upper distribution matches the lower distribution continuously in the variables ρ and p ($p_t \ll p_a$, $\rho_t \simeq \rho_a$).

Let us assume that the upper layer is described by an isentrope with exponent $\gamma = 0$. Then the distributions in this layer are

$$\rho(X) = \frac{\rho_t}{1 + X/c_{t2}^2}, \quad p \approx p_t, \quad c = \sqrt{X + c_{t2}^2}, \tag{2.5}$$

where c_{t2} is the sound velocity at the matching point t on the upper-layer side. We have $c_{t2} \ll c_{t1}$, where c_{t1} is the sound velocity at the point t on the lower-layer side. The smallness of c_{t2}^2 is responsible for the rapid decrease of ρ in the tail (2.5), since the height of the homogeneous atmosphere $(\Delta X)_2 = c_{t2}^2$ in it is small compared with the analogous scale in the lower layer.

2.4. *Remark concerning multidimensional hydrodynamic stability.* The acceleration of the cold part of the foil by the pressure p_a isunstable with respect to the Rayleigh–Taylor instability. The ablation surface is unstable. The entropy distribution s inside the foil is uniform and therefore local quasihydrostatic equilibrium is neutrally stable. For low plasma density in the corona we have $\Gamma = \sqrt{gk}$, where Γ is the growth rate of the Rayleigh-Taylor instability, g is the acceleration, and k is the wave number. This dependence, which refers to the linear stage, is universal: It does not depend on the equation of state or the profile ρ.

At the nonlinear stage the instantaneous state of mixing is characterizedby a horizontal scale $\langle \lambda \rangle$—the average bubble size. The profile of ρ should affect the dynamics of thebubble motion if $\langle \lambda \rangle \sim H_{\text{eff}} = \sigma/\langle \rho \rangle$, where σ is the surface density and ρ the average density. The presence of two phases means that the foil is bounded by two sharp jumps in ρ at the points a and e (see Fig. 2). For $\rho_e \simeq \rho_a$ this profile is similar to that of a uniform ($\rho_e = \rho_a$) incompressible ($\gamma = \infty$) liquid, since in our case the exponent γ is quite large inside the foil. A detailed experimental investigation of the nonlinear Rayleigh–Taylor instability in a layer of incompressible liquid has been made in Ref.15. The lifetime t_{life} of the foil before it undergoes mixing and perforation is finite. It is known that the growth rate of the disturbances is such that the foil traverses several tens of H_{eff} before being perforated. The foil can undergo many acoustic oscillations over the time t_{life}. Therefore the quasihydrostatic description makes senseas a convenient zeroth approximation despite the Rayleigh–Taylor instability.

3. RAREFACTION WAVE

3.1. *Numerical integration.* Let us consider the expansion of an initially uniform layer into a vacuum. The expansion is described by a centered[8–10] rarefaction wave. The equations of gas dynamicsin this case are

$$\rho_t + (\rho u)_x = 0, \quad u_t + u u_x + p_x/\rho = 0, \quad p = p(\rho). \tag{3.1}$$

In Riemann variables the system (3.1) becomes

$$P_t + (u + c)P_x = 0, \quad M_t + (u - c)M_x = 0, \tag{3.2}$$

where $c = \sqrt{(\partial p/\partial \rho)_s}$, $P = u + I$, $M = u - I$, $I = \int_{\rho_a}^{\rho} c(\rho) \times d\rho/\rho$, and a reference level for ρ, which is convenient for what follows, is chosen immediately in the integral I. In contrast to the integral (2.1) the integral I (here the acceleration of gravity $g = 0$) contains the first power of $c(\rho)$.

A centered rarefaction wave is self-similar. The self-similar variable is $\xi = x/t$, where t is measured from the mo-

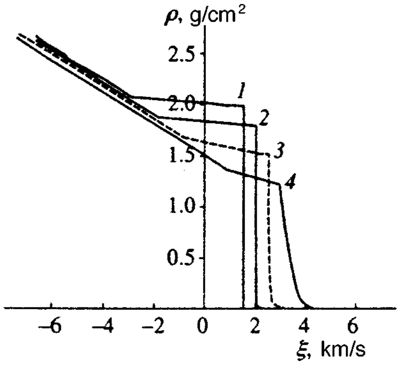

FIG. 7. Profiles of a centered rarefaction wave.

ment expansion starts, the point $x=0$ coincides with the position of the edge of the layer at $t=0$, and the x axis is directed toward the vacuum.

In the self-similar case the system (3.2) becomes

$$(-\xi+u+c)P'_{\xi}=0, \quad (-\xi+u-c)M'_{\xi}=0. \tag{3.3}$$

For the chosen direction of the x axis, the waves $u-c$ travel into the foil. In this case the function $P(\xi)$ is constant ($P'_{\xi}\equiv 0$), and the function $M(\xi)$ is nontrivial ($M'_{\xi}\neq 0$). Therefore from the second equation of Eqs. (3.3) follows $-\xi+u-c=0$. From the condition at the edge of a centered rarefaction wave and by virtue of the chosen calibration of the integral I we have $P(\xi)\equiv 0$, since $u(\xi)+I(\xi)=u(-c_a)+I(-c_a)=0$. Therefore $u(\xi)=-I(\xi)$.

On the isentrope we have $\rho=\rho(c)$. Therefore $I=I(\rho)=I[\rho(c)]=I(c)$. Correspondingly, the relation $-\xi+u-c=0$ becomes $\xi=u-c=-I(c)-c=-I[c(\rho)]-c(\rho)=-I(\rho)-c(\rho)$. The latter equation determines implicitly the self-similar profiles of the thermodynamic variables and the velocity u of the centered rarefaction wave. The computational results obtained under this program are presented in Fig. 7. The waves *1–4* refer to the isentropes *1–4* in Fig. 1.

3.2. *Analytic solution: power-law approximation.* For a power-law isentrope the systems (3.1) and (3.2) assume the simple form

$$Nc_t+cu_x+Nc_xu=0, \quad u_t+uu_x+Ncc_x=0, \tag{3.4}$$

where $N=2/(\gamma-1)$. In the self-similar case the system (3.4) becomes

$$cu'/N+(u-\xi)c'=0, \quad (u-\xi)u'+Ncc'=0, \tag{3.5}$$

where $u'\equiv u'_{\xi}$. Inhomogeneous solutions are obtained if the determinant of the system (3.5) vanishes. This condition has the form $(u-\xi)^2=c^2$ or $\xi=u\pm c$. In our case $\xi=u-c$. Since the equations of the acoustic characteristics are $dx/dt=u\pm c$, this means that the characteristics of the family $u-c$ are rectilinear, and inside the centered rarefaction wave they coincide with the lines $\xi=\text{const}$. Substituting $\xi=u-c$ into the system (3.5), we find the general solution[8–10]

$$u=\frac{N\xi}{N+1}+A, \quad c=-\frac{\xi}{N+1}+A, \tag{3.6}$$

where A is an arbitrary constant determined by the boundary conditions.

Let us examine a centered rarefaction wave in the case of a two-step isentrope. On each step the isentrope is a power-law with exponents γ_1 and γ_3, respectively. The two-step isentrope is continuous at the point e where the steps join. At the point e the derivative of the isentrope is discontinuous.

The rarefaction wave consists of three sections. The first one is acentered rarefaction wave, adjoining a uniform rest zone, where undisturbed matter is present. The conditions of matching with this zone determine the constant A in Eq. (3.6). On the first section the solution (3.6) is

$$u_1(\xi)=\frac{N_1(\xi+c_a)}{N_1+1}, \quad c_1(\xi)=\frac{-\xi+N_1c_a}{N_1+1}, \tag{3.7}$$

where the index 1 refers to the first section and $c_a\equiv c_{sa}$ is the sound velocity in the uniform zone.

The section 1 of the centered rarefaction wave lies in the segment $-c_a<\xi<\xi_{1e}$. Along the characteristic $\xi=-c_a$ this section adjoins the undisturbed foil ($u_1(-c_a)=0$, $c_1(-c_a)=c_a$).

Let us examine the other boundary of section 1, located at the point ξ_{1e}. To each of the two steps of the isentrope there is associated a separate centered rarefaction wave of the form (3.6). The first stepcovers the segment $a-e$ of the isentrope (see Fig. 1). The second step covers the part of the isentrope which passes along the two-phase region from $\rho=\rho_e$ to $\rho=0$. The sign of the kink in the isentrope at the point e is such that the sound velocity decreases abruptly at the transition through the point e from right to left (see Fig. 1), i.e., from the single- to the two-phase region. The corresponding inequality is

$$c_{3e}<c_{1e},$$
$$c_{3e}=\sqrt{\left(\frac{\partial p(\rho_e-0)}{\partial\rho}\right)_s}, \quad c_{1e}=\sqrt{\left(\frac{\partial p(\rho_e+0)}{\partial\rho}\right)_s}, \tag{3.8}$$

where c_{1e} is the sound velocity at the end of section 1 of the centered rarefaction wave at the point ξ_{1e} and c_{3e} is the sound velocity at the start of section 3 at the point ξ_{3e}. Since the characteristics $u-c$ run, relative to Lagrangian particles, in the direction of negative values of ξ, the start of the centered rarefaction wave lies at lower values of ξ than the end of the wave.

In the rarefaction wave under consideration we have $\xi=u-c$. Therefore

$$\xi_{1e}=u_{1e}-c_{1e}, \quad \xi_{3e}=u_{3e}-c_{3e}, \tag{3.9}$$

where $u_{1e}\equiv u_1(\xi_{1e})$, $c_{1e}\equiv c_1(\xi_{1e})$, $u_{3e}\equiv u_3(\xi_{3e})$, and $c_{3e}\equiv c_3(\xi_{3e})$. A region of uniform flow—a plateau—can exist between the end of section 1 and the start of section 3. Then $u_{1e}=u_{3e}=u_2$. It is easy to see that from this equality, Eqs. (3.9), and the inequality (3.8) follows $\xi_{3e}>\xi_{1e}$. This means that the sections 1 and 3 of the centered rarefaction wave cannot be matched directly with one another and that a region of uniform flow does indeed lie in the segment $\xi_{1e}<\xi<\xi_{3e}$. Indeed, p is continuous at the point *e*. Therefore u is also continuous at this point (∇p and the acceleration of the Lagrangian particles are finite). The sound velocity decreases abruptly. Hence the sound wave ξ_{3e} lags behind the wave ξ_{1e}. We shall consider the correspondence between the points $\{\rho,p\}$ on the isentrope andthe points $\{\xi,\rho\}$ in the rarefaction wave. We can see that the point e in the $\{\rho,p\}$ plane transforms into a finite segment in the $\{\xi,\rho\}$ plane. The left- and right-hand neighborhoods of the point e from $\{\rho,p\}$ form the ends of the plateau on $\{\xi,\rho\}$.

Let us now estimate the position of the plateau on the basis of the parameters of the tabulated isentrope. Calculating the second of the functions (3.7) at the point ξ_{1e}, we find

$$\xi_{1e}=N_1c_a-(N_1+1)c_{1e}, \tag{3.10}$$

where the velocities c_a and c_{1e} are found from the tabulated isentrope. Now we write the first of the functions (3.7) at the point ξ_{1e}. Substituting the expression (3.10) into it we find the mass velocity in the region of uniform flow

$$u_{1e}=u_2=u_{3e}=N_1(c_a-c_{1e}). \tag{3.11}$$

It is obvious from what we have said above that the parts of the wave are the section 1, the uniform flow (plateau, section 2), and the section 3.

Let us now find the width of the plateau. We have $\xi_{3e}=u_2-c_{3e}$ [see Eqs. (3.9) and (3.11)]. The value of c_{3e} is found from the isentrope. The velocity c_{3e} is very low, so $\xi_{3e}\approx u_2$. From $\xi=u-c$ and Eq. (3.11) we find the width of the plateau as

$$\xi_{3e}-\xi_{1e}=c_{1e}-c_{3e}\approx c_{1e}.$$

Let us now consider the section 3. This rarefaction wave is unusual, because the exponent $\gamma=(\partial\ln p/\partial\ln\rho)_s$ is small here compared to 1. In ordinary situations we have $\gamma>1$ (a power-law rarefaction wave bounded by a vacuum edge) or $\gamma=1$ (isothermal exponential rarefaction wave continuing to infinity).

Let us determine the constant A in Eqs. (3.6) for a centered rarefaction wave on the section 3. For this we use the point ξ_{3e} and the sound velocity c_{3e} there. As a result we obtain

$$u_3(\xi)=\frac{N_3\xi+N_1(c_a-c_{1e})+N_3c_{3e}}{N_3+1}\approx 2\xi-N_1(c_a-c_{1e}), \tag{3.12}$$

$$c_3(\xi)=\frac{-\xi+N_1(c_a-c_{1e})+N_3c_{3e}}{N_3+1}\approx \xi-N_1(c_a-c_{1e}). \tag{3.13}$$

In Eqs. (3.12) and (3.13) approximate expressions are obtained for $c_{3e}\approx 0$ and $N_3\approx -2$.

From Eq. (3.13) and the isentrope follows an expression for ρ

$$\rho_3(\xi)=\rho_2\left[\frac{-\xi+N_1(c_a-c_{1e})+N_3c_{3e}}{(N_3+1)\quad c_{3e}}\right]^{N_3}$$
$$=\frac{\rho_e\{[-(N_3+1)]c_{3e}\}^{2/(1-\gamma_3)}}{\xi-\xi_{3e}+[-(N_3+1)]c_{3e}}\approx\frac{\rho_2c_{3e}^2}{(\xi-\xi_{3e}+c_{3e})^2}, \tag{3.14}$$

where ρ_2 is the density on the section 2, i.e., on the plateau.

For low sound velocities c_{3e} in the two-phase region the density (3.14) in the section 3 decreases rapidly with ξ increasing rightward from the left-hand boundary ξ_{3e} of the section (see Fig. 7). This section resembles the tail of the atmosphere studied in Sec. 2. Its width is $(\Delta\xi)_3=c_{3e}$.

As we can see, the relative width of the tail in the rarefaction wave (i.e., the width of the tail scaled to the total width of the wave) is of order c_{3e}/c_a. It is much greater than the relative thickness of the atmospheric tail, which in order of magnitude is $(c_{3e}/c_a)^2$. However, ρ in the tail decreases with increasing x more rapidly in the rarefaction wave than in hydrostatics. Specifically, $\rho\propto 1/x$ in the atmosphere, whereas $\rho\propto 1/x^2$ in the rarefaction wave [compare Eqs. (2.5) and (3.14)].

3.3. *Analytic solution with a linear approximation.* We shallconsider the model with a linear approximation of the isentrope in the single-phase region. Let the pressure in this region be

$$p(\rho)=(p_a-p_e)\frac{\rho-\rho_e}{\rho_a-\rho_e}+p_e. \tag{3.15}$$

Let $c_{ae}^2=(p_a-p_e)/(\rho_a-\rho_e)$. It is easy to show that the profiles of the variables in this case are given by the formulas for an isothermal rarefaction wave. They have the form

$$u_1(\xi)=\xi+c_{ae}\quad\text{and}\quad \rho_1(\xi)=\rho_a\exp[-(\xi+c_{ae})/c_{ae}].$$

The rarefaction wave once again consists of three sections: an isothermal wave, a plateau, and a tail. The coordinates of the boundaries between them are $\xi_{1e}=-c_{ae}+c_{ae}\ln(\rho_a/\rho_e)$ and $\xi_{3e}=c_{ae}\ln(\rho_a/\rho_e)-c_{3e}$, where once again the indices 1, 2, and 3 denote the numbers of the sections, a is the initial point, and e is a point on the curve of boiling. The indices $1e$ and $3e$ refer to neighborhoods of the point e on the side of the single- and two-phase regions, respectively.

The problem of boiling up of a liquid in rarefaction waves has a definite bearing on the questions examined in Sec. 3 (Ref. 16). Itwas studied in connection with the problems of safety of nuclear power production. We are talking about accidental rupture of a pipeline carrying coolant at high pressure. We underscore that physically the coolant and laser-heatingcases are very different. In the former p_a/p_c is only 10% greater than1,[16] whereas in the latter $p_a/p_c\sim 10^2$, (see Sec. 2.1). In thefirst case $\rho_e\approx\rho_a$. Therefore u_2 is small compared to c_a (the rarefaction wave in section 1 is short and the acceleration of matter is small). As a result, sound runs rapidly through the tube, while matter exhausts relatively slowly. In the second case matter on the plateau moves much more rapidly ($u_2\sim c_a$).

In the rupture problem information about the process is obtained from pressure sensors. In this problem p varies abruptly on the short section 1. Conversely, a sharp decrease of ρ on the section 3 under our conditions (see Fig. 7) is smoothed by the piezoelectric element, since p in the two-phase region approximately equals p on the plateau. At the same time, for applications in optical diagnostics of laser ablation by ultrashort laser pulses[17,18] jumps in ρ (great values of $|\nabla\rho|$), which reflect light, are important. Such jumps in ρ have been observed in Refs. 17 and18. In addition, in

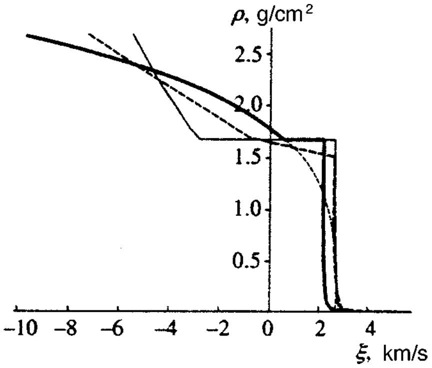

FIG. 8. Structural features of a two-phase wave: plateau and tail. The parameters are constant in the plateau region. Thermodynamically this entire portion of matter is in a state corresponding to the point of the kink in the isentrope on the vaporization curve. Asharp decrease of the density occurs in the tail.

the rupture problem the section $x<0$ is usually studied (see Fig. 7), since the rupture point is located at $x=0$, while for optical measurements the entire wave is important.

3.4. *Structure of the wave.* In Fig. 8 the first and second analytic solutions are compared with the numerical solution. The heavy dashed curve shows the numerical solution in the case of the isentrope 3. In Fig. 7 it is distinguished by dashes. The heavy solid curve shows the first solution corresponding to a power-law approximation of the single-phase isentrope. The light solid curve is a linear approximation (3.15) of the same part of the isentrope (second solution). The light small dashes show the form that the centered rarefaction wave would have if the isentrope were described by a power law everywhere.

The first solution was constructed as follows. The points $(\ln p_a, \ln \rho_a)$ and $(\ln p_e, \ln \rho_e)$ were taken from the isentrope (see Fig. 1). The exponent γ_1 was found from the slope of the straight line connecting them. The velocities $c_a^2 = \gamma_1 p_a/\rho_a$ and $c_e^2 = \gamma_1 p_e/\rho_e$ were calculated. This method overestimates c_a and underestimates c_e, and in consequence it overestimates the width of section 1. Looking ahead, we note in contrast that in the method referring to the second solution c_a is underestimated, c_e is overestimated, and therefore the width of section 1 is underestimated. The fact that the linear approximation underestimates c_a and overestimates c_e is obvious from a comparison of the linear (dashed curve $a-e$)and the tabulated (solid curve $a-e$) isentropes presented in Fig.1. We note that the width of the plateau in the first solution is less and in the second solution greater than in the numerical calculation.

We also note that the plateau in the numerical calculations has a slope (see Figs. 7 and 8). This is due to the computational errors.

Further, in constructing the first solution ξ_{1e} and $\rho_1(\xi)$ were found, the end ξ_{3e} of the plateau was found, and γ_3 and c_{3e} were found from the isentrope, and the function $\rho_3(\xi)$ was found. To construct the second solution p_a, p_e, ρ_a, and ρ_e were chosen and c_{ae}, ξ_{1e}, and ξ_{3e} and the functions $\rho_1(\xi)$ and $\rho_1(\xi)$ were calculated.

Figure 8 compares waves with $s=3.76$ J/g·K (isentrope *3,* see Fig. 1). The values of the parameters are $\gamma_1=8.4$, $c_a=9.8$ km/s, $c_e=1.7$ km/s, $\xi_{1e}=0.487$ km/s, $\xi_{3e}=2.14$ km/s, $\gamma_3=0.195$, and $c_{3e}=0.052$ km/s. In the second solution the main parameters are $c_{ae}=5.5$ km/s, $u_{1e}=u_2=u_{3e}=2.6$ km/s, $\xi_{1e}=-2.9$ km/s, and $\xi_{3e}=2.6$ km/s.

A comparison shows that the tails of the analytic and numerical solutions agree well with one another (the agreement between the heavy dashed and light solid lines in Fig. 8).

4. SIMPLE WAVE AND INTERFERENCE OF TWO SIMPLE WAVES: FORMATION OF A GAP INSIDE THE VAPORIZED MATTER

An atmosphere, bounded by a density jump, and a simple wave, consisting of the standard centered rarefaction wave and a plateau, also bounded by the jump in ρ, was presented in Secs. 2 and 3. We shall now study the expansion of a layer heated uniformly by an ultrashort laser pulse ($\tau \ll H/c_a$). Simple waves propagate toward one another from the two boundaries of the layer (foil) and meet at time $H/2c_a$. If the foil is deposited on a rigid substrate, then at time $t_r=H/c_a$ the simple waveis reflected from this substrate. For definiteness, we shall study this case.

In the case of an isentrope without a phase transition and with an everywhere continuous sound velocity, for $t>t_r$ the rarefaction wave is divided into two parts by the reflected sound wave with instantaneous coordinate $x_r(t)$. The first part forms a centered rarefaction wave. This is the region $x_r(r)<x<x_v(t)$ with one simple wave. In the second part $(0<x<x_r(t))$ two simple waves—incident and reflected—interact. As $t\to\infty$ the mass of the first region $m_1\to 0$. Integrating the reflected characteristic $(dx_r/dt=u+c)$ with a power-law approximation of the isentrope gives $x_r=H+[Nc_a-(Nc_a+H/t_r)(t_r/t)^{2/(N+1)}]t$. Here x is measured from the rigid wall. With this reference the film edge at $t=0$ is located at the point $x=H=c_a t_r$.

We are especially interested in the interaction of the reflected wave with the plateau region. The wave $x_r(t)$ reaches it at the moment of freezing t_f. The reason why it is called this will be explained below. For a power-law approximation $t_f/t_r=[(Nc_a+H/t_r)/(Nc_a-\xi_{1e})]^{(N+1)/2}$. A difference calculation of the isentropes *1–4* presented in Fig. 1 gives $t_f/t_r=\exp[\int d\xi/2c(\xi)]=1.44$ (1), 1.63 (2), 1.97 (3), and 2.63 (4), where the integral is taken from $-c_a$ to ξ_{1e} (the number of the isentrope is indicated in parentheses). As s increases, the difference $\xi_{1e}-\xi_a$ and the time t_f/t_r increase.

The rate of stretching, characterized by the derivative $\partial u/\partial x$ or the total derivative $(-D\ln\rho/Dt)$ $(D/Dt=\partial_t+u\partial_x)$, is greater in the two-wave region than in the one-wave region. In a simple wave with a phase transition (see Sec. 3) $u_x'>0$ on section 1, and $u_x'\equiv 0$ on section 2. This is the frozen region (plateau), where matter coasts as a whole and the thermodynamic quantities are constant as a function of x and t.

Let us see what happens for $t>t_f$. The unusual nature of the the interaction of the reflected sound with the plateau region should be underscored. Behind the reflected wave $(0<x<x_r(t))$ the density $\rho(x,t)$ decreases with time, while

on the plateau it is ρ_e and does not change. At $t=t_f$ $\rho(x,t) \approx \rho_e$ in the reflected wave ($\rho(x_r,t_f)=\rho_e$). We shall call the region $0<x<x_{1e}$ a gap, where x_{1e} is the instantaneous left-hand coordinate of the plateau. By the time $t\approx t_f$ a phase transition has already occurred in allmatter in the gap. After the transition $|\nabla p|$ decreases sharply. As a result, motion in the gap becomes frozen—a coasting regime is established. This means that the dependence on t in the velocity $u(m,t)$ written as a function of the Lagrangian coordinate m vanishes.

At the moment freezing occurs $u'_m\equiv 0$ on the plateau and $u'_m>0$ in the gap. Therefore for $t>t_f$ the motion of the boundary between the plateau and the gap along the Lagrangian coordinate m stops (accumulation of matter by the plateau region stops; the maximum ratio of the plateau mass to the total mass is $(\rho_e/\rho_a)(\Delta\xi/c_a)(t_f/t_r)$, $\Delta\xi=\xi_{3e}-\xi_{1e}$). The density ρ in the gap continues to decrease, since here $u'_m>0$ and is frozen. As a result, ρ in the gap becomes smaller than the density ρ_e in the plateau region. Thus the distribution of ρ over x or m becomes nonmonotonic. On the plateau $c=c_{1e}$ holds for compression waves and $c=c_{3e}$ for rarefaction waves. In the gap for $t>t_f$ the motion is hypersonic. It is evident that the boundary between the plateau and the gap becomes a second (with respect to the boundary x_{3e}) boundary where the density changes sharply. For $t\gg t_f$ the density ρ in the gap decreases $\approx(t_f/t)\rho_e$. The plateau decays slowly, since the expansion velocities are low ($\sim c_{3e}$).

5. CONCLUSIONS

A comparative analysis of the structure of ablation-accelerated foils was performed including and neglecting phase transformations. A phase transition leads to the appearance of important physical features. Whenthe unloading isentrope passes through the two-phase region, a vaporization boundary forms and ρ above this boundary decreases sharply (atmospheric tail). For rarefaction waves the two-phase nature of the medium has the effect that a quite extended region of uniform flow (a plateau), ending with a sharp dropoff of ρ (the tail of the rarefaction wave) arises. In the zone of the sharp dropoff of density the matter is in a two-phase state (vapor bubbles in a liquid, droplet condensate in vapor). The phase composition of the vapor–liquid changes away from the vaporization boundary (from the outer boundary of the plateau). Expansion of a layer of finite thickness is accompanied by the formation of two plateaus separated by a gap. The density of matter in the plateau remains constant for a long time.

The structures of the tails are different in the atmospheric and wave cases. They differ by the width and asymptotic behavior of $\rho(x)$ (see end of Sec. 3.2). An analytic description of the profiles of the tails was obtained. It agrees well with numerical calculations.

Ordinarily, the fraction of matter behind the vaporization boundary is small compared to 1 under the experimental conditions considered. This is due to the smallness of the ratio of the pressure p_e on the binodal to the ablation pressure p_a. We note that a direct numerical simulation is impeded because after a phase transition the sound velocity decreases to extremely low values.

We thank A. M. Oparin and V. V. Zhakhovskiĭ for helpful discussions. Financial support for this work was provided by the Russian Fund for Fundamental Research (Project 98-02-17441-a) and the program for support of leading science schools (Project 96-15-96448).

[*)] E-mail: nail@landau.ac.ru

[†)] E-mail: anisimov@itp.ac.ru

[‡)] E-Mail: i2042803@aix3.thp.nat.tu-bs.de

[1)] It should be noted that the accuracy of the semiempirical equations of state in the two-phase and near-critical regions is comparatively low for most metals. Metals with low critical temperatures (mercury and alkali metals), for which direct experimental data are available, are exceptions. In this connection, the numerical results obtained with different equations of state can differ. However, this does not change the qualitative picture of the phenomena studied. Laser and beam experiments open up interesting possibilities for refining the equations of state in regions where the more conventional methods are of limited applicability.

[1] A. M. Prokhorov, S. I. Anisimov, and P. P. Pashinin, Usp.Fiz. Nauk **119**, 401 (1976) [Sov. Phys. Usp. **19**, 547 (1976)].

[2] S. I. Anisimov, A. M. Prokhorov, and V. E. Fortov, Usp. Fiz. Nauk **142**, 395 (1984) [Sov. Phys. Usp. **27**, 181 (1984)].

[3] J. J. Duderstadt and G. A. Moses, *Inertial Confinement Fusion* (Wiley, New York, 1982) [Russian translation, Énergoatomizdat, Moscow, 1985].

[4] V. E. Fortov and I. T. Iakubov, *Physics of Nonideal Plasma*, Hemisphere, New York, 1990 [Russian original, Énergoatomizdat, Moscow, 1994].

[5] K. Tanaka (1998), private communication.

[6] R. Bock, I. Hofmann, and R. Arnold, Nucl. Sci. Appl. **2**, 97 (1984).

[7] R. C. Arnold and J. Meyer-ter-Vehn, Rep. Prog. Phys. **50**(5), 559 (1987).

[8] R. Courant and K. O. Friedrichs, *Supersonic Flow and Shock Waves* (Interscience, New York, 1948] [Russian translation, Inostr. Lit., Moscow, 1950].

[9] Ya. B. Zel'dovich and Yu. P. Raĭzer, *Physics of Shock Waves and High-Termperature Hydrodynamic Phenomena* Vols. 1 and 2, translation of 1st Russian edition (Academic Press, New York, 1966, 1967) [Russian original, 2nd edition, Nauka, Moscow, 1966].

[10] L. D. Landau and E. M. Lifshitz, *Fluid Mechanics* (Pergamon Press, New York) [Russian original, Nauka, Moscow, 1986].

[11] A. V. Bushman and V. E. Fortov, Usp. Fiz. Nauk **140**, 177 (1983) [Sov. Phys. Usp. **26**, 465 (1983)].

[12] V. A. Agureĭkin, S. I. Anisimov, A. V. Bushman, G. I. Kanel, V. P. Karyagin, A. B. Konstantinov, B. P. Kryukov, V. F. Minin, S. V. Razorenov, R. Z. Sagdeev, S. G. Sugak, and V. E. Fortov, Teplofiz. Vys. Temp. **22**, 964 (1984) [High Temp. **22**, 761 (1984)].

[13] R. Yamamoto and K. Nakanishi, Phys. Rev. B **49**(21), 14958 (1994).

[14] A. Onuki, J. Phys.: Condens. Matter **9**, 6119 (1997).

[15] M. D. Kamchibekov, E. E. Meshkov, N. V. Nevmerzhitsky, and E. A. Sotskov, in *Proceedings of the 6th International Workshop on ''The Physics of Compressible Turbulent Mixing,''* edited by G. Jourdan and L. Houas, Marseille, France, 1997, p. 238.

[16] R. I. Nigmatullin, *Dynamics of Multiphase Media* (Nauka, Moscow, 1987), Part II, p. 136.

[17] D. von der Linde, K. Sokolowski-Tinten, and J. Bialkowski, Appl. Surf. Sci. **109/110**, 1 (1996).

[18] K. Sokolowski-Tinten, J. Bialkowski, A. Cavalleri, D. von der Linde, A. Oparin, J. Meyer-ter-Vehn, and S. I. Anisimov, Phys. Rev. Lett. **81**(1), 224 (1998).

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Multiquantum triple-pulse spin-echo signals from quadrupole nuclei in magnetically ordered substances

V. N. Berzhanskiĭ,*) A. I. Gorbovanov, and S. N. Polulyakh

Simferopol State University, 330007 Simferopol, Ukraine

(Submitted 25 June 1998)
Zh. Éksp. Teor. Fiz. **115**, 2106–2112 (June 1999)

We analyze theoretically the formation of NMR pulse responses from a quadrupole spin system in which the inhomogeneous broadening of a spectral line is due to both magnetic and electric quadrupole interactions. We derive formulas for the moments of formation of multiquantum echo signals in the case of three exciting pulses. For the first time we detected in experiments multiquantum spin-echo signals from copper nuclei in ferromagnetic copper sulfochromite in the cases of double-pulse and triple-pulse excitations. We find that there is good agreement between the calculated and experimentally observed moments of echo signal formation. © *1999 American Institute of Physics.* [S1063-7761(99)01506-1]

1. INTRODUCTION

The magnetic hyperfine interactions between the electron and nuclear spin subsystems in magnetically ordered substances lead to the emergence of local magnetic fields on the nuclei, fields that determine the NMR frequencies.[1] More than that, NMR in magnetically ordered substances is characterized by natural inhomogeneous broadening of a spectral line, due to which the main method of experimental NMR spectroscopy is the spin-echo method.[1–3] In the case of magnetic resonance of nuclei with spins $I>1/2$, the electric quadrupole interactions lead to the emergence of additional spectral lines, to quadrupole satellites, and to formation of multiquantum echo signals.[2,4]

The moments of formation of multiquantum echo signals from quadrupole nuclei in magnetically ordered substances are well known for the case of two exciting pulses.[5–8] On the other hand, triple-pulse trains are widely used to excite echo signals in inhomogeneously broadened spin systems.[2,3] In particular, they are used in studies of magnetic relaxation processes.

The aim of the present work is to analyze theoretically the moments of formation of multiquantum responses of a quadrupole spin system under triple-pulse excitation. To verify the theoretical results in experiments, we used multiquantum echo signals from ^{63}Cu nuclei in the ferromagnet $CuCr_2S_4:Sb$ at $T=77$ K.

2. THEORY

We analyze the moments of formation of triple-pulse echo signals theoretically by using the approach developed for the case of two exciting pulses.[4–7] We begin by representing an inhomogeneously broadened spectral line in the form of a set of isochromatic spin groups. We write the Hamiltonian of an isochromatic spin group in a rotating reference frame as follows:

$$H=-\hbar\Delta\omega I_z+\hbar\omega_q\left(I_z^2-\frac{I(I+1)}{3}\right), \tag{1}$$

where I_z is the operator of projection of spin I on the quantization axis z. Each isochromatic spin group is characterized by its own detuning $\Delta\omega$ and quadrupole splitting ω_q of the NMR spectrum.

During the action of the exciting pulses, the Hamiltonian of the spin system, H_1, consists of the Hamiltonian (1) and the term describing the interaction between the spin system and the alternating magnetic field:

$$H_1=H+\hbar\omega_1 I_x, \tag{2}$$

where ω_1 is the amplitude of the alternating magnetic field expressed in units of frequency.

Using the method of the density-matrix operator,[2,4–7] we arrive at an expression for the transverse component of the nuclear magnetization, $M_+=M_x+iM_y$, at time t:

$$M_+(t)=\sum_{\substack{m,m_1,m_2,\\ m_3,m_4}} V\left\{\exp\left[-\frac{it}{\hbar}(E_m-E_{m+1})-\frac{i\tau_{23}}{\hbar}(E_{m_1}-E_{m_4})\right]\exp\left[-\frac{i\tau_{12}}{\hbar}(E_{m_2}-E_{m_3})\right]\right\}. \tag{3}$$

Here time t is measured from the moment when the third exciting pulse ends, τ_{12} is the time interval between the first and second exciting pulses, τ_{23} is the time interval between the second and third exciting pulses (Fig. 1), and $E_{m'}=\langle m'|H|m'\rangle$ is the eigenvalue of the Hamiltonian (1) in the state with the magnetic quantum number m'. Each term in (3) describes the echo signal in the case where the exponent vanishes at all values of $\Delta\omega$ and ω_q. This is possible for various values of the magnetic quantum numbers m, m_1, m_2, m_3, and m_4. The amplitude V of the corresponding echo signal is

$$V=\sqrt{I(I+1)-m(m+1)}\langle m|R_3|m_1\rangle\langle m_1|R_2|m_2\rangle \times\langle m_2|R_1I_zR_1^{-1}|m_3\rangle\langle m_3|R_2^{-1}|m_4\rangle\langle m_4|R_3^{-1}|m+1\rangle, \tag{4}$$

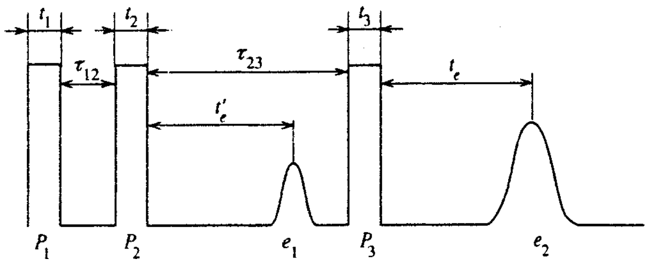

FIG. 1. Time diagram of formation of a triple-pulse echo signal: t_j is the length of the jth exciting pulse; P_1, P_2, and P_3 are the exciting pulses; e_1 is double-pulse echo; and e_2 is triple-pulse echo.

where $R_j^{\pm 1} = \exp\{\mp it_j H_1/\hbar\}$ are the operators describing the action of the jth exciting pulse, with t_j the pulse length.

Assuming that the exciting pulses are much shorter than the time intervals τ_{12} and τ_{23} and using the fact that the exponent in (3) must vanish at the moment of formation of the echo signal, we find that

$$t_e = \tau_{12}\frac{-\Delta\omega(m_2-m_3)+\omega_q(m_2^2-m_3^2)}{-\Delta\omega+\omega_q(2m+1)} + \tau_{23}\frac{-\Delta\omega(m_1-m_4)+\omega_q(m_1^2-m_4^2)}{-\Delta\omega+\omega_q(2m+1)}. \tag{5}$$

This expression (5) describes the moment of echo signal formation when t_e is independent of $\Delta\omega$ and ω_q. By multiquantum echo signals we mean such responses of the spin system for which at least one condition, $|m_2-m_3|>1$ or $|m_1-m_4|>1$, is met. If the first condition is met, multiquantum coherence forms in the time interval τ_{12}. If the second condition is met, multiquantum coherence forms in the time interval τ_{23}. Ordinary triple-pulse echo signals form if magnetic quantum numbers obey the condition

$$m = m_1 = m_2 = m_3 - 1 = m_4 - 1. \tag{6}$$

3. EXPERIMENTAL RESULTS AND A DISCUSSION

Experiments were done with a pulsed incoherent NMR spectrometer. The NMR signals were those generated by ^{63}Cu nuclei in a polycrystalline ferromagnet $CuCr_{1.98}Sb_{0.02}S_4$ cooled to $T = 77$ K. In the double-pulse response, in addition to the ordinary echo signal formed at time $t'_e = \tau$, we detected multiquantum echo formation at time $t'_e = 3\tau$, where τ is the time interval between the exciting pulses, and t'_e is measured from the moment at which the second exciting pulse ceases to act. The maximum in the echo signal amplitude at $t'_e = 3\tau$ was observed near 100.7 MHz, with the first and second pulse lengths being 5 and 3 μs, respectively, and the amplitude of the rf pulses being $U = 220 \pm 20$ V.

The ^{63}Cu nucleus is a quadrupole nucleus with spin $I = 3/2$. The moments of occurrence of triple-pulse echo signals, t_e, calculated by (5) with $I = 3/2$ are listed in Table I for the case of multiquantum signals. All echo signals can be divided into three groups. The first group consists of signals for which multiquantum coherence is formed in the time interval τ_{12} (signals 1–4 in Table I). The second group consists of echo signals for which multiphoton coherence is formed in the time interval τ_{23} (signals 5–8). The third group consists of echo signal with multiphoton coherence formed in both time intervals, τ_{12} and τ_{23} (signals 9–11). Equation (5) describes the occurrence of a multiquantum echo signal at time $t_e = 3\tau_{12}$ for any (of the four possible) values of the magnetic quantum number m_1. The moment of echo formation, $t_e = 3\tau_{23}$, is realized at all possible values of the magnetic quantum number m_2.

In our experiments we used triple-pulse trains in which at least two pulses were of the same length. The amplitudes U of different pulses belonging to the same chain were the same. The search for multiphoton echo signals for the case of triple-pulse excitation was done at the frequency corresponding to the maximum in the amplitude of the double-pulse echo, 3τ.

Ordinary triple-pulse echo signals were observed at times $t_e = \tau_{12}$, τ_{23}, $\tau_{12} - \tau_{23}$, $\tau_{23} - \tau_{12}$, and $\tau_{12} + \tau_{23}$ (Fig. 2). Moreover, all (eleven) theoretically expected multiquantum echo signals were experimentally detected in the triple-pulse response. The observation of a theoretically predicted echo signal is possible if $t_e > 0$. Equation (5) implies that it is impossible to select the values of the time intervals between the exciting pulses, τ_{12} and τ_{23}, so that all eleven calculated signals are observed simultaneously. A variety of values of τ_{12} and τ_{23} were used to observe various echo signals in experiments. The relationship between the moment of formation of the experimentally observed echo signal, t_e, and the

TABLE I. Moments of formation of triple-pulse multiquantum echo signals.

	Theory					Experiment						
No.	m_1	m_2	m_3	m_4	t_e	t_1, μs	t_2, μs	t_3, μs	τ_{12}, μs	τ_{23}, μs	t_e, μs	U, volts
1	any	3/2	$-3/2$	m_1	$3\tau_{12}$	5	3	5	50	75	150	220 ± 10
2	$-1/2$	3/2	$-3/2$	1/2	$3\tau_{12} - \tau_{23}$	5	3	5	50	10	140	220 ± 10
3	1/2	3/2	$-3/2$	$-1/2$	$\tau_{23} + 3\tau_{12}$	5	2	5	15	45	90	220 ± 10
4	1/2	$-3/2$	3/2	$-1/2$	$\tau_{23} - 3\tau_{12}$	5	3	5	10	110	80	220 ± 10
5	3/2	any	m_2	$-3/2$	$3\tau_{23}$	5	2	5	15	45	135	220 ± 10
6	3/2	$-1/2$	1/2	$-3/2$	$3\tau_{23} - \tau_{12}$	1	1.1	1	49	75	175	600 ± 10
7	3/2	1/2	$-1/2$	$-3/2$	$3\tau_{23} + \tau_{12}$	1	0.9	1	10	55	175	500 ± 10
8	$-3/2$	1/2	$-1/2$	3/2	$\tau_{12} - 3\tau_{23}$	1	1.1	1	90	10	60	600 ± 10
9	3/2	3/2	$-3/2$	$-3/2$	$3\tau_{12} + 3\tau_{23}$	5	1	5	10	55	190	220 ± 10
10	$-3/2$	3/2	$-3/2$	3/2	$3\tau_{12} - 3\tau_{23}$	5	5	3	80	35	135	220 ± 10
11	3/2	$-3/2$	3/2	$-3/2$	$3\tau_{23} - 3\tau_{12}$	5	5	5	12	85	220	220 ± 10

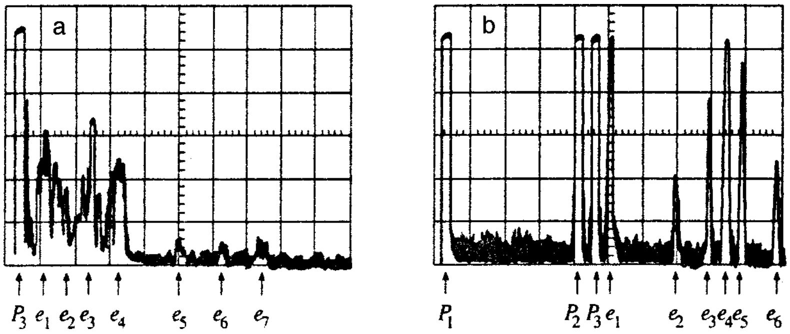

FIG. 2. Oscillograms of a triple-pulse response. The horizontal scale is 20 μs per scale division. P_1, P_2, and P_3 are the exciting pulses. The pulse lengths are (a) $t_1=5\ \mu$s, $t_2=3\ \mu$s, and $t_3=5\ \mu$s; and (b) $t_1=1\ \mu$s, $t_2=1\ \mu$s, and $t_3=1\ \mu$s. The time intervals between the pulses are (a) $\tau_{12}=43\ \mu$s and $\tau_{23}=17\ \mu$s; and (b) $\tau_{12}=80\ \mu$s and $\tau_{23}=10\ \mu$s. Ordinary triple-pulse echo signals are formed at the following moments: (a) $t_e=\tau_{23}$ (e_1), $\tau_{12}-\tau_{23}$ (e_2), τ_{12} (e_3), and $\tau_{12}+\tau_{23}$ (e_4); and (b) $t_e=\tau_{13}$ (e_1), $\tau_{12}-\tau_{23}$ (e_3), τ_{12} (e_4), and $\tau_{12}+3\tau_{23}$ (e_5). Multiquantum triple-pulse echo signals are formed at the following moments: (a) $t_e=\tau_{12}+3\tau_{23}$ (e_5), $3\tau_{12}$ (e_6), and $3\tau_{12}+\tau_{23}$ (e_7); and (b) $t_e=\tau_{12}-3\tau_{23}$ (e_2) and $\tau_{12}+3\tau_{23}$ (e_6).

time intervals between the exciting pulses was found by varying the values of the time intervals τ_{12} and τ_{23}. As an example, Table I lists the moments of formation of echo signals and the values of the time intervals τ_{12} and τ_{23} at which the corresponding echo signal was observed separately from other signals. The repetition frequency of the pulse trains amounted to 0.1 s.

For nonquadrupole spin systems with an inhomogeneously broadened spectral line there exists a vector model that provides a graphic description of the process of echo signal formation.[2,3] No such model exists, however, for multiquantum echo signals. We believe that the formation of echo signals in an inhomogeneously broadened quadrupole system (and in a nonquadrupole system) amounts to the following. The precursor pulse (or pulses) generates a transverse component of nuclear magnetization. Inhomogeneous broadening leads to decay of the transverse component of the magnetization, due to which the magnetizations of different isochromatic spin groups oscillate with different frequencies. The last exciting pulse leads to ''time reversal'' in the system, with the result that at a time t_e there forms a peak in the transverse component of magnetization, which is observed as an echo signal.

For a quadrupole spin system with $I=3/2$ the exciting pulse preceding the multiquantum coherence interval couples the states with $\Delta m=3$ and $\Delta E/\hbar=3\Delta\omega$ (see Table I). In the multiquantum coherence interval, the magnetizations of the isochromatic spin groups become dephased with a frequency equal to three times the detuning $\Delta\omega$. Such oscillations, however, are not observed in experiments, since they correspond to the frequency of a ''forbidden'' transition. The last exciting pulse couples the states with $\Delta E/\hbar=\Delta\omega$ and leads to phasing with a frequency that is three times smaller than the dephasing frequency. Such oscillations correspond to the frequency of an ''allowed'' transition ($\Delta m=1$) and are observed in experiments. Since the phasing process proceeds three times slower than the dephasing process, the moment of occurrence of a multiquantum echo signal is three times the value of the multiquantum coherence interval (see Table I).

The amplitude of multiquantum echo signals observed in experiments was found to be much smaller than the amplitude of ordinary echo signals. For signals 1–5 and 9–11 of multiquantum triple-pulse echo (Table I), the experimentally observed signal-to-noise ratio was small and varied from 1.5 to 3 (signals e_5, e_6, and e_7 in Fig. 2a). The maximum of the amplitude of these echo signals was observed when the length of the first exciting pulse and the pulse amplitude U were such that the amplitude of the double-pulse echo was at its maximum, 3τ (see Table I).

Optimum excitation of multiquantum echo signals 6–8 was achieved at a larger amplitude and smaller length of the exciting pulse (Table I). The amplitude of these echo signals was substantially larger than the amplitude of the other multiquantum triple-pulse echo signals (signals e_2 and e_6 in Fig. 2b). For the case of three-quantum double-pulse echo in a quadrupole spin system with $I=3/2$ it is known[5,8] that optical excitation of the signal is achieved when the first pulse is twice as long as the second. For signals 6–8 of multiquantum echo, the multiquantum coherence state is prepared by the first and second exciting pulses, whose total length is approximately twice the length of the last signal. Thus, signals 6–8 of multiquantum echo can be seen as analogs of multiquantum double-pulse echo signals. However, by comparing the conditions for the formation of these echo signals with the conditions for the formation of signals 5 (Table I) we can see that multiquantum triple-pulse echo is similar to double-pulse echo if there is an interval of ordinary one-quantum coherence ($\Delta m=1$ and $\Delta E/\hbar=\Delta\omega$) in the time interval between the first and second exciting pulses.

The dependence of the amplitude of the echo signal on the frequency of the alternating field in the exciting pulses was used to detect the magnetic resonance spectra. Additional studies have shown that the frequency spectra of multiquantum triple-pulse echo signals coincide with the frequency spectrum of the multiquantum double-pulse echo signal 3τ. Probably, the reason is that in the formation of both multiquantum double- and triple-pulse echo signals the exciting pulses couple the states with $\Delta E/\hbar=\Delta\omega$ and $\Delta E/\hbar=3\Delta\omega$, so that electric quadrupole interactions are effectively excluded from the process of echo signal formation and quadrupole interactions in the spectra of multiquantum echo signals are suppressed.

As a result of studies of the relaxation properties of double-pulse echo signals it was found that the amplitude of the multiquantum double-pulse echo signal 3τ decreases e-fold as the time interval τ between the exciting pulses increases by $70\pm5\ \mu$s. Such a decay rate is several times larger than that of ordinary echo at $t_e'=\tau$. To separate the moments of formation of different triple-pulse echo signals, we were forced to make the values of the time intervals τ_{12}

and τ_{23} very large (Table I), which could be the reason the amplitude of multiquantum echo signals is so small.

Moreover, the amplitude of the triple-pulse response [Eq. (4)] depends on a large number of parameters: the length t_j of each exciting pulse, the amplitude ω_1 of the alternating magnetic field, the average value of the quadrupole splitting of the NMR spectrum, etc. With such a large number of parameters one cannot exclude the possibility of nonoptimal signal excitation, which may be another reason the amplitude of multiquantum echo signals is so small.

4. CONCLUSION

We have analyzed theoretically the moments of formation of multiquantum signals of nuclear spin echo for the case of three exciting pulses. The situation involved both magnetic and quadrupole inhomogeneous spectral-line broadening. The formulas we derived can also be used in the limits of a purely magnetic and purely quadrupole inhomogeneous broadening. To do this we must drop the term with ω_q or with $\Delta\omega$ in (5).

We have verified by experiments the theoretical results for the quadrupole nucleus ^{63}Cu (spin $I=3/2$) in ferromagnetic copper sulfochromite alloyed with antimony. The lengths of the exciting pulses that ensure optimum conditions for the formation of multiquantum coherence in a quadrupole spin system have been established experimentally.

Our results can be used to study the relaxation characteristics of an inhomogeneously broadened spin system with quadrupole interaction. However, such studies require a theoretical analysis of the effect of fluctuations of the nonsecular part of the spin Hamiltonian, which is responsible for spin–lattice relaxation processes, on the rate of decay of echo signals. For the case of selective excitation of echo signals from quadrupole nuclei such an analysis has yet to be made.

This work was made possible by grants from ISSEP (Grants APU072083 and SPU062005).

*)E-mail: roton@ccssu.crimea.ua

---

[1] M. I. Kurkin and E. A. Turov, *NMR in Magnetically Ordered Materials and Its Applications* [in Russian], Nauka, Moscow (1990).

[2] A. Abragam, *The Principles of Nuclear Magnetism*, Clarendon Press, Oxford (1961).

[3] A. A. Vashman and I. S. Pronin, *Nuclear Magnetic Relaxation Spectroscopy* [in Russian], Énergoatomizdat, Moscow (1986).

[4] I. Solomon, Phys. Rev. **110**, 6 (1958).

[5] G. N. Abelyashev, V. N. Berzhanskiĭ, S. N. Polulyakh, N. A. Sergeev, and Yu. V. Fedotov, Zh. Éksp. Teor. Fiz. **100**, 1981 (1991) [Sov. Phys. JETP **73**, 1096 (1991)].

[6] G. N. Abelyashev, V. N. Berzhanskiĭ, S. N. Polulyakh, N. A. Sergeev, and Yu. V. Fedotov, Zh. Éksp. Teor. Fiz. **100**, 1255 (1991) [Sov. Phys. JETP **73**, 693 (1991)].

[7] V. I. Tsifrinovich, Zh. Éksp. Teor. Fiz. **94**(7), 208 (1988) [Sov. Phys. JETP **67**, 1413 (1988)].

[8] P. P. Man, Phys. Rev. B **52**, 9418 (1995).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Subdiffusion and stable laws

V. V. Uchaĭkin*)

Ul'yanovsk State University, 422700 Ul'yanovsk, Russia

(Submitted 18 September 1998)
Zh. Éksp. Teor. Fiz. **115**, 2113–2132 (June 1999)

This paper examines particle diffusion in N-dimensional Euclidean space with traps of the return type. Under the assumption that the random continuous-diffusion time has a finite mean value, it is established that subdiffusion (which is characterized by an increase in the width of the diffusion packet with time according to the t^{α}-law, where $\alpha<1$; for normal diffusion $\alpha=1$) emerges if and only if the distribution density of the random time a particle spends in a trap has a tail of the power-law type $\propto t^{\alpha-1}$. In these conditions the asymptotic expression for the distribution density of a diffusing particle is found in terms of the density of a one-sided stable law with a characteristic exponent α. It is shown that the density is a solution of subdiffusion equations in fractional derivatives. The physical meaning of the solution is discussed, and so are the properties of the solution and its relation to the results of other researchers in the field of anomalous-diffusion theory. Finally, the results of numerical calculations are discussed.
[S1063-7761(99)01606-6]

1. INTRODUCTION

Usually by anomalous diffusion one means a random-walk process involving a particle whose diffusion packet $\Delta(t)$ [i.e., the width of the distribution density $\rho(x,t)$ with the initial condition $\rho(x,0)=\delta(x)$] grows in time according to the law

$$\Delta t \propto t^{\nu}, \quad t\to\infty, \tag{1}$$

where the exponent ν differs from 1/2, a value that corresponds to normal diffusion. When $\nu>1/2$, we have superdiffusion, and when $\nu<1/2$, we have subdiffusion (see the reviews by Bouchaud and Georges[1], Isichenko,[2] and West and Deering[3]). The first type of anomalous diffusion is associated with anomalously long particle paths in the medium, ξ, while the second is associated with anomalously long times that the particle spends in a trap, τ. By anomalously long we mean that the random quantities ξ and τ are such that $\langle \xi^2\rangle=\infty$ and $\langle \tau\rangle=\infty$. The power law (1) emerges when the distributions of ξ and/or τ have tails of the power-law type. In Ref. 4 it is shown that in the asymptotic limit of very large times, superdiffusion is described by an equation with a fractional Laplacian whose solution is a symmetric stable distribution. The present paper is devoted to a theoretical study of the subdiffusion model, whose numerous applications to physical processes have been discussed in Refs. 1–3,5–10, and others.

If we assume that the random variables ξ_i and τ_i are mutually independent, subdiffusion can be described by integral equations. If we want the subdiffusion equation to look like an ordinary diffusion equation, we must consider equations in fractional derivatives.[5–10] Schneider and Wyss,[11] Glöckle and Nonnenmacher,[12] and West *et al.*[13] used Fox functions[14] to represent the solutions of such equation, but since Fox functions are peculiarly ill-suited to numerical problems, such representations are no better than Fourier–Laplace or Mellin transforms. The densities of the spatial distribution of a subdiffusive particle have yet to be found numerically.

The idea to seek the solution of subdiffusion equations in terms of stable laws emerged on the basis of two facts that are not well known to physicists. The first is that the Gaussian distribution is but one representative of an infinitely large set of stable laws, whose common property is that all these laws describe limiting distributions of sums of independent random quantities, with each sum being normalized in a special way.[15–17] The Gaussian distribution emerges as the limiting distribution only if the terms have finite or logarithmically divergent variances. The second fact is that a relationship exists[18,19] between stable laws and Fox functions.

Since different approaches to the problem of anomalous diffusion invoke different variants of the equations, we begin with a complete description of the model under consideration, based on integral equations.

2. INTEGRAL EQUATIONS OF THE DIFFUSION OF PARTICLES IN A MEDIUM WITH TRAPS

We use a model in which a particle can be in one of two states: a state of ordinary diffusion (state 1), or a state of rest (state 0) after it has landed in a trap. Subdiffusion is a process in which the particle state changes successively at random moments in time. We assume that the random time intervals within which the particle is in one of the two states are mutually independent and are distributed with densities $q_1(\tau)$ and $q_0(\tau)$.

The distribution of the time a particle stays in a trap, $q_0(\tau)$, which is determined by the trapping mechanism and the statistical spread of the trap properties, will not be nailed down for now. As for $q_1(\tau)$, we only assume that the mean time interval between a particle's leaving a trap arriving at the next trap is finite:

$$\overline{\tau_1}=\int_0^\infty \tau q_1(\tau)\,d\tau<\infty. \tag{2}$$

The medium is assumed spatially homogeneous and time-invariant.

Let $p(x,t)$ be the spatial distribution of the probability in the case of continuous diffusion. In N-dimensional space,

$$p(x,y)=\frac{1}{(4\pi Dt)^{N/2}}\exp\left\{-\frac{x^2}{4Dt}\right\},\quad x\in R^N, \tag{3}$$

where D is the diffusion coefficient. Next, by $\rho_0(x,t)$ we denote the particle distribution at time t, where the particle's history begins at $t=0$ when it lands in the trap at $x=0$, and by $\rho_1(x,t)$ we denote the particle distribution at time t, where the particle's history begins at $t=0$ when it leaves the trap at $x=0$. These two distributions are related by a pair of integral equations:

$$\rho_0(x,t)=Q_0(t)\,\delta(x)+\int_0^t d\tau\, q_0(\tau)\rho_1(x,t-\tau), \tag{4}$$

$$\rho_1(x,t)=Q_1(t)p(x,t)+\int_0^t d\tau\, q_1(\tau)p(x,\tau)*\rho_0(x,t-\tau), \tag{5}$$

where $Q_i(t)=\int_t^\infty q_i(\tau)\,d\tau$, and $*$ denotes spatial convolution:

$$p(x,\tau)*\rho_0(x,t-\tau)\equiv\int p(x',\tau)\rho_0(x-x',t-\tau)\,d^Nx'. \tag{6}$$

The system of equations (4) and (5) normally describes a more general class of processes, since it holds for an arbitrary distribution $p(x,t)$. In particular, instead of using the diffusion regime (3) in the interval between two traps, we can use the ballistic regime or, say, the superdiffusion regime. In this paper we limit ourselves to the study of solutions of Eqs. (4) and (5) with distribution (3). We select the distribution of the time a particle spends in a trap, $q_0(\tau)$, in a form that ensures that the subdiffusion regime prevails.

3. NECESSARY AND SUFFICIENT CONDITION FOR SUBDIFFUSION

Let us find the condition that the distribution $q_0(\tau)$ must meet so that the model leads to subdiffusion (1). Introducing, for the sake of brevity, the notation $s_i(t)=\int|x|^2\rho_i(x,t)\,d^Nx$ and using (4) and (5), we obtain

$$s_0(t)=\int_0^t d\tau\, q_0(\tau)s_1(t-\tau), \tag{7}$$

$$s_1(t)=Q_1(t)s(t)+\int_0^t d\tau\, q_1(\tau)[s(\tau)+s_0(t-\tau)], \tag{8}$$

where

$$s(t)=\int |x|^2p(x,t)\,d^Nx=at,\quad a=2ND. \tag{9}$$

Taking the Laplace transform of Eqs. (7) and (8), $s_i(\lambda)=\int_0^\infty \exp\{-\lambda t\}s_i(t)\,dt$, we obtain a system of algebraic equations for the components $s_0(\lambda)=q_0(\lambda)s_1(\lambda)$ and $s_1(\lambda)=K(\lambda)+q_1(\lambda)s_0(\lambda)$, where

$$K(\lambda)=\int_0^\infty dt\,\exp\{-\lambda t\}\left[Q_1(t)s(t)+\int_0^t d\tau\, q_1(\tau)s(\tau)\right].$$

Its solution has the form

$$s_0(\lambda)=\frac{q_0(\lambda)K(\lambda)}{1-q_0(\lambda)q_1(\lambda)}, \tag{10}$$

$$s_1(\lambda)=\frac{K(\lambda)}{1-q_0(\lambda)q_1(\lambda)}. \tag{11}$$

Using (9), we can transform $K(\lambda)$ into

$$\begin{aligned}K(\lambda)&=-a\frac{d}{d\lambda}\int_0^\infty Q_1(t)\exp\{-\lambda t\}dt-\frac{a}{\lambda}\frac{dq_1\lambda}{d\lambda}\\&=-a\frac{d}{d\lambda}\frac{1-q_1(\lambda)}{\lambda}+\frac{a}{\lambda}\frac{d}{d\lambda}[1-q_1(\lambda)]\\&=\frac{a[1-q_1(\lambda)]}{\lambda^2}.\end{aligned} \tag{12}$$

According to Tauber's theorem, $s_i(t)\sim A_it^\alpha$, $t\to\infty$, implies

$$s_i(\lambda)\sim\Gamma(\alpha+1)A_i\lambda^{-\alpha-1},\quad \lambda\to 0, \tag{13}$$

and conversely.[20] By virtue of (2) we have

$$1-q_1(\lambda)\sim\overline{\tau_1}\lambda,\quad Q_1(\lambda)=\frac{1-q_0(\lambda)}{\lambda}\sim\overline{\tau_1}, \tag{14}$$

so that $K(\lambda)\sim a\overline{\tau_1}/\lambda$ as $\lambda\to 0$. Substituting (13) into (10) and (11) and solving the resulting equations for $1-q_0(\lambda)$, we find the necessary condition for subdiffusion:

$$1-q_0(\lambda)\sim b\lambda^\alpha,\quad \lambda\to 0,\quad b=\frac{a\overline{\tau_1}}{\Gamma(\alpha+1)A},\quad \alpha<1, \tag{15}$$

with $A_1=A_2=A$ (the asymptotic behavior of the width of the subdiffusion packet is independent of the initial particle state). By virtue of the reciprocity of Tauber's theorem, the condition (14) is also sufficient.

To reformulate the condition (14) for the distribution density $q_0(\tau)$, we again turn to Tauber's theorem and apply it to the function $Q_0(t)$, with the result that $Q_0(\lambda)=[1-q_0(\lambda)]/\lambda$. We obtain

$$Q_0(t)=\int_t^\infty q_0(\tau)\,d\tau\sim Bt^{-\alpha},\quad t\to\infty,$$

$$B=\frac{a\overline{\tau_1}}{[\Gamma(1-\alpha)]^2A}, \tag{16}$$

or, for the density,

$$q_0(t)\sim\alpha Bt^{-\alpha-1},\quad t\to\infty. \tag{17}$$

Thus, in the model considered, subdiffusion emerges if and only if the distribution of the times particles stay in traps exhibits asymptotic behavior of the power-law type (16) with an exponent $\alpha < 1$. This means, in particular, that the average time a particle stays in a trap is infinite: $\int_0^\infty \tau q_0(\tau)\,d\tau = \infty$, $\alpha < 1$. If it is finite, or $\int_0^\infty \tau q_0(\tau)\,d\tau = \overline{\tau_0}$, the asymptotic behavior of $q_0(\lambda)$ is

$$q_0(\lambda) \sim 1 - \overline{\tau_0}\lambda, \quad \lambda \to 0. \tag{18}$$

Substituting (14) and (18) into (11), we see that in this case

$$s_1(\lambda) \sim \frac{a}{\lambda^2[1+\overline{\tau_0}/\overline{\tau_1}]}, \quad \lambda \to 0,$$

with the result that the effect of traps reduces to a variation in the diffusion coefficient, $D \to D/(1+\overline{\tau_0}/\overline{\tau_1})$, and that the temporal variation of the mean square $s_1(t)$ remains linear. It can be shown that the diffusion packet in this case remains Gaussian.

4. DIFFERENTIAL EQUATIONS OF SUBDIFFUSION

We go back to Eqs. (4) and (5) and take Fourier and Laplace transforms with respect to position and time, respectively:

$$\rho_i(k,\lambda) = \int_0^\infty dt \int d^N x \exp\{-\lambda t + ikx\}\, p_i(x,t),$$

with $k \in R^N$. This yields

$$\rho_0(k,\lambda) = Q_0(\lambda) + q_0(\lambda)\rho_1(k,\lambda),$$

$$\rho_1(k,\lambda) = Q_1(\lambda + Dk^2) + q_1(\lambda + Dk^2)\rho_0(k,\lambda).$$

The solution of this system has the form

$$\rho_0(k,\lambda) = \frac{Q_0(\lambda) + q_0(\lambda)Q_1(\lambda + Dk^2)}{1 - q_0(\lambda)q_1(\lambda + Dk^2)}, \tag{19}$$

$$\rho_1(k,\lambda) = \frac{Q_1(\lambda + Dk^2) + Q_0(\lambda)q_1(\lambda + Dk^2)}{1 - q_0(\lambda)q_1(\lambda + Dk^2)}. \tag{20}$$

Combining the conditions (15)–(17) with Eqs. (19) and (20), we obtain an expression for the leading asymptotic terms,

$$\rho^{\mathrm{as}}(k,\lambda) = \frac{\lambda^\alpha}{\lambda[D'k^2 + \lambda^\alpha]}, \quad D' = \frac{\overline{\tau_1}D}{b}, \tag{21}$$

which is independent of the initial state. We postpone taking the inverse transform to Sec. 5. Here we write the above relationship in three equivalent forms:

$$\lambda^\alpha \rho^{\mathrm{as}}(k,\lambda) = -D'k^2\rho^{\mathrm{as}}(k,\lambda) + \lambda^{\alpha-1}, \tag{22}$$

$$\lambda \rho^{\mathrm{as}}(k,\lambda) = -D'k^2\lambda^{1-\alpha}\rho^{\mathrm{as}}(k,\lambda) + 1, \tag{23}$$

$$\rho^{\mathrm{as}}(k,\lambda) = -D'k^2\lambda^{-\alpha}\rho^{\mathrm{as}}(k,\lambda) + \lambda^{-1}. \tag{24}$$

As is known, on a suitable class of functions the Laplace transform $F(\lambda)$ of the Riemann–Liouville fractional derivative

$$F(t) = \frac{d^\mu f(t)}{dt^\mu} \equiv \frac{1}{\Gamma(-\mu)}\frac{d}{dt}\int_0^t \frac{f(\tau)\,d\tau}{(t-\tau)^\mu}, \quad \mu < 1, \tag{25}$$

is related to $f(\lambda)$ of the differentiable function $f(t)$:[21,22]

$$F(\lambda) = \lambda^\mu f(\lambda). \tag{26}$$

When $\mu < 0$, the expression (25) is a fractional integral of order $|\mu|$. Using this notation in the inverse Fourier–Laplace transform of Eqs. (22)–(24), we obtain an equation in fractional derivatives that describes the asymptotic behavior of the subdiffusion process:

$$\frac{\partial^\alpha \rho^{\mathrm{as}}}{\partial t^\alpha} = D'\nabla^2\rho^{\mathrm{as}} + \frac{t^{-\alpha}}{\Gamma(1-\alpha)}\delta(x), \tag{27}$$

$$\frac{\partial \rho^{\mathrm{as}}}{\partial t} = D'\nabla^2\frac{\partial^{1-\alpha}\rho^{\mathrm{as}}}{\partial t^{1-\alpha}} + \delta(x)\delta(t), \tag{28}$$

$$\rho^{\mathrm{as}} = D'\nabla^2\frac{\partial^{-\alpha}\rho^{\mathrm{as}}}{\partial t^{-\alpha}} + \delta(x). \tag{29}$$

These equations have a general solution with the Fourier–Laplace transform found earlier, which is represented by Eq. (21).

Note that the special case of Eq. (27) corresponding to $\alpha = 1/2$ was obtained by Nigmatullin[5] in connection with diffusion in fractal structures of the Koch-tree type, which models porous and disordered media. The one-dimensional analog of Eq. (28) was written out by Compte,[10] and the integral equation (29) was solved by Schneider and Wyss.[11] We now discuss their solution.

5. SUBDIFFUSION DISTRIBUTION DENSITY

Let $D' = 1$ in Eq. (29) and write the latter

$$\rho(x,t) = \delta(x) + \frac{1}{\Gamma(\alpha)}\int_0^t \frac{dt\nabla^2\rho(x,t)}{(t-\tau)^{1-\alpha}}, \quad 0 < \alpha < 1. \tag{30}$$

Equation (30) was studied by Schneider and Wyss,[11] who expressed its solution as a function of the distance $r = |x|$ in terms of the Fox functions:

$$\rho(r,t) = \frac{1}{\alpha^1\pi^{N/2}r^N}H^{20}_{12}\left(\left(\frac{r}{2}\right)^{2/\alpha}\frac{1}{t}\,\middle|\,\begin{matrix}(1, & 1\\ (N/2,1/\alpha), & (1,1/\alpha)\end{matrix}\right). \tag{31}$$

They also found the explicit form of the Mellin component in r:

$$\rho(s,t) = \int_0^\infty r^{s-1}\rho(r,t)\,dr$$

$$= \frac{2^{s-N-1}}{\pi^{N/2}}t^{\alpha(s-N)/2}\frac{\Gamma(s/2)\Gamma((s-N)/2)}{\alpha\Gamma(\alpha(s-N)/2)}. \tag{32}$$

We establish another form of the solution that relates the form to stable distributions. This will make it possible not only to carry out a qualitative analysis but also to understand the physics of the solution.

We write (21) as

$$\rho^{\mathrm{as}}(k,\lambda) = \lambda^{\alpha-1}\int_0^\infty \exp\{-[D'k^2 + \lambda^\alpha]y\}\,dy, \tag{33}$$

with $\alpha<1$, and also write the inverse Laplace transform:

$$\rho^{\mathrm{as}}(k,t)=\int_0^\infty dy\,\exp\{-D'k^2y\}$$

$$\times(2\pi i)^{-1}\int_\gamma d\lambda\,\lambda^{\alpha-1}\exp\{\lambda t-\lambda^\alpha y\}.$$

Evaluating the innermost integral by parts, we obtain

$$\rho^{\mathrm{as}}(k,t)=\frac{t}{\alpha}\int_0^\infty dy\,\exp\{-D'k^2y\}y^{-1}$$

$$\times(2\pi i)^{-1}\int_\gamma \exp\{\lambda t-\lambda^\alpha y\}d\lambda.$$

In the innermost integral we transform to the new variable $s=y^{1/\alpha}\lambda$:

$$\rho^{\mathrm{as}}(k,t)=\frac{t}{\alpha}\int_0^\infty dy\,\frac{\exp\{-D'k^2y\}}{y^{1+1/\alpha}}$$

$$\times\left[(2\pi i)^{-1}\int_\gamma \exp\{sy^{-1/\alpha}t-s^\alpha\}\,ds\right].$$

The expression in square brackets is a one-sided stable density with characteristic exponent $\alpha<1$ [see Eq. (A1)]:

$$g^{(\alpha)}(t)=(2\pi i)^{-1}\int_\gamma \exp\{st-s^\alpha\}\,ds. \tag{34}$$

Thus,

$$\rho^{\mathrm{as}}(k,t)=\alpha^{-1}t\int_0^\infty dy\,\frac{\exp\{-D'k^2y\}}{y^{1+1/\alpha}}\,g^{(\alpha)}(y^{-1/\alpha}t).$$

Introducing the integration variable $\tau=y^{-1/\alpha}t$, we find that

$$\rho^{\mathrm{as}}(k,t)=\int_0^\infty d\tau\,\exp\left\{-\frac{D'k^2t^\alpha}{\tau^\alpha}\right\}g^{(\alpha)}(\tau).$$

Finally, taking the inverse Fourier transform, we obtain

$$\rho^{\mathrm{as}}(x,t)=\frac{1}{(D't^\alpha)^{N/2}}\Psi_N^{(\alpha)}(|x|/\sqrt{D't^\alpha}), \tag{35}$$

where

$$\Psi_N^{(\alpha)}(r)=\frac{1}{(4\pi)^{N/2}}\int_0^\infty d\tau\,\exp\left\{-\frac{r^2\tau^\alpha}{4}\right\}\tau^{N\alpha/2}g^{(\alpha)}(\tau), \tag{36}$$

with $\alpha<1$, is a function of the distance and depends on two parameters: the subdiffusion exponent α and the dimensionality N of the space.

Note that both (21) and (33) have the same meaning at the limit $\alpha=1$ and lead to normal diffusion with the same coefficient D'. This means that the function $\Psi_N^{(\alpha)}(r)$ can be redefined so that it holds at $\alpha=1$:

$$\Psi_N^{(\alpha)}(r)=(4\pi)^{-N/2}\exp\left\{-\frac{r^2}{4}\right\}.$$

The subdiffusion distribution in the form (35) can be defined as an ordinary diffusion distribution from simple probabilistic considerations based on the central limit theorem in its generalized form (A7) and (A8).[16,17,23] Ignoring the particle dwell time in the diffusion state in our calculations of the distribution of the number of trap events over the observation time $t\to\infty$, we find that

$$p_n\approx Q_0^{(n+1)}(t)-Q_0^{(n)}(t)$$

$$=G^{(\alpha)}((nB^*)^{-1/\alpha}t)-G^{(\alpha)}([(n+1)B^*]^{-1/\alpha}t).$$

Representing the argument of the subtracted function in the form

$$[(n+1)B^*]^{-1/\alpha}t=[nB^*]^{-1/\alpha}t-[nB^*]^{-1/\alpha}t(n\alpha)^{-1}$$

and expanding in series, we obtain the asymptotic expression

$$p_n\sim[nB^*]^{-1/\alpha}t(n\alpha)^{-1}g^{(\alpha)}([nB^*]^{-1/\alpha}t),\quad t\to\infty.$$

When n is fixed, the conditional distribution of the coordinates of a particle can be expressed in terms of the ordinary diffusion density as $\rho(x,t|n)\sim p(x,n/\mu)$. Here the random diffusion time is replaced by the mean value n/μ for understandable reasons. Averaging over the number of continuous-diffusion events, $\rho(x,t)=\Sigma_n\rho(x,t|n)p_n$, and replacing summation over n by integration with respect to the variable $\tau=[nB^*]^{-1/\alpha}t$, we obtain the distribution (35).

A convenient way to compare our solution with the one obtained by Schneider and Wyss[11] is to compare the Mellin transforms of the two solutions. According to (35),

$$\rho^{\mathrm{as}}(s,t)=\frac{1}{2}\pi^{-N/2}(4D't^\alpha)^{(s-N)/2}\Gamma\left(\frac{s}{2}\right)$$

$$\times\int_0^\infty \tau^{(N-s)\alpha/2}g^{(\alpha)}(\tau)\,d\tau.$$

Expressing the surviving integral in terms of gamma functions via (A2) and comparing the result with (32) at $D'=1$, we see that the solutions are identical.

6. ANALYSIS OF SUBDIFFUSION DISTRIBUTIONS

In this section we discuss some properties of the solutions, examine their asymptotic behavior at small and large distances, and discuss the results of numerical calculations.

We begin with the spatial moments, which can be explicitly expressed in terms of the Mellin transforms discussed above:

$$\langle|x|^{2n}\rangle=\frac{2\pi^{N/2}}{\Gamma(N/2)}\int_0^\infty t^{2n+N-1}\rho^{\mathrm{as}}(r,t)\,dr$$

$$=\frac{\Gamma(n+1)\Gamma(N/2+n)}{\Gamma(\alpha n+1)\Gamma(N/2)}(4D't^\alpha)^n.$$

The second moment

$$\langle|x|^2\rangle=\frac{2ND'}{\Gamma(\alpha+1)}t^\alpha, \tag{37}$$

which is the same as the one calculated by Schneider and Wyss (Eq. (1.14) in Ref. 11), increases with time in proportion to t^α, $0<\alpha<1$, a hallmark of subdiffusion. The ratio

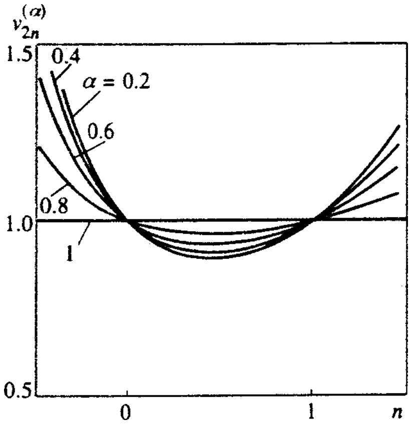

FIG. 1. Ratio of dimensionless moments, $\nu_{2s}^{(\alpha)}$, for $\alpha=0.2$, 0.4, 0.6, 0.8, and 1.

$$\mu_{2n}^{(\alpha)}=\frac{\langle|x|^{2n}\rangle}{\langle|x|^2\rangle^n}=\frac{2^n\Gamma(n+1)\Gamma(N/2+n)[\Gamma(\alpha+1)]^n}{N^n\Gamma(\alpha n+1)\Gamma(N/2)},$$

which represents dimensionless moments of orders higher than the second, does not depend on time. This suggests that the shape of the distribution remains unchanged [incidentally, this follows immediately from (35)]. At $\alpha=1$ the ratio yields the dimensionless moments of the normal distribution:

$$\mu_{2n}^{(1)}=\frac{2^n\Gamma(N/2+n)}{N^n\Gamma(N/2)}.$$

Figure 1 depicts the ratio of the dimensionless absolute moments of order $2s$ $(-1/2<s<3/2)$,

$$\nu_{2s}^{(\alpha)}=\frac{\mu_{2s}^{(\alpha)}}{\mu_{2s}^{(1)}}=\frac{\Gamma(s+1)[\Gamma(\alpha+1)]^s}{\Gamma(\alpha s+1)},$$

which characterizes the difference between the shape of the subdiffusion distribution $\Psi_N^{(\alpha)}(r)$ for $\alpha<1$ and the normal distribution $\Psi_N^{(1)}(r)$. Close to zero $(s<0)$ and at large distances $(s>1)$, $\Phi_N^{(\alpha)}(r)$ exceeds the normal distribution, while in the transitional region $0<s<1$ the opposite is true.

Moving on to analyze the shape of the distribution $\Psi_N^{(\alpha)}(r)$, we note first and foremost that by differentiating (36) with respect to r we can easily obtain a relationship between the distributions in N- and $N+2$-dimensional space:

$$\Psi_{N+2}^{(\alpha)}(r)=-\frac{1}{2\pi r}\frac{d\Psi_N^{(\alpha)}(r)}{dr}. \tag{38}$$

Let $r=(x_1,\dots,x_N)$ be an N-dimensional vector, so that $r=\sqrt{x_1^2+\dots+x_N^2}$. Integrating (36) with respect to the variables $x_{n+1},\dots,x_N$, where $1<n<N$, we obtain

$$\int dx_{n+1}\cdots\int dx_N\,\Psi_N^{(\alpha)}(\sqrt{x_1^2+\dots+x_N^2})$$
$$=\Psi_n^{(\alpha)}(\sqrt{x_1^2+\dots+x_n^2}),$$

which means that the behavior of the projection of a random point in subdiffusive motion in N-dimensional space onto an n-dimensional subspace is described by an n-dimensional subdiffusion equation with the same characteristic exponent α. In this respect the situation is similar to that in normal diffusion, but there is an important difference. In normal diffusion, the coordinates X_1 and X_2 of a diffusing particle are mutually independent, while in subdiffusion the joint distribution of the particles,

$$P\{X_1\in dx_1,X_2\in dx_2\}$$
$$=(4\pi)^{-1}\int_0^\infty d\tau\exp\left\{-\frac{(x_1^2+x_2^2)\tau^\alpha}{4}\right\}\tau^\alpha g^{(\alpha)}(\tau),$$

does not reduce to the product $P\{X_1\in dx_1\}P\{X_2\in dx_2\}$, with the result that the random coordinates X_1 and X_2 cease to be independent. The nature of their statistical dependence at small and large distances can be clarified by examining the asymptotics.

Equation (36) implies that in the one-dimensional $(N=1)$ case the distribution density at the origin exists:

$$\Psi_1^{(\alpha)}(0)=(4\pi)^{-1/2}\int_0^\infty\tau^{\alpha/2}g^{(\alpha)}(\tau)\,d\tau,$$

since $g^{(\alpha)}(\tau)$ has finite moments of order less than α. Making use of Eq. (A2), we obtain

$$\Psi_1^{(\alpha)}(0)=[2\Gamma(1-\alpha/2)]^{-1}.$$

In spaces with $N\geqslant2$, the subdiffusion density (in contrast to the normal diffusion density) has an integrable singularity. When $N=2$, this singularity is logarithmic, as can be easily verified by splitting the integral in (36) into two parts, the transient and the asymptotic, and replacing the density $g^{(\alpha)}(\tau)$ in the latter by the leading term in the expansion (A5):

$$\Psi_2^{(\alpha)}(r)\approx(4\pi)^{-1}\left(\int_0^T\exp\left\{-\frac{r^2\tau^\alpha}{4}\right\}\tau^\alpha g^{(\alpha)}(\tau)\,d\tau\right.$$
$$\left.+\frac{\Gamma(1+\alpha)\sin\pi\alpha}{\pi}\int_T^\infty\exp\left\{-\frac{r^2\tau^\alpha}{4}\right\}\tau^{-1}d\tau\right).$$

When $r\to\infty$, the second term in this sum dominates, which leads to a logarithmic singularity:

$$\Psi_2^{(\alpha)}(r)\sim[4\pi\Gamma(1-\alpha)]^{-1}E_1(r^2T^{\alpha/4})$$
$$\sim[2\pi\Gamma(1-\alpha)]^{-1}|\ln r|,\quad r\to0.$$

In a space with $N\geqslant3$, the singularity at the origin is hyperbolic:

$$\Psi_N^{(\alpha)}(r)\sim\frac{\Gamma(\alpha+1)\sin\pi\alpha}{\pi(4\pi)^{N/2}}\int_0^\infty\exp\left\{-\frac{r^2\tau^\alpha}{4}\right\}T^{(N/2-1)\alpha-1}\,d\tau$$
$$=\frac{1}{4\pi^{N/2}}\frac{\Gamma(N/2-1)}{\Gamma(1-\alpha)}r^{-(N-2)},\quad r\to0. \tag{39}$$

Setting $N=1$ in (38) and substituting the asymptotic expression (39) into the left-hand side, we find that

$$\frac{d\Psi_1^{(\alpha)}(r)}{dr}\to-\frac{2}{\Gamma(1-\alpha)}$$

as $r\to\infty$. This implies that at $x=0$ the derivative of the function $\Psi_1^{(\alpha)}(x)$ has a finite discontinuity, i.e., rather than

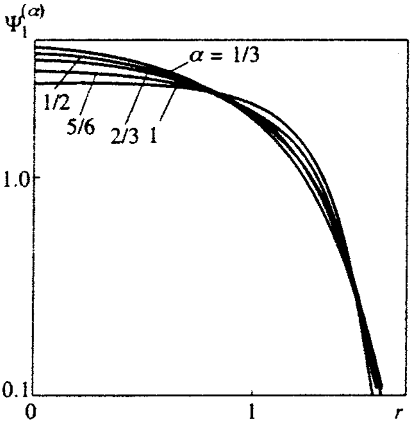

FIG. 2. One-dimensional distributions $\Psi_1^{(\alpha)}(r)$ for $\alpha=2/6$, 3/6, 4/6, 5/6, and 1.

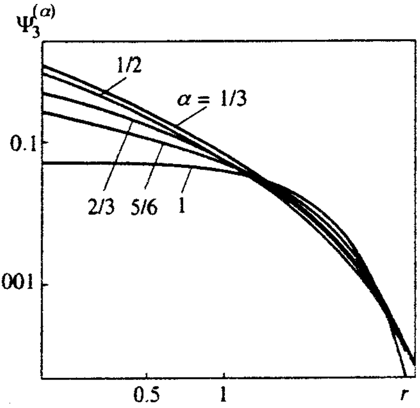

FIG. 4. Three-dimensional distributions $\Psi_3^{(\alpha)}(r)$ for the same values of α as in Fig. 2.

being smooth (as in normal diffusion), the peak of the distribution in the one-dimensional case is a cusp. As $\alpha\to1$, the derivative vanishes and the vertex becomes smooth.

When r is large, the exponential in the integrand of (36) rapidly decreases, so that to make an asymptotic estimate of the integral we use the expression (A6), which approximates the stable density at small values of the argument. We calculate the resulting integral

$$\Psi_N^{(\alpha)}(r)\approx(4\pi)^{-N/2}A\int_0^\infty \exp\left\{-\frac{r^2\tau^\alpha}{4}-br^{-\delta}\right\}T^{N\alpha/2-\gamma}\,d\tau \tag{40}$$

by the Laplace method. We begin by introducing the notation

$$\varphi(\tau)=-\frac{r^2\tau^\alpha}{4}-b\tau^{-\delta}. \tag{41}$$

Using the condition $\varphi(\overline{\tau})=0$, we find the position of the maximum of this function:

$$\overline{\tau}=\left(\frac{4b\delta}{\alpha r^2}\right)^{1/(\delta+\alpha)}. \tag{42}$$

Proceeding in the usual way, we find that

$$\int_0^\infty \exp\left\{-\frac{r^2\tau^\alpha}{4}-b\tau^{-\delta}\right\}T^{(N\alpha/2-\gamma)}d\tau$$
$$\sim\frac{A\overline{\tau}^{-(N\alpha/2-\gamma)/(\delta+\alpha)}}{\pi^{N/2}}\sqrt{\frac{2\pi}{|\varphi''(\overline{\tau})|}}\exp\{\varphi(\overline{\tau})\}. \tag{43}$$

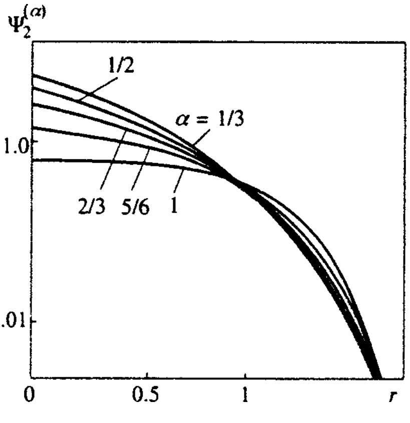

FIG. 3. Two-dimensional distributions $\Psi_2^{(\alpha)}(r)$ for the same values of α as in Fig. 2.

Substituting (41) and (42) into (43), and the result into (40), we obtain the asymptotic expression (40) in the form

$$\Psi_N^{(\alpha)}(r)\sim\frac{1}{(4\pi)^{N/2}}\frac{\alpha^{[(N+1)\alpha/2-1]/(2-\alpha)}}{\sqrt{2-\alpha}}$$
$$\times\left(\frac{r}{2}\right)^{-N(1-\alpha)/(2-\alpha)}$$
$$\times\exp\left\{-(2-\alpha)\alpha^{\alpha/(2-\alpha)}\left(\frac{r}{2}\right)^{2/(2-\alpha)}\right\}. \tag{44}$$

A comparison with exact calculations (carried out below) shows that for $\alpha>1/2$, Eq. (44) provides a satisfactory approximation of the distribution over the entire region, except at small distances, and as $\alpha\to1$ it turns into the normal distribution $\Psi_N^{(1)}(r)=(4\pi)^{-N/2}\exp\{-r^2/4\}$, i.e., it becomes exact.

Figures 2–4 provide an accurate idea of the shape of the distributions. They depict the subdiffusion distributions $\Psi_N^{(\alpha)}(r)$ for several values of α, including the limit $\alpha=1$ corresponding to normal diffusion [the variances of these distributions are different and depend on α according to (37)]. An important difference in the shape of subdiffusion distributions that sets them apart from normal distributions is the higher concentration of probability at both small and large distances. But if these features do not play a significant role in a specific problem, in the one-dimensional case with $\alpha>1/2$ subdiffusion distributions can indeed be approximated by a Gaussian with subdiffusion variance, as demonstrated by Klimontovich[24] (see Fig. 5). For spaces of higher dimensionality, the normal approximation fails to yield satisfactory results.

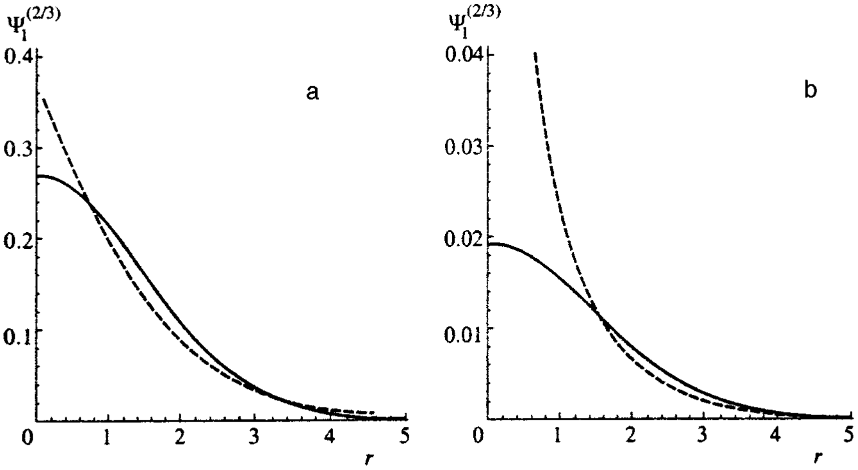

FIG. 5. Comparison of subdiffusion distributions (solid curves) and normal distributions (dashed curves) with the same variance, $\alpha=2/3$. (a) $N=1$, and (b) $N=3$.

7. DISCUSSION

The fact that there are different formulations of the anomalous-diffusion problem and different ways in which the results can be presented sometimes leads to a situation in which researchers fail to see the logical relations among them, even when the same problem is being studied. The approach developed in the present paper—from the random-walk model based on integral equations to the asymptotic part of solutions of these equations that satisfy equations in fractional derivatives—makes it possible not only to express the coefficients of anomalous diffusion in terms the characteristics of ''elementary distributions,'' but also to establish a relationship among the solutions obtained by various means. That such a problem exists can easily be seen by comparing some of the papers devoted to anomalous diffusion[1,11] with Ref. 13.

As noted above, Schneider and Wyss[11] used the integral equation (29) with a multidimensional Laplacian, found the Mellin and Laplace transforms, expressed the solution in terms of Fox functions, found an approximate expression [that is exactly equivalent to Eq. (44)] for the density at large distances, and obtained in the one-dimensional case an exact expression for the density of the form (in our notation)

$$\Psi_1^{\alpha}(r)=\frac{1}{\alpha r^{1+2/\alpha}}g^{(\alpha/2)}(r^{-2/\alpha}). \tag{45}$$

If in (36) we put $N=1$ and use (A3), we obtain the same formula, which advantageously differs from (36) in that there is no need to integrate now. Note that, according to the property (38), we can express the distributions in spaces with a large odd number of dimensions in terms of the density $g^{(\alpha/2)}$ and its derivatives. As is well known, however, integration is preferable to differentiation in numerical calculations.

At $\alpha=2/3$, the distribution (45) can be expressed, according to (A4), in terms of modified Bessel functions of the second kind:

$$\Psi_1^{(2/3)}(r)=\frac{1}{3\pi}\sqrt{r}\,K_{1/3}\frac{2r^{3/2}}{\sqrt{27}}.$$

In Sec. 1.2.3.1 of their review, Bouchaud and Georges[1] discuss the problem in which a particle hops suddenly from one point to another with a density characterized by a finite value of the rms hop length, and by a distribution of time intervals between successive hops satisfying the condition (16). Using the central limit theorem (as we did at the end of Sec. 5 of the present paper), they found the Laplace transform $\rho^{\mathrm{as}}(x,\lambda)$, which is the same as the one obtained by Schneider and Wyss[11] [Eqs. (2.8) and (2.10)], established the self-similar behavior of the distribution, i.e., introduced the function $\Psi_N(r)$, and once more derived Eq. (45) without any mention of Ref. 11.

The work of Schneider and Wyss[11] is mentioned in the Introduction to the paper by West *et al.*,[13] who nevertheless believed that they solved a different problem [not the one solved by Schneider and Wyss[11] with the use of Eq. (28)]. They expressed their solution (only the one-dimensional case was considered) in terms of Fox functions and gave the approximate expression (44), derived earlier by Schneider and Wyss.[11] West *et al.*[13] gave neither a general formula for the multidimensional case nor the exact solution (45) for the one-dimensional case, and the well known review of Bouchaud and Georges[1] was not cited at all. It must also be noted that by writing the equation in the form

$$\frac{\partial\sigma_0(x,t)}{\partial t}=C\frac{\partial^2}{\partial x^2}\frac{\partial^{\beta}\sigma_0(x,t)}{\partial t^{\beta}}, \tag{46}$$

West *et al.*[13] used a nonstandard notation for the fractional derivative

$$\frac{\partial^{\beta}f(t)}{\partial t^{\beta}}=\frac{1}{r(1-\beta)}\int_0^t\frac{f(t')\,dt'}{(t-t')^{\beta}},$$

instead of the standard notation (25). As a result, Eq. (46) corresponds to Eq. (28) at $\beta-1=1-\alpha$, i.e., $\beta=2-\alpha$. With this modification, the results of West *et al.*[13] are identical to their analogs in the cited papers and in the present paper, but the improperly defined order of the fractional derivative led them to believe that for $\alpha>1$ their solution describes subdiffusion (see the remark to Eq. (44) and Figs. 2–5 in Ref. 13, which depict the distributions for $\alpha>1$). In actuality, however, as Sec. 4 of the present paper suggests, the parameter α cannot exceed unity in this problem: even if we put $\alpha>1$ in the distribution (16) where it first appears, the transforms

(19) and (20) will lead, due to (18), to an ordinary diffusion equation, i.e., to Eq. (28) with $\alpha=1$. As Ref. 4 suggests, the superdiffusion regime is described by equations containing fractional derivatives with respect to the spatial variables (incidentally, the same is stated in the last section of Ref. 13).

In conclusion we note that the representation of subdiffusion distributions in terms of steady distributions (in contrast to Fox functions) appears to be more convenient, physically clear, and logically justified (in the sense of the limit theorem). The properties of stable distributions have been thoroughly studied and the densities have been tabulated, so that they can be added to the class of special functions.[19]

The author is grateful to S. A. Korobko for doing the necessary numerical calculations of the distributions in this paper. The work was supported by a grant from the Russian Fund for Fundamental Research (Grant No. 98-01-03307).

APPENDIX ONE-SIDED STABLE LAWS

One-sided stable densities $g^{(\alpha)}(t)$, $\alpha<1$, comprise a subset of the family of strictly stable laws defined in the following way: the density $g(t)$ is strictly stable if and only if for any two positive numbers b_1 and b_2 there exists a positive number b such that

$$\frac{1}{b_1 b_2} g\left(\frac{t}{b_1}\right) * g\left(\frac{t}{b_2}\right) = \frac{1}{b} g\left(\frac{t}{b}\right).$$

In other words, the shape of strictly stable distributions is invariant under convolutions (the best known representative of this class of distributions is the Gaussian distribution, corresponding to $\alpha=2$, but it is not a member of the subset of one-sided distributions considered here).

The characteristic functions of one-sided stable distributions have the simple form (form (B) in Ref. 19)

$$\varphi^{(\alpha)}(k) = \int_0^\infty \exp\{ikt\} g^{(\alpha)}(t)\, dt$$
$$= \exp\left\{ -|k|^\alpha \exp\left[-\frac{i(\alpha\pi/2)k}{|k|} \right] \right\}.$$

According to Lemma 2.2.1 of Ref. 19, the analytic continuation of the function $\varphi^{(\alpha)}(k)$ from the entire real k axis to the complex z plane with a cut along the ray $\arg z = -(3/4)\pi$ is given by the function

$$\varphi^{(\alpha)}(z) = \exp\{-(-iz)^\alpha\}, \quad \alpha<1,$$

which implies that $g^{(\alpha)}(\lambda) = \int_0^\infty \exp\{-\lambda t\} g^{(\alpha)}(t)\, dt$, the Laplace transform of the one-sided stable density $g^{(\alpha)}(t)$, has the form

$$g^{(\alpha)}(\lambda) = \varphi^{(\alpha)}(i\lambda) = \exp\{-\lambda^\alpha\}. \tag{A1}$$

The Mellin transform of the stable density can be expressed in terms of the ratio of two gamma functions:[17,18]

$$g^{(\alpha)}(s) \equiv \int_0^\infty t^s g^{(\alpha)}(t)\, dt = \frac{\Gamma(1-s/\alpha)}{\Gamma(1-s)}. \tag{A2}$$

We have the relationship

$$\frac{1}{\sqrt{4\pi}} \int_0^\infty \exp\left\{ -\frac{r^2 \tau^\alpha}{4} \right\} g^{(\alpha)}(\tau) \tau^{\alpha/2}\, d\tau = \frac{g^{(\alpha/2)}(r^{-2/\alpha})}{\alpha r^{1+2/\alpha}}, \tag{A3}$$

whose validity can easily be proved by taking the Mellin transform and using (A2).

The density $g^{(\alpha)}(t)$ can be expressed in terms of elementary functions only if $\alpha=1/2$:

$$g^{(1/2)}(t) = \frac{1}{2\sqrt{\pi}} t^{-3/2} \exp\left\{ -\frac{1}{4t} \right\}.$$

This is the Lévy distribution, also known as the Smirnov distribution[17,25] (named after N. V. Smirnov). At $\alpha=1/3$ and $\alpha=2/3$ the stable density (34) can be expressed in terms of the modified Bessel function and Whittaker function:[17]

$$g^{(1/2)}(t) = (3\pi)^{-1} t^{-3/2} K_{1/3}(2/\sqrt{27t}),$$
$$g^{(2/3)}(t) = \sqrt{3/\pi}\, t^{-1} \exp\{-u/2\} W_{1/2,1/6}(u), \tag{A4}$$

where $u=(4/27)t^{-2}$. When α is rational, $g^{(\alpha)}(t)$ can be represented by a finite sum of generalized hypergeometric functions,[26], e.g.,

$$g^{(3/4)}(t) = -\frac{1}{t} \frac{8}{3\pi} \sum_{n=1}^{3} \sin\left(\frac{3n\pi}{4}\right) z^n {}_2F_2\left(\frac{1}{3}+\frac{n}{4}, \frac{2}{3} + \frac{n}{4}; \frac{1}{2} + \frac{n(n-1)}{8}, \frac{n(7-n)}{8}; -z^4\right),$$

with $z=-(3/t)^{3/4}/4$. For an arbitrary function of α, the functions $g^{(\alpha)}(t)$ are related to the Fox functions by[18]

$$g^{(\alpha)}(x) = \alpha^{-1} x^{-2} H^{10}_{11}\left(x^{-1} \middle| \begin{matrix} (-1, & 1) \\ (-\alpha^{-1}, & \alpha^{-1}) \end{matrix} \right).$$

For numerical purposes, it is convenient to represent the density as the integral of a nonoscillating function. This representation was obtained by Zolotarev[27] by deforming the integration contour in (34) in a special way:

$$g^{(\alpha)}(t) = \frac{\alpha t^{1/(\alpha-1)}}{\pi(1-\alpha)} \int_{-\pi/2}^{\pi/2} U_\alpha(\varphi) \exp\{-t^{\alpha/(\alpha-1)} U_\alpha(\varphi)\}\, d\varphi,$$

where

$$U_\alpha(\varphi) = \left[\frac{\sin\alpha(\varphi+\pi/2)}{\cos\varphi} \right]^{\alpha/(1-\alpha)} \frac{\cos[\pi\alpha/2-(1-\alpha)\varphi]}{\cos\varphi}.$$

It is also convenient to expand the density in a series that convergence for any positive t,

$$g^{(\alpha)}(t) = \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n!} \Gamma(1+n\alpha) \sin(\pi n\alpha) t^{-n\alpha-1}, \tag{A5}$$

and to find the leasing term in the asymptotic expansion when $t\to 0$:

$$g^{(\alpha)}(t) = A t^{-\gamma} \exp\{-b t^\delta\}, \tag{A6}$$

where

$$A = \frac{\alpha^{1/2(1-\alpha)}}{\sqrt{2\pi(1-\alpha)}},$$

$$\gamma=\frac{1-\alpha/2}{1-\alpha},\quad b=\frac{\alpha-1}{\alpha^{\delta}},\quad \delta=\frac{\alpha}{\alpha-1}.$$

Formula (A6), which is exact at $\alpha=1/2$, provides a fairly accurate approximation in the middle of the interval $(0,1)$.

The expectation value and higher-order moments of these densities are infinite, but moments of order $\mu<\alpha$ (including negative orders) exist and are given by Eq. (A2).

Tables of the distribution functions $G^{(\alpha)}(t)=\int_0^t g^{(\alpha)}(\tau)\,d\tau$ and densities can be found in Refs. 28 and 29, respectively. Note that Holt and Crow[29] use form (A) of the stable distribution, which is related to form (B) of Bol'shov *et al.*[28] used here, $g_A^{(\alpha)}(t)=c(\alpha)g^{(\alpha)}(c(\alpha)t)$, where $c(\alpha)=[\cos(\pi\alpha/2)]^{1/\alpha}$. Moreover, the second parameter β of stable laws is defined differently by these researchers: according to Holt and Crow[29], $\beta=-1$ for a one-sided distribution on the positive semiaxis, while according to Bol'shov *et al.*[28] $\beta=1$ for the same case. It may have been these differences that prevented Holt and Crow[29] from comparing their results with the earlier results of Bol'shov *et al.*[28] (at least they attempted to do so at the end of their paper; see p. 163 in Ref. 29). In any case, our comparison of the results of the two groups, with full allowance for the foregoing, has shown that there is good agreement.

Stable laws play the same role in the summation of independent random quantities with infinite variances as the ordinary Gaussian law does in the case of finite variances. In particular, if independent random quantities $T_i\leqslant 0$ are distributed with a density $q_0(t)$ satisfying the condition (16) for $\alpha<1$, for large values of n the normalized sum

$$S_n=\sum_{i=1}^{n}\frac{T_i}{[nB\Gamma(1-\alpha)]^{1/\alpha}} \tag{A7}$$

is distributed with a density $g^{(\alpha)}(t)$. In other words, in the asymptotic region of large n, the distribution density $q_0^{(n)}(t)$ of the sum $\Sigma_{i=1}^{n}T_i$ has the form

$$q_0^{(n)}(t)\sim[nB^*]^{-1/\alpha}g^{(\alpha)}((nB^*)^{-1/\alpha}t), \tag{A8}$$

where $B^*=B\Gamma(1-\alpha)$.

*)E-mail: uchaikin@sv.uven.ru, uchaikin@themp.univ.simbirsk.su

---

[1] J.-P. Bouchaud and A. Georges, Phys. Rep. **195**, 127 (1990).
[2] M. B. Isichenko, Rev. Mod. Phys. **64**, 961 (1992).
[3] B. J. West and W. Deering, Phys. Rep. **246**, 1 (1994).
[4] V. M. Zolotarev, V. V. Uchaĭkin, and V. V. Saenko, Zh. Éksp. Teor. Fiz. **115**, 1411 (1999) [JETP **88**, 780 (1999)].
[5] R. R. Nigmatullin, Phys. Status Solidi B **123**, 739 (1984).
[6] W. Wyss, J. Math. Phys. **27** 2782 (1986).
[7] T. F. Nonnenmacher and D. J. F. Nonnenmacher, Acta Phys. Hung. **66**, 145 (1989).
[8] G. M. Zaslavsky, Physica D **76**, 110 (1994).
[9] K. B. Chukbar, Zh. Éksp. Teor. Fiz. **108**, 1875 (1995) [JETP **81**, 1025 (1995)].
[10] A. Compte, Phys. Rev. E **53**, 4191 (1996).
[11] W. R. Schneider and W. Wyss, J. Math. Phys. **30** 134 (1989).
[12] W. G. Glöckle and T. F. Nonnenmacher, J. Stat. Phys. **71**, 741 (1993).
[13] B. J. West, P. Grigolini, R. Metzler, T. F. Nonnenmacher, Phys. Rev. E **55**, 99 (1997).
[14] C. Fox, Trans. Am. Math. Soc. **98**, 395 (1961).
[15] P. Lévy, *Processes stochastiques et mouvement Brownien*, 2nd ed., Gauthier-Villars, Paris (1965).
[16] B. V. Gnedenko and A. N. Kolmogorov, *Limiting Distributions for Sums of Independent Random Variables*, Addison-Wesley, Reading, MA (1954).
[17] V. M. Zolotarev, *One-Dimensional Stable Distributions*, American Mathematical Society, Providence, R.I. (1986).
[18] W. R. Schneider, in *Stochastic Processes in Classical and Quantum Systems* (Proc. 1st Ascona-Como Int. Conf., held in Ascona, Ticino, Switzerland, June 24–29, 1985), S. Albeverio, G. Casati, and D. Merlini (eds.), Springer, Berlin (1986), p. 497.
[19] W. R. Schneider, in *Lecture Notes in Physics*, Vol. 1250, Springer, Berlin (1987), p. 269.
[20] W. Feller, *An Introduction to the Probability Theory and Its Applications*, Vol. 2, 2nd ed., Wiley, New York (1971).
[21] A. A. Kilbas, O. I. Marichev, and S. G. Samko, *Fractional Integrals and Derivatives: Theory and Applications*, Gordon & Breach, Philadelphia (1993).
[22] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York (1993).
[23] A. A. Borovkov, *A Course in Probability Theory* [in Russian], Nauka, Moscow (1972).
[24] Yu. L. Klimontovich, *Statistical Theory of Open Systems*, Kluwer, Dordrecht (1995).
[25] E. Lukacs, *Characteristic Functions*, Hafner, New York (1970).
[26] H. Scher and E. W. Montroll, Phys. Rev. E **12**, 2455 (1975).
[27] V. M. Zolotarev, in *Proc. (Trudy) V. A. Steklov Institute of Mathematics* [in Russian], Vol. 71 (1964), p. 47.
[28] L. N. Bol'shev, V. M. Zolotarev, E. S. Kedrova, and M. A. Rybinskaya, Teor. Veroyatn. Primen. **XV**, 309 (1970).
[29] D. R. Holt and E. L. Crow, J. Res. Natl. Bur. Stand., Sect. B **77B**, 143 (1973).

Translated by Eugene Yankovsky

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Feasibility of measuring the average sizes of diamagnetic domains by the μSR^2 method

A. S. Baturin*) and V. N. Gorelkin

Moscow Physicotechnical Institute, 141700 Dolgoprudnyi, Moscow Province, Russia

(Submitted 24 September 1998)
Zh. Éksp. Teor. Fiz. **115**, 2133–2142 (June 1999)

One of the simplest examples of possible application of the μSR^2 method for estimating the sizes of diamagnetic domains is analyzed in detail. The domains have been observed for the first time by means of the μSR method in beryllium [G. Solt, C. Baines, V. S. Egorov *et al.*, Hyperfine Interactions **104**, 257 (1997)]. Results are given from a computer simulation of a μSR^2 experiment to measure domain sizes in Be. An algorithm is described for processing the experimental results. It is graphically demonstrated that domain sizes can be estimated within the accelerator operating time allocated for an ordinary μSR experiment. © *1999 American Institute of Physics.* [S1063-7761(99)01706-0]

1. INTRODUCTION

The μSR method is customarily used to measure the probability density function of magnetic fields $w(b)$ and to estimate their correlation times. Data averaged over the volume or surface of a target are obtained in μSR experiments. Even if another coordinate of a decay point $\mathbf{r}$ is fixed during the compilation of a histogram, only the reconstruction of the density function $w(b,\mathbf{r})$ can be formally addressed. A muon stopping point or decay point can be fixed to within at best 10^{-3} cm. The actual errors to within which a coordinate is fixed are $10^{-1}-10^{-2}$ cm. Such distance scales clearly have a more important bearing on solid state physics than the correlation radii of magnetic inhomogeneities, because it is impossible to investigate space-time correlations by such a ''direct'' method.

However, it has been shown[1–3] that the temporal correlation function of ordinary μSR histograms compiled from several different regions of the investigated target already contain information about the space-time correlations of the magnetic fields. In principle, the temporal correlation functions of μSR histograms compiled from several counters set up at different angles relative to the initial direction of polarization of the muon spin contain information about the space-time correlations of the second-order, third-order, and in part fourth-order magnetic fields. In the same papers it has been shown that slow ($\tau_{\mathrm{cor}} \gg 10^{-6}$ s), long-wavelength space-time correlations with a characteristic scale $r_{\mathrm{cor}} \gg 10^{-6}$ cm can already be successfully investigated at the intensity levels of present-day accelerators. In other words, space-time correlations can be investigated in ranges that are essentially inaccessible by the neutron method. Such scales are important in the investigation of domain structure, spin-glass states, fluctuations at phase transition points (in particular, for the experimental measurement of critical indices), vortex structure in superconductors, and the segregation of samples into different fractions, etc.

Since the μSR signal is squared as it occurs in the correlation function, the given approach is called the μSR^2 method. We emphasize that this method can actually be used to measure fluctuations of the muon spin polarization vector. The principal condition for observing them is the detection of a sufficient number of paired events (the decay of at least two muons) from one correlation volume during compilation of the μSR histogram. It has been proposed in the above-cited papers that either strong pulsed beams of muons or an apparatus that fixes the muon decay point be used to obtain information of interest about magnets. Naturally, this approach requires new experimental procedures.

In this paper we have sought to focus attention on experiments of another kind, in which the acquisition of additional information does not require modification of the standard μSR procedure.

For example, the first application of the μSR method to observe the onset of diamagnetic domains in Be has been reported.[4] In the cited experiment the total volume of different types of domains has been observed to depend on the external magnetic field and the temperature. From these data, however, it is impossible to deduce the size of the diamagnetic domains.

Nonetheless, as will be shown below, by compiling several ordinary μSR histograms from various points of the target, i.e., in essence transcending the conventional procedure used in Ref. 4, it is possible to estimate the domain size after special processing of the experimental data.

Fluctuations of the domain size lead to fluctuations of the volumes of corresponding fractions in which the muon spin precesses at different frequencies. On the basis of real physical models, information about the domain size can be reconstructed from measurements of the fluctuations of the relative volumes of the fractions at a fixed total volume.

Analogous situations are encountered in the investigation of vortex structure in rigid superconductors, in the segregation of a magnet into zones of differing magnetization, and, finally, in paramagnetic states produced by a muon trapped by impurities or defects of a crystal lattice. Different fractions with their own distinct muon spin precession frequencies coexist in all these cases.

2. RELATIONSHIP OF THE μSR^2 CORRELATION FUNCTION TO FLUCTUATIONS OF THE RELATIVE FRACTIONS

The zone irradiated by the muon beam changes several times during the μSR^2 experiment so that the probability $W^{(a)}$ of detecting decays from different fractions also varies in an inhomogeneous target. In the ensuing discussion we define the ith pulse or block of information as a set of N_i detected decays of muons that occur in one zone and for which $W_i^{(a)}$ can be regarded as constant. Here $n_i(t)$ is the number of events in the ith pulse in the jth time channel, $j=[t/\Delta t]$ and width $\Delta t = T_{\max}/N_{\mathrm{ch}}$, where N_{ch} is the number of counter channels, and $T_{\max}$ is the decay time corresponding to the channel with the number N_{ch} ([. . .] denotes the integer part of a number). The expected value of the random variable $n_i(t)$ for a channel width t small in comparison with all characteristic times and for a counter having a small solid angle can be written as

$$n_i(t)=\frac{N_i\Delta t}{\gamma\tau_\mu}\exp\left(-\frac{t}{\tau_\mu}\right)\left[1-K_\alpha\sum_a W_i^{(a)}M_{\alpha\beta}^{(a)}(t)P_\beta(0)\right], \tag{1}$$

where K_α is a vector defining the direction toward the counter, and the normalization factor γ depends on the maximum ($T_{\max}$) and minimum ($T_{\min}$) muon decay times and can be deduced from the normalization condition

$$\int_{T_{\min}}^{T_{\max}} n_i(t)\,dt=1.$$

The behavior of the spin polarization $P_\beta(t)$ in the fraction of type a is specified by the tensor $M_{\alpha\beta}^{(a)}(t)$, and its contribution to the total polarization is proportional to $W_i^{(a)}$. Of course, $\Sigma_a W_i^{(a)}=1$. If the quantities $W_i^{(a)}$ vary during compilation of the statistics for a fixed target position, this time interval must be divided into subintervals in which the probabilities $W_i^{(a)}$ can be regarded as constant. The subintervals in this case must be much greater than the muon lifetime $\tau_\mu = 2.2\times 10^{-6}$ s. One situation in which this condition fails has been discussed in two papers,[2,5] where the potential use of pulsed muon beams in μSR^2 experiments has been investigated. In domain measurement experiments in ferromagnets and very likely in the investigation of diamagnetic domains in Be such conditions can be observed when the target position is fixed and the external fields are constant during the compilation of more than 10^6 events. Cases in which this requirement might not be satisfied will be discussed below in Secs. 6 and 7.

We denote the number of histograms, which is equal to the number of ''pulses,'' by N_{blocks}. The total number of single events detected in the entire experiment is then $N_{\mathrm{single}} = N_{\mathrm{blocks}}N_i$ for an identical number of events N_i in the pulse. We can construct an ordinary μSR histogram averaged over all pulses, $n(t)=\Sigma_i n_i(t)/N_{\mathrm{blocks}}$, and find the average probabilities $\langle W^{(a)}\rangle$.

In the usual processing of histograms from a physical model the tensors $M_{\alpha\beta}^{(a)}(t)$, which involve the precession frequencies, are specified, and the most probable precession amplitudes of the components of the various fractions are selected. The averages $\langle W^{(a)}\rangle$ are actually determined in this way.

In μSR^2 experiments we can obtain an estimate of the correlation function

$$B(\tau)=\sum_{i=1}^{N_{\mathrm{blocks}}}\int_{T_{\min}}^{T_{\max}/2}\frac{n_i(t+\tau)n_i(t)}{N_i(N_i-1)}dt - \int_{T_{\min}}^{T_{\max}/2}\frac{n(t+\tau)n(t)}{N_i^2}dt, \tag{2}$$

where $n(t)=\Sigma_i n_i(t)$. The expected value of the correlation function,

$$B_T(\tau)=\frac{\Delta t}{\gamma^2\tau_\mu}\exp\left(-\frac{\tau}{\tau_\mu}\right)\sum_{a,a'}\langle\Delta W^{(a)}\,\Delta W^{(a')}\rangle C^{(aa')}(\tau), \tag{3}$$

is proportional to the correlations of the deviations of the quantities $W^{(a)}$ from their averages. The function $C^{(aa')}(\tau)$ is uniquely determined by the behavior of the polarization $M_{\alpha\beta}^{(a)}(t)P_\beta(0)$ in the various fractions, which can be determined from the ordinary μSR signal:

$$C^{(aa')}(\tau)=\int_{T_{\min}}^{T_{\max}/2}\exp\left(-\frac{t}{\tau_\mu}\right)K_\alpha M_{\alpha\beta}^{(a)}(t+\tau)\times P_\beta(0)K_{\alpha'}M_{\alpha'\beta'}^{(a')}(t)P_{\beta'}(0)\,dt. \tag{4}$$

For sufficient statistics the least-squares method or Fourier analysis can be employed, in principle, to find $\langle\Delta W^{(a)}\Delta W^{(a')}\rangle$ from the correlation function. Usually the correlations $\langle\Delta W^{(a)}\Delta W^{(a')}\rangle$ simply reduce to fluctuations of the total volume of the fractions and are associated with the number of grains or domains in the muon stopping zone and with the scatter of their sizes. The relationship of the sizes of the domains to their number must be established on the basis of a physical model. For a stochastic scatter of domain sizes comparable in order of magnitude with the average or for random variations of the density of grains, their fluctuations are proportional to the volume V of the muon stopping zone and, accordingly, the quantity $\langle\Delta W^{(a)}\Delta W^{(a')}\rangle$ is proportional to $1/V$. The number of detected events is proportional to V, so that the error of measurement of the correlation function is essentially independent of the size of the beam and is in fact determined by its intensity.

Equation (4) provides a means of readily obtaining a theoretical estimate of the average correlation function when the number of detected paired events, the number of which in the experiment is $N_{\mathrm{double}} = N_{\mathrm{blocks}}N_i(N_i-1)$, is formally infinite. A finite number of events produces statistical scatter in the experimentally determined correlation function. A published estimate[2] of the number of paired events necessary for the measurement of correlation effects yields a gross overestimate, and depends not only on the total number of paired events, but also on the number of events in a single pulse.

When $\langle(\Delta W^{(a)})^2\rangle$ are small, a computer simulation of the μSR^2 experiment must be performed to obtain a realistic estimate of the error of measurement of the correlations and the optimum number of events in a single ''pulse.'' The total number of paired events is squared in N_i, and it is preferably made as large as possible. In some cases, however, simulation shows that it may be pointless to increase N_i. For narrow beams the intensity distribution over the profile of the beam must also be taken into account. A μSR^2 experiment simulation program has been implemented and, in particular, has provided a means for choosing the conditions under which it is possible to estimate satisfactorily the sizes of diamagnetic domains in beryllium.

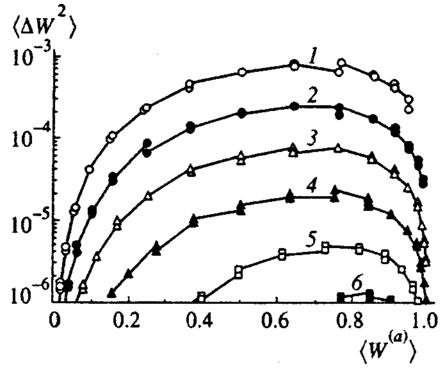

FIG. 1. Network of graphs for the determination of $R_{\max1}$ from experimental values of $\langle W^{(a)}\rangle$ and $\langle\Delta W^2\rangle$: (*1*) $R_{\max1}/\sigma=3.162\times10^{-3}$; (*2*) 10^{-3}; (*3*) 3.162×10^{-4}; (*4*) 10^{-4}; (*5*) 3.162×10^{-5}; (*6*) 10^{-5}.

3. APPLICATION OF THE μSR^2 METHOD FOR ESTIMATING THE SIZES OF DIAMAGNETIC DOMAINS IN BERYLLIUM

A straightforward model describing the structure of diamagnetic domains in Be does not exist; we therefore consider the simplest model of linear domains, which rests on the assumption that when an external magnetic field is applied, the target is partitioned into two types of domains, which differ in the strength of the magnetic field in the interior of the domain. The following simplifications are adopted in this model:

1) The field inside each domain is constant (the difference between the fields in adjacent domains for beryllium is 30–40 G for an applied external field of 27.4 G).

2) A domain comprises a strip of infinite length in the direction of the y axis and of constant width independent of y and z.

3) The width of a domain is distributed equiprobably in the interval from $R_{\min}$ to $R_{\max}$ ($R_{\min}$ and $R_{\max}$ differ for different types of domains, so we denote by $R_{\max1}$ and $R_{\max2}$ the maximum widths of domains of the first and second types, respectively; $R_{\min}$ can serve as the width of the domain walls, which are not explicitly taken into account in the given model). The width of the domain wall cannot be smaller in order of magnitude than the Larmor radius of the electron, which is equal to 10^{-4} cm in such fields; accordingly, the maximum domain width must be of the same order or greater than 10^{-3} cm.

The contribution of domains of each type to the μSR histogram depends, though not strongly, on the profile of the beam. Let us assume that the target is irradiated by a beam with a Gaussian intensity distribution characterized by the parameter σ. The input parameters of the model are therefore the minimum and maximum widths of the domains of each type and the parameter of the Gaussian intensity distribution σ.

Under the condition $R_{\min}\ll R_{\max}$ the problem is essentially independent of the minimum width. In this model we have constructed the network of graphs shown in Fig. 1, which can be used to determine $R_{\max1}$ directly from $\langle W^{(a)}\rangle$ and $\langle\Delta W^2\rangle$ and then to determine $R_{\max2}$ from Fig. 2.

The general equations (3) and (4) can be simplified considerably for diamagnetic domains in Be. If the beam polarization and the direction to the counter are perpendicular to the magnetic field in the domains, we can introduce a phase ϕ, which takes into account the relative position of the counter (vector $\mathbf{K}$) and the initial polarization $\mathbf{P}(0)$.

As a result, the tensor expression $K_\alpha M^{(a)}_{\alpha\beta}(t+\tau)P_\beta(0)$ can replace the scalar function $P_0\cos(\Omega_a t+\phi)$, where $P_0=|\mathbf{P}(0)|$, and Ω_a is the precession frequency in a domain of type a.

Taking into account the presence of only two fractions,

$$\langle(\Delta W^{(1)})^2\rangle=\langle(\Delta W^{(2)})^2\rangle=-\langle\Delta W^{(1)}\Delta W^{(2)}\rangle$$
$$=-\langle\Delta W^{(2)}\Delta W^{(1)}\rangle=\langle\Delta W^2\rangle,$$

it is a simple matter to obtain an analytic expression for the expected value of the correlation function from the general equations (3) and (4):

$$B_T(\tau)=\frac{\Delta\Omega\,\tau_\mu^2}{4}\left(\frac{\Delta\Omega\,\tau_\mu}{4}k_\alpha\cos\frac{\Delta\Omega\,\tau}{2}+k_\beta\sin\frac{\Delta\Omega\,\tau}{2}\right)\cos(\Omega\,\tau),\tag{5}$$

where

$$k_{\alpha,\beta}=F_{\alpha,\beta}\left(\frac{T_{\max}}{\tau_\mu}\right)-F_{\alpha,\beta}\left(\frac{2\,T_{\min}}{\tau_\mu}\right),$$

$$F_\alpha=-(x^2+2\,x+2)\,e^{-x},\quad F_\beta=-(x+1)\,e^{-x},$$

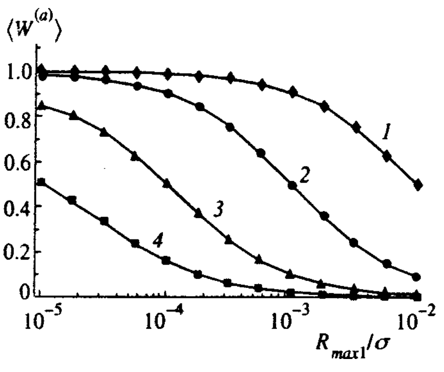

FIG. 2. Dependence of $\langle W^{(a)}\rangle$ on the maximum radius $R_{\max1}$ of domains of the first type for various maximum radii $R_{\max2}$ of domains of the second type: (*1*) $R_{\max2}/\sigma=10^{-2}$; (*2*) 10^{-3}; (*3*) 10^{-4}; (*4*) 10^{-5}.

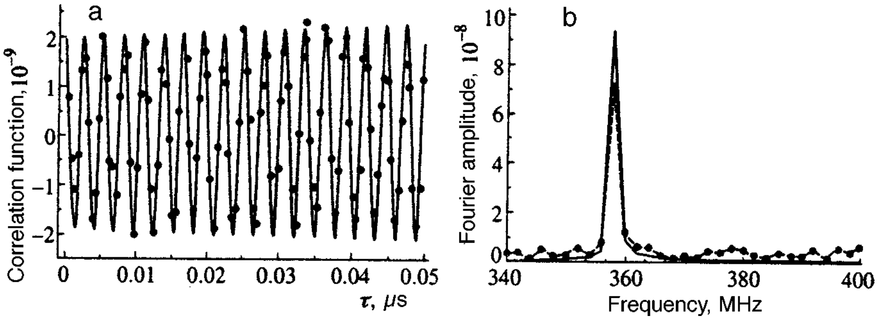

FIG. 3. Theoretical (solid curve) and modeled (dashed curve and points) correlation functions (a) and their Fourier transforms (b) for 2×10^8 (a) and 2×10^7 (b) events in the case $R_{\max}=10^{-3}$ cm, $R_{\min}=10^{-4}$ cm, $\sigma=0.1$ cm.

$\Delta\Omega=\Omega_1-\Omega_2$ is the difference between the precession frequencies in adjacent domains, and $\Omega=(\Omega_1+\Omega_2)/2$ is the average precession frequency. Equation (5) has been derived on the assumption that $\Delta\Omega T_{\max}\ll1$ and $\Omega T_{\max}\gg1$.

4. SIMULATION PROGRAM

The validity of the μSR^2 method for estimating domain sizes in Be has been tested by Monte Carlo computer simulation of a μSR experiment. The program simulating the μSR^2 experiment consists of three main parts. The first part is designed to calculate the distribution of the domains over the target in the above-described one-dimensional model of linear domains. The final result is the values of $W^{(a)}$.

The second part of the program is designed to simulate the compilation of μSR statistics. The only parameter carried over from the first part to the second part of the program is $W^{(1)}$, because $W^{(2)}=1-W^{(1)}$.

The simulation result is a μSR histogram, i.e, an array of integers, whose ith element is equal to the number of muons in the corresponding channel of the counter. The resulting array can be placed in a file for subsequent processing by other programs, displayed on the screen for comparison with a normalized theoretical curve, or transferred into the third module of the program, where the correlation function is calculated.

A cumulative computational algorithm is used, which permits the correlation function to be determined after the compilation of statistics in each pulse. This kind of algorithm does not require storage of the sequence of μSR histograms for all pulses.

5. DETERMINATION OF THE AMPLITUDE OF THE μSR^2 CORRELATION FUNCTION IN THE CASE OF LARGE STATISTICS

For a sufficiently large number of events the correlation function obtained by computer simulation of the μSR^2 experiment can be brought into agreement with the theoretical equation. For diamagnetic domains having an identical maximum width of 10^{-3} cm, irradiated by a Gaussian beam of muons with a half-width (standard deviation) of 0.1 cm in the computer simulation of 20 histograms compiled from various points of the target, we have $\langle\Delta W^2\rangle=1.3\times10^{-4}$ and $\langle W^{(a)}\rangle\approx0.5$. Figure 3a shows the correlation function obtained by simulation and its theoretical estimate calculated from Eqs. (2)–(4), in which the value of $\langle\Delta W^2\rangle$ known from the first simulation stage is substituted.

The ‘‘experimental’’ points essentially fit the theoretical curve with the correct value of $\langle\Delta W^2\rangle$, i.e., $\langle\Delta W^2\rangle$ is determined indirectly for $\langle W^{(a)}\rangle\approx0.5$, and from the curves in Figs. 1 and 2 we readily obtain the dimensionless ratio $R_{\max}/\sigma=10^{-2}$, which for an effective beam diameter of 0.1 cm gives an estimate of the diamagnetic domain width as 10^{-3} cm. Such an experiment conducted on a modern accelerator would require 10–20 h, but there is no need to compile 2×10^8 events, because a satisfactory estimate of the domain widths can be obtained with significantly smaller statistics. The amplitude of the correlation function at the frequency known from an ordinary μSR experiment can be discriminated for statistics at least an order of magnitude smaller. If the total number of events is 2×10^7 and the same number of ‘‘pulses’’ is used, the Fourier amplitude of the ‘‘experimental’’ correlation function does not deviate too far from the theoretical function (Fig. 3b), so that the domain widths can be estimated.

If the above-described effect is not observed experimentally, this means that the structure of the domains in the target changes during the compilation of a single histogram, corresponding to 10^6 events in the case discussed above.

The invariance of $W^{(a)}$ during the compilation of events in each block pulse signifies in this case that the domains must preserve their configuration during a time of the order of 10^2 s. This requirement can be relaxed by decreasing the number of single events in the block or by using a pulsed beam. However, an order-of-magnitude reduction of the number of events in one block leads to a hundredfold increase in the time to compile the same number of paired events, making it necessary to consider the possibility of obtaining information about the domain sizes for ‘‘small’’ statistics.

6. DISCRIMINATION OF THE AMPLITUDE OF THE μSR^2 CORRELATION FUNCTION IN THE PRESENCE OF SMALL STATISTICS

The statistics compiled in an individual block might not be sufficient to obtain a ‘‘good’’ correlation function, which can be extremely cluttered with noise, and its amplitude cannot possibly be determined from the Fourier transform. In this event a technique analogous to synchronous detection in radio engineering can be used to discriminate the useful signal.

The technique essentially entails convolution of the signal containing useful information with a certain *a priori*

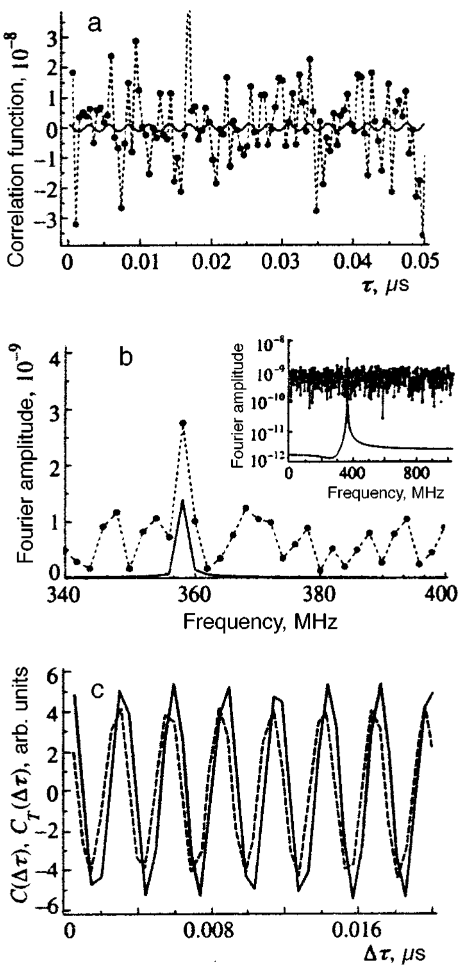

FIG. 4. Discrimination of the amplitude of the correlation function for statistics of 5000 blocks of 20 000 muons each for the case $R_{\max}=10^{-3}$ cm, $R_{\min}=10^{-4}$ cm, $\sigma=0.1$ cm. Comparison of the theoretical (solid curve) and modeled (dashed curve and points) correlation functions (a) and their Fourier transforms (b). Result of convolution with the theoretical correlation function (c): autocorrelation function of the kernel $C_T(\Delta\tau)$ (solid curve) and convolution of the signal $C(\Delta\tau)$ (dashed curve).

known theoretical function describing the useful signal. The convolution yields a periodic function, whose amplitude carries information about the amplitude of the useful signal, and the phase shift can be used to assess the influence of noise.

Figure 4 shows how the amplitude of the μSR^2 correlation function is reconstructed for the statistics in a single block pulse of 20 000 muons. The number of pulses in this case is 5000, so that the number of single events is of the same order as in the cases represented in Fig. 3, but now the number of paired events governing the error of reconstruction of the correlation function is much lower. The noise amplitude in the correlation function is greater than the useful signal and is comparable with it in the Fourier spectrum.

In the given situation the signal is an experimental estimate of the μSR^2 correlation function $B(\tau)$ described by Eq. (2). The expected value of the convolution $B_T(\tau)$ (3) is adopted as its kernel.

Comparing the autocorrelation function of the kernel

$$C_T(\Delta\tau)=\int B_T(\tau)B_T(\tau+\Delta\tau)d\tau \tag{6}$$

with the convolution of the signal

$$C(\Delta\tau)=\int B(\tau)B_T(\tau+\Delta\tau)d\tau, \tag{7}$$

both of which are shown in Fig. 4c, we can see that the curves all but coincide. They agree because the known value of $\langle\Delta W^2\rangle$ has been substituted into the kernel determined from Eq. (3). In a real experiment $\langle\Delta W^2\rangle$ is unknown and must be determined from the ratio of the amplitudes of the convolutions. In contrast with the correlation functions themselves, whose form in an experiment depends on the number of paired events and in the presence of small statistics resemble white noise, their convolutions are in the form of periodic functions, even when a useful signal is simply nonexistent. In this case the convolution of the theoretical and experimental correlation functions actually give the amplitude of random noise at the frequencies of the theoretical correlation function. The test of whether a useful signal is present is smallness of the phase shift. For a high noise amplitude the phase shift can be large and must undergo random variation during compilation of the statistics. In Fig. 4c the indicated phase shift is noticeable, but small, indicating a low noise amplitude at the frequency of the useful signal. This fact is not so obvious in the Fourier spectrum in Fig. 4b.

The small statistics of paired events ($5000\times 20\,000^2=2\times 10^{12}$) in the above example are the result of the comparatively small number of solitary events (20 000) in a single ''pulse,'' despite the large total number of single events, 2×10^8, and their compilation requires approximately 10 h of acceleration operating time. To estimate the domain sizes, the same result can be obtained in a few minutes provided that the structure of the domains does not change significantly during such time. For example, the statistics can be compiled eight times from 5×10^5. Of course, the random deviations of $W^{(a)}$ from the average over eight measurements do not give the exact variance in this case, but an order-of-magnitude estimate is clearly in order. For 5000 ''pulses'' the variance is determined very accurately, and the characteristic time of required invariance of the domain structure can be of the order of a second.

7. CONCLUSION

We have shown that large-scale magnetic inhomogeneities with characteristic dimensions of the order of 10^{-3} cm can be measured essentially without any modification of the conventional procedure used to perform μSR experiments on surface beams of muons, even when the magnetic field variations have the small amplitude typical of diamagnetic domains. The μSR^2 method reduces to a technique for processing ordinary μSR histograms compiled from several zones of the investigated target.

The limitations inherent in the conventional procedure — first of all, the annoyances incurred by having to change the position of the beam relative to the target and, secondly, the utilization of a small part of the beam, so that no more than 10^4 events can be detected — severely restrict the capabilities of the μSR^2 method. The application of pulsed beams or instruments that fix the muon decay or stopping point for μSR investigations makes it possible to surmount these difficulties and opens up expanded possibilities for the investigation of space-time correlations. Such instruments and beams have been in existence for some time now,[6,7] but so far have not been used for these applications. By fixing the decay coordinate even within 1-mm error limits it is possible to speed up the compilation of statistics by two orders of magnitude on a target of diameter 1 cm without having to change the position of the target, since a hundred of the μSR^2 experiments described in this article are actually performed simultaneously. It is entirely within the realm of possibility to reduce the coordinate-fixation error to the order of 10^{-2} cm, so that μSR^2 experiments can be successfully performed on existing Russian-made accelerators. The planned time structure of the Moscow Meson Factory beam would afford the capability of measuring the temporal and well as spatial correlations of magnetic fields with correlation times of $10^{-7}-10^{-5}$ s and, most auspiciously, slow correlations of 10^{-2} s or more. Such experiments were first proposed for this very accelerator many years ago.

The primary objective of the present article has been to visibly demonstrate, by means of a computer experiment on an object of interest in metal physics, the accessibility to experimental teams of this new direction in materials research by the μSR method.

*E-mail: baturin@quant.ffke.mipt.ru

---

[1] V. A. Gordeev and V. N. Gorelkin, Hyperfine Interact. **65**, 1129 (1990).

[2] V. A. Gordeev and V. N. Gorelkin, in *Physics of the Atomic Nucleus and Elementary Particles: Proceedings of the 24th Winter School* [in Russian], LIYaF, Leningrad (1989), p. 159.

[3] V. A. Gordeev and V. N. Gorelkin, *Muons and Pions in Matter: Proceedings of the Third International Symposium on Problems in the Interaction of Muons and Pions with Matter* [in Russian], Dubna (1995), p. 237.

[4] G. Solt, C. Baines, V. S. Egorov *et al.*, Hyperfine Interact. **104**, 257 (1997).

[5] V. A. Gordeev, V. N. Gorelkin, and N. E. Vtorova, Hyperfine Interact. **106**, 265 (1997).

[6] P. Podini and R. Tedeschi, Hyperfine Interact. **17–19**, 917 (1984).

[7] V. A. Gordeev, O. V. Savchenko, V. M. Abazov *et al.*, JETP Lett. **57**, 270 (1993).

Translated by James S. Wood

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Critical behavior of a degenerate ferromagnet in a uniaxial random field: exact results in a space of arbitrary dimension

D. É. Fel'dman*)

L. D. Landau Institute of Theoretical Physics, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 12 November 1998)
Zh. Éksp. Teor. Fiz. **115**, 2143–2159 (June 1999)

The critical behavior of the transverse (with respect to the field) magnetization component in classical degenerate magnets with only nearest-neighbors interaction in a uniaxial random magnetic field at zero temperature is found exactly. For a Gaussian distribution of the random field the asymptotic transverse magnetization in strong fields does not depend on the dimension of the space and is of the form $m_\perp \propto \ln h_0/h_0^2$, where h_0 is the width of the distribution. For a bimodal distribution, where only the field direction is random and the amplitude is fixed, the transverse magnetization behaves as $m_\perp \propto \exp(-\text{const}/(H_c-H)^{D/2})$, where H is the amplitude of the random field, D is the dimension of the space, and H_c is the critical field. © *1999 American Institute of Physics.* [S1063-7761(99)01806-5]

1. INTRODUCTION

Phase transitions in systems with random magnetic-field disorder[1–3] remain of interest even after 20 years. The Ising model in a random field was proposed at the end of the 1970s for describing doped antiferromagnets placed in a uniform external field.[4–6] Later, a variety of other systems with random-field disorder was discovered. These include various structurally disordered materials,[7–9] classical[10] and quantum[11–13] liquids and liquid crystals[14,15] in porous matrices, and vortex phases of dirty superconductors.[16,17]

Theoretical investigations of these systems have encountered serious difficulties. The early works[18–20] devoted to an Ising magnet in a random field led to an elegant result known as the Parisi–Sourlas reduction: the critical exponents of a dirty system in a space of dimension D are the same as in a pure $(D-2)$-dimensional magnet. This result implies that long-range order is absent in a three-dimensional ferromagnet in a random field. However, experiments with doped antiferromagnets in a uniform external field have not confirmed this prediction of the theory.[5] Several years later the existence of long-range order in a three-dimensional Ising model in a random field was proved rigorously.[21–23] Subsequent experiments[1] and numerical simulations[24] showed that even the critical exponents had been predicted incorrectly in all dimensions. The collapse of the theory is evidently due to the complicated structure of the energy landscape of a disordered system.[25,26] As a result, the perturbation theory used in the early works does not work. Since a complicated energy relief is typical for mean-field systems where replica symmetry is broken[27–30] it is tempting to use the concept of replica symmetry breaking in disordered systems with a finite interaction length. In recent years progress has been made along this avenue[31–35] by means of the variational method, taking account of the possibility of replica-symmetry breaking. However, an approximate variational approach is insufficient for solving the phase-transition problem.

In the absence of a reliable systematic method exact results that can be obtained for specific models become important. For systems with random fields and a finite interaction length, exact solutions can be obtained in two cases: one-dimensional models[36] and the spherical approximation.[37–40] Since there is no phase transition in the one-dimensional case, exact results on the critical behavior exist only for the spherical model, where the number of components of the order parameter is infinite.

In the present paper the critical behavior of classical degenerate magnets with a finite number of components (XY and Heisenberg magnets) and a finite interaction length in a uniaxial random magnetic field at zero temperature is studied. The critical behavior of the transverse (with respect to the field) magnetization component near the average amplitude of the random field for which the magnetization vanishes is studied. This problem is interesting because it can be solved exactly and mathematically rigorously in a space of arbitrary dimension. A brief report on this subject has been published in Ref. 41. The present paper contains a detailed exposition of the results presented in Ref. 41. The method is based on a proof of strict upper and lower estimates for the magnetization. These estimates can be made close enough so that the asymptotic behavior of the magnetization near a phase-transition point can be extracted from them.

Two types of distributions of the random field are studied in this paper: a Gaussian distribution and a bimodal distribution in which only the direction of the field is random and the absolute magnitude of the field is fixed. For a bimodal distribution of the random field the model can be used to describe dirty antiferromagnets in a uniform magnetic field. Quantum fluctuations are neglected below, i.e., the spins are assumed to be large.

The critical behavior of the present model at nonzero temperature has been studied in Refs. 42–44 by the renormalization group method. It is found that the critical behavior of the magnetization at zero temperature differs substantially from a power law, predicted by renormalization-group calculations for finite temperature, and in addition there is a large difference between the Gaussian and bimodal cases.

The model is described by the Hamiltonian

$$H=-J\sum_{\langle ij\rangle}\mathbf{S}_i\cdot\mathbf{S}_j-\sum_i H_iS_i^z, \tag{1}$$

where $\mathbf{S}_i$ are spin vectors of unit length, H_i are the random fields, and $\Sigma_{\langle ij\rangle}$ denotes summation over pairs of neighboring sites in a D-dimensional cubic lattice. Below the distance between the neighboring spins is assumed to be unity. Two types of distributions of the random field are considered:

1) a Gaussian distribution of width h_0

$$P(H_i)=\frac{1}{\sqrt{2\pi}h_0}\exp\left(-\frac{H_i^2}{2h_0^2}\right); \tag{2}$$

2) a bimodal distribution

$$P(H_i)=c\,\delta(H_i+H)+(1-c)\,\delta(H_i-H), \tag{3}$$

where c and $1-c$ are the probabilities of two opposite directions of the random field $\pm H$.

For a bimodal distribution of the random field the Hamiltonian (1) can be obtained by a gauge transformation of the Hamiltonian of a Mattis spin glass[45,46] in a uniform magnetic field H. It can be verified that the results found in the bimodal case also hold for antiferromagnets with random-bond disorder in a uniform external field.

The magnetization component transverse to the field, the case of interest below, arises by the same mechanism as the order parameter component transverse to a uniform external field in antiferromagnets.[47] For a Gaussian distribution of the random field the transverse magnetization $m_\perp$ is different from zero for arbitrary distribution width $h_0<\infty$. It is shown rigorously below that for large h_0 the transverse magnetization follows the law

$$m_\perp\propto\mathrm{const}\frac{\ln h_0}{h_0^2} \tag{4}$$

for a space of arbitrary dimension D. The exact formulation of the result consists of the inequalities (16) for the disorder- and volume-averaged magnetization in the thermodynamic limit. For a bimodal distribution of the random field the transverse magnetization vanishes in strong fields, $H>H_c=4DJ$. As will be shown below, for random field amplitude H close to the critical field H_c the transverse magnetization satisfies

$$m_\perp\propto\exp\left(-\frac{\mathrm{const}}{(H_c-H)^{D/2}}\right). \tag{5}$$

The rigorous result lies in the estimates (33).

This paper is organized as follows. The critical behavior of the transverse magnetization (4) and (5) is elucidated in Sec. 2 on the basis of qualitative considerations. A rigorous derivation of the upper and lower bounds for the magnetization with a Gaussian random field is given in Sec. 3. The bimodal case is analyzed in Sec. 4. Two lemmas which are used in so doing are given in an appendix. A discussion of the results is contained in Sec. 5.

Auxiliary statements required in what follows are formulated below.

In the first place, we note that in the ground state the transverse (with respect to the field) magnetization components of all spins point in the same direction. Indeed, the interaction energy with the random field does not depend on the direction of the transverse components. However, the exchange energy is minimum for spins oriented in the same direction. Hence it follows that if for all spins the transverse components are different from zero, then they all point in the same direction. The case where in the ground state some spins are parallel to the random field requires a separate analysis. In this case the system could break up into clusters in which the transverse magnetization points in different directions. However, it is easy to see that in the ground state the transverse component of the magnetization can vanish only for all spins simultaneously. To show this, rotate the transverse components of all spins so that they point in the same direction. This will not increase the energy of the system. Now take a spin S which is parallel to the field and has a neighbor with a nonzero transverse component. An effective field, including the random field and the Weiss field due to the exchange interaction with neighboring spins acts on the spin S. This effective field possesses a component transverse to the random field. The energy of the system will decrease if the spin S is rotated, without changing the direction of all other spins, so that it points in the same direction as the effective field. The assertion follows.

We have arrived at the conclusion that in the ground state all spins lie in the same plane. This explains the fact that the critical behavior (4) and (5) is the same for any number of spin components. For this reason, in what follows we can assume that the spins $\mathbf{S}_i$ are two-component and the spin component transverse to the field is positive:

$$S_i^z=\cos\phi_i,\quad S_i^x=\sin\phi_i\geqslant 0. \tag{6}$$

Now our problem is to calculate in the ground state the transverse component of any single spin averaged over the realizations of disorder

$$m_\perp=\overline{\sin\phi_i}. \tag{7}$$

If the ground state is degenerate, then any ground state can be chosen, since the estimates obtained below hold for all ground states. Of course, the average (7) is independent of the choice of spin only in the thermodynamic limit, where boundary effects are negligible. In the problem at hand the transition to the thermodynamic limit does not present any difficulties. For the transition to the thermodynamic limit in the Gaussian case it should be noted that the estimates derived in Sec. 3 hold for all spins located farther than 5 (the distance between neighboring spins is taken as 1) from the boundary of the system. In the bimodal case boundary effects are negligible for spins located at distances $L>\exp(1/(H_c-H))$ from the boundary.

We will use repeatedly two inequalities whose derivation is given below. In the ground state each spin $\mathbf{S}_k$ is directed along the effective field, which depends on the random field H_k and the directions of the neighboring spins $\mathbf{S}_l$. This is expressed by the equation

$$\sin\phi_k=\left(\cos\phi_k J\sum_l \sin\phi_l\right)\Big/\left(H_k+J\sum_l \cos\phi_l\right). \tag{8}$$

Hence for sufficiently large H_k

$$\sin\phi_k\leqslant\left(J\sum_l \sin\phi_l\right)/(|H_k|-2DJ). \tag{9}$$

Since $\sin\phi_l\leqslant 1$, it follows from the last inequality that

$$\sin\phi_k\leqslant 2DJ/(|H_k|-2DJ). \tag{10}$$

2. QUALITATIVE ESTIMATES

The purpose of this section is to present intuitive considerations elucidating the origin of the results obtained in this work. Strict proofs, which are contained in Secs. 3 and 4, are obtained by formalizing these guiding considerations.

2.1. *Gaussian case.*

For a Gaussian distribution of the random field the result (4) does not depend on the dimension of the space. For this reason, we start with the zero-dimensional case.

For this, we consider the two-spin model with the Hamiltonian

$$H=-J\mathbf{S}_1\cdot\mathbf{S}_2-H_1S_1^z-H_2S_2^z, \tag{11}$$

where H_1 and H_2 are random fields. We assume that

$$H_2\gg J, \tag{12}$$

and the field H_1 is such that

$$|H_1+J|<\epsilon J^2/|H_2|, \quad \epsilon\ll 1. \tag{13}$$

Let us compare the energies of two equilibrium states [i.e., such that the spins are directed along the local effective field and their components satisfy Eq. (8)]: the state A where $S_1^x=S_2^x=0$ and the state B where $S_1^x\approx 1$ and $S_2^x\approx J/|H_2|$. A simple calculation shows that the state B is deeper: the loss in the energy of interaction of the spin S_2 with the magnetic field is compensated by a gain in the interaction energy of the spin components transverse to the field. Therefore the transverse magnetization of the system in the ground state is of the order of 1. For a Gaussian distribution of random fields H_1 and H_2 with width $h_0\gg J$, the configuration of random fields (12) and (13) has a probability $P\sim\ln h_0/h_0^2$. However, if the inequality (13) is not satisfied, then the state A is deeper and the transverse magnetization is zero. The result (4) follows.

In the D-dimensional situation we note that for a large width h_0 of the Gaussian distribution rare regions where the random field is of the order of the exchange constant make the main contribution to the transverse magnetization. It can be shown that the transverse magnetization decreases exponentially away from such regions. Therefore the magnetization of the system is concentrated in rare clusters consisting of several spins. The clusters can be treated as zero-dimensional systems. As a result, the critical behavior (4) remains in a space of arbitrary dimension.

2.2. *Bimodal case.*

Just as in the Gaussian case, rare clusters make the main contribution to the magnetization, but now the magnitude of the magnetic field is fixed. Clusters where the magnetic field is zero on the average over the volume play the main role. In such clusters the tendency of the exchange interaction to align all spins in the same direction competes with the tendency of the spins to align along the magnetic field. A rigorous analysis shows that the transverse magnetization is concentrated in clusters where the random field is oriented upwards and downwards in a chessboard fashion.

It is easy to find the energy of such a cluster with volume $V=L^D$ (L is the size of a cluster) when the transverse component of all spins $m_\perp$ is the same. Neglecting surface effects, for small $m_\perp$ the energy can be expressed as

$$E=\text{const}\,V-V\frac{H_c-H}{2}m_\perp^2, \tag{14}$$

where H is the amplitude of the random field and $H_c=4DJ$. However, the state where the magnetization is nonuniform over the volume of a cluster is more advantageous, since the energy loss due to the boundary of the cluster is lower in this state. Assuming that the magnetization varies smoothly over the volume from zero at the boundary to a certain value m at the center of the cluster, we estimate the energy as

$$E\sim\text{const}\,V-V\frac{H_c-H}{2}m^2+JL^{D-2}m^2, \tag{15}$$

where L is the cluster size. A state with nonzero L is advantageous for $L>1/\sqrt{H_c-H}$, i.e., $V>(H_c-H)^{-D/2}$. Since large clusters are exponentially rare as a function of their volume, the result (5) follows.

The results (4) and (5) are rigorously substantiated below.

3. GAUSSIAN DISTRIBUTION

We shall estimate the disorder-averaged transverse component $\overline{\sin\phi_0}$ of the spin $\mathbf{S}_0$ at the lattice site S_0. Our aim is to prove the inequalities

$$\overline{\sin\phi_0}>C_1\frac{\ln h_0}{h_0^2}, \tag{16a}$$

$$\overline{\sin\phi_0}<C_2\frac{\ln h_0}{h_0^2}, \tag{16b}$$

where C_1 and C_2 are constants.

A derivation of Eq. (16a) is contained in the next section. In addition, in this section a convention is introduced for the notation used in subsequent sections.

In what follows const and the letters α, γ, δ, ϵ, Ω, C, C_1, and C_2 denote constants which depend on the dimension D of the space but not on the amplitude of the random field.

3.1. *Lower bound on the magnetization*

Let us consider a configuration of random fields for which

$$0<H_0+2DJ<\epsilon\sum_k\frac{J^2}{H_k}, \tag{17}$$

$$H_k>\Omega J, \quad H_p>\Omega J. \tag{18}$$

In these formulas H_0 is the field at the site S_0, H_k is the field at sites S_k neighboring S_0, H_p is the field acting on the spins $\mathbf{S}_p$ $(p\neq 0)$ which are neighbors of the spin $\mathbf{S}_0$ (and are different from $\mathbf{S}_0$), and the constants $\epsilon \ll 1$ and $\Omega \gg 1$. We shall use a similar convention for the spin indices k and p everywhere below without explanation. The bounds on the constants ϵ and Ω will be clear from the scheme of the proof presented in the present section. For large h_0 the probability of the configuration (17) and (18) is of the order of $\ln h_0/h_0^2$. For this reason, to prove Eq. (16a) it is sufficient to show that $\sin\,\phi_0 \sim 1$. To this end we compare the energy of two states: the state *A*, which has the lowest possible energy with the restriction $\sin\,\phi_0 < \epsilon$, and the state *B*, which will be described below and in which $\sin\phi_0 = 1$. In the state *B* all spins except $\mathbf{S}_0$ and its neighbors $\mathbf{S}_k$ point in the same direction as in the state *A*. The spins $\mathbf{S}_k$ are assumed to be directed along the local effective field:

$$\sin\phi_k = \left(\cos\phi_k J \sum_l \sin\phi_l\right) \Big/ \left(H_k + J\sum_l \cos\phi_l\right), \quad (19)$$

where Σ_l denotes summation over all neighbors of the spin $\mathbf{S}_k$, including the spin $\mathbf{S}_0$. Similar notation will be used below without explanation. The equation (19) also holds in the state *A*, since this equation expresses the condition that the energy has an extremum with respect to the angles ϕ_k.

As will be shown below, the state *B* is deeper. This is because the loss in the interaction energy of the spins $\mathbf{S}_k$ with the random field is less than the gain in the exchange interaction energy of their transverse components with spin $\mathbf{S}_0$.

An estimate of the form (9) for the quantities ϕ_k follows from Eq. (19). In the state *A* a similar estimate also holds for the spins $\mathbf{S}_p$ (after all, in the state *A* they point in the same direction as the local field). The inequality (10) for $\sin\,\phi_p$ follows from this estimate. It is easy to verify that the spins $\mathbf{S}_k$ and $\mathbf{S}_p$ make anglesless than $\pi/2$ with the corresponding random fields. In combination with Eqs. (9) and (10), this makes it possible to estimate $\cos\phi_{(k,p)}$.

Now it is easy to estimate the contribution depending on the spins $\mathbf{S}_0$ and $\mathbf{S}_k$ to the energy of the state *A*. To within a constant that is independent of the spins $\mathbf{S}_0$ and $\mathbf{S}_k$, this contribution is

$$\begin{aligned} E(\phi_0,\phi_k) = &-(H_0+2DJ)\cos\phi_0 + \sum_k (H_k + J\cos\phi_0) \\ &\times(1-\cos\phi_k) - J\sum_k \sin\phi_k \sin\phi_0 \\ &+ J\sum_k (1-\cos\phi_k)\sum_p \cos\phi_p \\ &- J\sum_k \sin\phi_k \sum_p \sin\phi_p . \end{aligned} \quad (20)$$

In this formula the notation $\Sigma_k(\dots)\Sigma_p$ implies summation over pairs of neighboring spins $\mathbf{S}_k$ and $\mathbf{S}_p$ only. We shall adhere to a similar convention below. The estimates enumerated above permit concluding that in the case *A*

$$E > -\alpha \sum \frac{J^2}{H_k}, \quad \alpha \ll 1. \quad (21)$$

The lower bound of the angles ϕ_k in the state *B* can be found by analogy with Eq. (9):

$$\sin\phi_k \geq \frac{J\cos\phi_k}{H_k + (2D-1)J}. \quad (22)$$

The estimates obtained above make it possible to estimate the energy (20) in the case *B*. It is found that

$$E < -b\sum_k \frac{J^2}{H_k}, \quad b \approx 1/2. \quad (23)$$

We see that the state *B* is deeper. Hence we conclude that the ground state satisfies $\sin\phi_0 > \epsilon$ and we extract the required estimate (16a).

The proof that the transverse magnetization is different from zero for any width of the Gaussian distribution is completely analogous.

3.2. *Upper bound of the magnetization*

We now proceed to the derivation of Eq. (16b). If the magnetic field H_0 at the site S_0 is high compared with the exchange constant, then the spin $\mathbf{S}_0$ is oriented almost in the direction of the field. For this reason, only spins located in a weak random field contribute to the transverse magnetization. This contribution depends on the field acting on a given spin and on the fields at neighboring sites. However, this contribution does not depend on the random field acting on distant spins. The reason is that the transverse spin components decrease rapidly away from the weak-field region. Therefore the correlations between distant spins are weak. It turns out that in the limit of large h_0 the contribution of the spin $\mathbf{S}_0$ to the transverse magnetization is determined mainly by the configuration of the random field inside a cube Γ with edge length 9 (the distance between neighboring spins is taken to be 1) and centered at the site S_0.

Four possibilities for the distribution of the random field in the cube Γ are considered below:

1) the random field $H_i < \Omega J$, where the constant satisfies $\Omega \gg 1$, in at least two points of the cube Γ;

2) $H_i > \Omega J$ holds at all points of the cube;

3) the random field satisfies $H_i < \Omega J$ at one point of the cube, which is not the center of the cube S_0;

4) the random field satisfies $H_0 < \Omega J$ and $H_i > \Omega J$ holds elsewhere in the cube.

In all cases our problem is to estimate the contribution of the corresponding field configurations to $\overline{\sin\phi_0}$.

1) This case is very simple. It occurs with probability of order $1/h_0^2$, and $\sin\phi_0 \leq 1$. Therefore the corresponding contribution m_1 to the transverse magnetization is

$$m_1 < \mathrm{const}/h_0^2. \quad (24)$$

2) In this case the spin $\mathbf{S}_0$ behaves almost as if the system were in a strong uniform external field. Therefore the transverse magnetization is small. The scheme of the arguments in the second case is as follows: applying Eq. (9) to the spin $\mathbf{S}_0$ we obtain

$$\sin\phi_0 < \left(J\sum_k \sin\phi_k\right)/(|H_0|-2DJ), \tag{25}$$

where the notation was explained in the preceding section. Applying Eq. (9) to the angles ϕ_k and Eq. (10) to the angles ϕ_p, we find

$$\sin\phi_0\left(1-\frac{1}{\Omega-2D}\,\frac{2D}{\Omega-2D}\right) < \frac{J}{|H_0|-2DJ} \times\sum_k \frac{J}{|H_k|-2DJ}\sum_p \frac{2DJ}{|H_p|-2DJ}. \tag{26}$$

It remains to integrate the inequality (26) over the distribution function of the random field. As a result,

$$m_2 < \mathrm{const}\left(\frac{\ln h_0}{h_0}\right)^3 \tag{27}$$

is obtained for the contribution m_2 to the transverse magnetization.

3) There are two variants here:

a) the random fields $H_i > \Omega J$ at all points of the cube Δ with edge length 5 and centered at the site S_0;

b) the random field $H_1 < \Omega J$ at some point S_1 of the cube Δ.

The case 3a) can be examined identically to case 2). This is because information about random fields outside the cube Δ was not used to analyze case 2). The inequality

$$m_{3a} < \mathrm{const}\left(\frac{\ln h_0}{h_0}\right)^3 \tag{28}$$

is obtained for the contribution to the transverse magnetization.

In the case 3b) once again we argue similarly to case 2). Here a difference from case 2) arises. It is due to the fact that now Eq. (9) cannot be used to estimate the transverse component $\sin\phi_1$ of the spin at the site S_1. Therefore the structure of Eq. (26) changes. Now the indices p and k cannot assume the value 1, and an additional term proportional to $\sin\phi_1$ appears. Its upper bound is $\mathrm{const}\cdot\sin\phi_1$. Integrating over the distribution function of the random field gives the following estimate of the contribution to the transverse magnetization:

$$m_{3\mathrm{b}} < \mathrm{const}\left(\frac{\ln h_0}{h_0}\right)^3 + \mathrm{const}(p\,\overline{\sin\phi_1}), \tag{29}$$

where the overbar denotes an average over the realizations of the disorder for which $H_1 < \Omega J$ and p is the probability of the configuration b). The expression for $p\,\overline{\sin\phi_1}$ can be estimated by the same method that will be used to analyze the case 4).

4) Once again there are two variants:

$$\text{a)}\quad \left|H_0+J\sum_k \operatorname{sign}H_k\right| < C\sum_k \frac{J^2}{|H_k|},$$

$$\text{b)}\quad \left|H_0+J\sum_k \operatorname{sign}H_k\right| > C\sum_k \frac{J^2}{|H_k|},$$

where the notation is the same as in Eq. (17) and the constant satisfies $C \gg 1$.

The probability of the configuration 4a) is of the order of $\ln h_0/h_0^2$. This is sufficient to obtain the estimate

$$m_{4\mathrm{a}} < \mathrm{const}\,\frac{\ln h_0}{h_0^2}. \tag{30}$$

In the case 4b) the z component $H_0+J\Sigma_k\cos\phi_k$ of the effective field at the site S_0 is essentially identical to $H_0+J\Sigma_k\,\operatorname{sign}H_k$. This can be verified by noting that the differences of the z components of the spins $\mathbf{S}_k$ from $\operatorname{sign}H_k$ are small, as $1/H_k^2$ (or even stronger). The rest of the proof is the same as in case 2). The only difference will be that in an estimate of the type (9) for the angle ϕ_0 the quantity $|H_0+J\Sigma_k\operatorname{sign}H_k|/2$ must be substituted for $|H_0|-2DJ$. The final result for the contribution to the transverse magnetization is

$$m_{4\mathrm{b}} < \mathrm{const}\left(\frac{\ln h_0}{h_0}\right)^3. \tag{31}$$

Using the inequalities (24) and (27)–(31) we obtain the estimate (16b). This completes the analysis of the Gaussian case.

4. BIMODAL DISTRIBUTION

We shall call a site of this set of lattice a connected region if any lattice site can be reached from any other lattice site by moving along lattice edges while always remaining inside the region. In the bimodal case connected regions where the random field is oriented upwards and downwards in a chessboard fashion make the main contribution to the magnetization. We shall call chessboard regions connected regions if the random field at each point of the region points in a direction opposite to the field at all neighboring points of the lattice. In chessboard regions the Hamiltonian (1) can be obtained from the Hamiltonian of an antiferromagnet in a uniform external field by inverting the spins $\mathbf{S}_i \to -\mathbf{S}_i$ in one of the two chessboard sublattices. Therefore in sufficiently large chessboard regions a transverse magnetization appears for the same critical field as in an antiferromagnet. In our model this field is $H_c = 4DJ$, where D is the dimension of the space.

It is easy to show that the transverse magnetization vanishes for $H > H_c$. For this, the inequality (9) must be applied several times in succession. As a result, we obtain for arbitrary spin, given by the angle ϕ,

$$\sin\phi < \left(\frac{2DJ}{H-2DJ}\right)^M, \tag{32}$$

where M is an arbitrarily large number. It follows that $\sin\phi = 0$ for $H > 4DJ$.

Our goal is to prove the two inequalities

$$m_\perp > \exp\left(-\frac{C_1}{(H_c-H)^{D/2}}\right), \tag{33a}$$

$$m_{\perp}<\exp\left(-\frac{C_2}{(H_c-H)^{D/2}}\right), \tag{33b}$$

where C_1 and C_2 are constants. The inequality (33b) will be proved below for $D\neq 2$. The weaker inequality

$$m_{\perp}<\exp\left(-\frac{\text{const}}{(H_c-H)\ln(1/(H_c-H))}\right) \tag{33c}$$

will be proved in the two-dimensional case.

It would be interesting to learn whether in the two-dimensional case logarithmic corrections are indeed present or the inequality (33b) always holds.

The lower bound can be easily derived. For this it is necessary to consider a chessboard region in the shape of a cube of size $L=\Omega/\sqrt{1-H/H_c}$, $\Omega\gg 1$. The concentration of such regions is exponentially low $\exp(-\text{const}\,L^D)$. In a pure antiferromagnet the magnetization near H_c behaves according to the mean-field law $\sin\phi\sim\sqrt{1-H/H_c}$. Comparing different contributions to the energy shows that in a disordered system the state with $\sin\phi\sim\sqrt{1-H/H_c}$ at the center of the cube is indeed energetically favorable. Hence obtains the required bound. The formal proof can be constructed by comparing the energy of the deepest state A from among the states for which $\phi<(1-H/H_c)^4$ everywhere inside the cube with the energy of the state B for which all spins outside and on the boundary of the cube point in the same direction as in the state A, while inside the cube $\sin\,\phi$ varies smoothly, reaching at the center the maximum value $\epsilon\sqrt{1-H/H_c}$, where ϵ is a small constant which does not depend on H.

The upper bounds are derived below. To make the exposition clearer we consider the case $c\ll 1$ first and then proceed to the general case. The simplification for small c arises because in this limit large chessboard clusters are exponentially rare. The exposition of the general case will be based on the case $c\ll 1$.

4.1. *Upper bound of the magnetization in the presence of weak disorder*

For small c the density of chessboard clusters with volume V is less than $\exp(-\text{const V})$. The argument proving this fact is based on an estimate of the number of connected regions with volume V that contain a given point. It is given in Appendix 1. It is obvious that the contribution from chessboard clusters with volume $V>\text{const}/(H_c-H)^{D/2}$ to the magnetization is exponentially small.

Outside chessboard clusters the transverse magnetization decreases rapidly with increasing distance to the nearest chessboard cluster. If H is close to H_c (this is the case of interest to us), then an analysis ofthe possible directions of the effective field shows that the magnetization at each site makes with the random field an angle less than $\pi/2$. This makes it possible to obtain from Eq. (8) the following inequality outside chessboard clusters:

$$\sin\phi_k<\frac{J\Sigma_l\sin\phi_l}{H-(2D-1)J}. \tag{34}$$

Applying the inequality (34) several times successively we obtain for the transverse component of the spin $\mathbf{S}_k$

$$\sin\phi_k<\left[\frac{2DJ}{H-(2D-1)J}\right]^S, \tag{35}$$

where S is the distance from the spin $\mathbf{S}_k$ to the nearest chessboard region. The derivation of Eq. (33b) requires an estimate with the structure of Eq. (35) but with a larger exponent.

To obtain such an estimate we consider a chessboard region with volume $V<V_c$, where

$$V_c=\alpha(H_c-H)^{-D/2},\quad D\neq 2, \tag{36a}$$

$$V_c=\alpha\frac{1}{(H_c-H)\ln(1/(H_c-H))},\quad D=2, \tag{36b}$$

$$\alpha\ll 1.$$

The magnetization in this region can be estimated in terms of the magnetization at the sites closest to it and lying outside the region, i.e., in terms of the magnetization at the boundary of this region. A lemma according to which inside the described region the transverse magnetization satisfies the inequality

$$\sin\phi<(1+\epsilon)(\sin\phi_b)_{\max},\quad \epsilon<\text{const}\,\alpha^{2/D}\ll 1, \tag{37}$$

where $(\sin\phi_b)_{\max}$ is the maximum magnetization at theboundary of the region, is proved in Appendix 2. This lemma permits writing instead of Eq. (34) (for spins located outside chessboard clusters and not on the boundary of the chessboard clusters of volume greater than V_c) the inequality

$$\sin\phi_k<\left(J(1+\epsilon)\sum_{u=1}^{2D}\sin\phi_u\right)/(H-(2D-1)J), \tag{38}$$

where the $\sin\phi_u$ denote the transverse components of certain spins located outside the chessboard regions and at a distance less than V_c+1 from the site S_k. Repeated application of the inequality (38) gives the estimate

$$\sin\phi_k<\left[\frac{2DJ(1+\epsilon)}{H-(2D-1)J}\right]^{S'/(V_c+1)}, \tag{39}$$

where V_c is the critical volume (36) and S' is the distance from the spin $\mathbf{S}_k$ to the nearest chessboard region with volume $V>V_c$. The inequality (39) refers to spins outside the chessboard clusters. Combining it with the inequality (37) we obtain a similar estimate for spins from small chessboard regions. Since large chessboard regions are exponentially rare, this makes it possible to derive Eqs. (33b and 33c).

4.2. *Upper bound of the magnetization in the presence of strong disorder*

If c is not small, then one cannot assert that large chessboard clusters are exponentially rare. Conversely, above the percolation threshold even infinite chessboard clusters can arise in the system. However, the estimates of the transverse magnetization (33b and 33c) remain valid. The point is that almost every chessboard cluster is ''larded'' with nonchessboard regions. Although the appearance of transverse magnetization inside a cluster is energetically favorable, the loss in energy due to ''nonchessboard impregnations'' can be greater. As a result, only the large chessboard clusters that are free of ''nonchessboard impregnations'' contribute to the magnetization, and the number of such clusters is exponentially small.

We now proceed to the proof. Let us cut the system into cubic cells with edges of large length A which is independent of H. The probability of such a cell being a chessboard cluster is small. In what follows, we shall retain the term ''chessboard cluster'' only for chessboard regions consisting entirely of one or several cells. With this definition large chessboard clusters will once again be exponentially rare. Therefore their contribution to the magnetization is once again exponentially small. Our next goal is to derive the analog of Eq. (39) for spins located outside chessboard clusters. Just as in the preceding subsection, the inequality (37) makes it possible to obtain an analogous formula for spins from small chessboard clusters. This will make it possible to give a complete proof.

To derive the analog of Eq. (39) we shall assume that the distance from the spin considered to the closest chessboard cluster with volume greater than V_c (36) is greater than $AD(V_c+1)$. The fraction of spins for which this condition is not satisfied is exponentially small as a function of V_c, so that their contribution to the transverse magnetization is certainly also exponentially small.

All further arguments are valid if the amplitude of the random field is sufficiently close to H_c, $H_c-H<B\ll H_c$, and the constant α (36) is sufficiently small (but independent of H). How small α and B must be will be evident from the proof given below.

Outside chessboard clusters an inequality similar to (34) (but weaker because of the new definition of a chessboard cluster) follows from Eq. (8). This inequality is

$$\sin\phi_k\leqslant\left(\sum_l \sin\phi_l\right)/2D(1-(H_c-H)/(2DJ)). \tag{40}$$

Each of our cells with edge length A, if the cell is not a chessboard cluster (this is the only case in which we are now interested), will contain a spin $\mathbf{S}_1$ for which the random field at one of the neighboring sites points in the same direction as at this spin. For this spin the inequality (34) holds. For arbitrary spin $\mathbf{S}_k$ from a nonchessboard cell we can write

$$\sin\phi_k\leqslant\frac{T_k}{2D}\left(\sum_l \sin\phi_l\right), \tag{41}$$

where $T_k/(2D)\Sigma_l \sin\phi_l$ denotes the right-hand side of the inequality (40) for spins at which the random fields at all neighboring sites point in a direction opposite to the direction at the spins themselves and the right-hand side of the inequality (34) for all other spins.

Using the lemma proved in Appendix 2 we derive from the inequality (41) the relation

$$\sin\phi_k\leqslant\frac{R_k}{2D}\left(\sum_{u=1}^{2D} \sin\phi_u\right), \tag{42}$$

where $R_k=(1+\epsilon)T_k$ and the spins $\mathbf{S}_u$ lie outside the chessboard clusters and are located at distances from the site S_k no greater than (V_c+1). The constant ϵ can be made as small as desired (independently of H) by appropriately choosing the constant α in the definition of the critical volume (36). Therefore, for a random field with amplitude sufficiently close to H_c it is possible to attain

$$R_k<1+\gamma, \tag{43}$$

where γ is an H-independent constant, which by choosing appropriate values of α and B can be made as small as desired. The estimate

$$R_1<\frac{1+\epsilon}{1+1/(4D)} \tag{44}$$

which is stronger than the inequality (43), holds for the above-mentioned spin S_1 with the property that the random field acting on it points in the same direction as the field at one of the neighboring sites .

Next we apply the inequality (42) DA times successively. The result is an estimate for the transverse magnetization $\sin\phi_k$

$$\sin\phi_k\leqslant\frac{1}{(2D)^{DA}}\sum_{w=1}^{(2D)^{DA}} P_w\sin\phi_w, \tag{45}$$

where all spins $\mathbf{S}_w$ lie at distances no greater than $DA(V_c+1)$ from the site S_k. Each factor P_w in Eq. (45) is a product of DA factors of the form R_i. The inequality (45) can be put into the form

$$\sin\phi_k\leqslant\frac{1}{(2D)^{DA}}\sum_w P_w(\sin\phi_w)_{\max}, \tag{46}$$

where $(\sin\phi_w)_{\max}$ is the largest of the quantities $\sin\phi_w$.

We now recall the inequalities (43) and (44). They make it possible to estimate P_w. It is important that since we have worked with a nonchessboard cell, a factor equal to the right-hand side of Eq. (44) appears in the estimate of one of the coefficients P_w. For sufficiently small γ this makes it possible to show that

$$\frac{1}{(2D)^{DA}}\sum_w P_w<1-\delta, \tag{47}$$

where δ is a small positive constant that does not depend on H. As a result we obtain from Eq. (46)

$$\sin\phi_k\leqslant(1-\delta)(\sin\phi_w)_{\max}. \tag{48}$$

Applying the latter formula repeatedly we obtain an estimate of the type (39). To complete the proof it remains to give an argument similar to that contained in the preceding section.

5. DISCUSSION

In the system studied above, rare regions give rise to ordering. This is reminiscent of a Griffiths phase.[48] However, our problem contains long-range order and spontaneous symmetry breaking. The appearance of long-range order is due to a weak ferromagnetic interaction of rare ordered clusters. Such an interaction is also present in other disordered systems undergoing a Griffiths transition, but even weak thermal fluctuations can destroy the order. A specific feature of the problem studied in the present paper is that there are no thermal fluctuations.

The mean-field approximation[42] ignores the contributions from rare regions and therefore leads to incorrect conclusions. In the Gaussian case the mean-field approximation predicts an incorrect exponential dependence of the magnetization on the field in the strong-field limit. In the bimodal case the mean-field approximation gives incorrect answers for the critical behavior and for the critical field in which a phase transition occurs.

The results obtained above are applicable to certain other systems. Specifically, the critical behavior (5) occurs in antiferromagnets with random-bond disorder in a uniform field.

The model studied above neglects quantum fluctuations. Strong quantum fluctuations can completely change the behavior of the system. For weak fluctuations there are two critical regimes: the classical regime studied in the present paper and a regime closer to a phase transition point—the quantum regime with a hypothetically slow power-law field dependence of the magnetization.

I thank V. S. Dotsenko, A. B. Kashuba, A. Yu. Kitaev, V. I. Marchenko, I. M. Suslov, and M. V. Feĭgel'man for helpful discussions. This work was supported by the Russian Fund for Fundamental Research (Project 96-02-18985) and the Program for the Support of Leading Scientific Schools (Project 96-15-96756).

APPENDIX 1

Here the upper bound on the number of connected regions with volume V which consist of the sites of a D-dimensional cubic lattice and contain a given site A is found. It is shown that this bound is an exponential function of the volume. Hence it follows for weak disorder that thedensity of chessboard regions with volume V is exponentially small as a function of V.

The number of broken lines of length $(2V-2)$, consisting of the edges of the lattice and starting at the point A, is less than $(2D)^{(2V-2)}$. Therefore it is sufficient to show that the number of regions of interest to us does not exceed the number of broken lines described above.

For this we note that each connected region can be put into correspondence in at least one way with a tree graph consisting of the edges of a lattice lying in the region and containing all of its sites. The length of the tree graph is $V-1$. It remains to note that a tree graph can be put into correspondence with a closed broken line passing through each edge of the graph twice.

APPENDIX 2

The lemma used in Sec. 4 is proved here.

Lemma

Consider a chessboard region with volume $V<V_c$, where V_c is defined by Eqs. (36). The transverse magnetization inside such a region satisfies the inequality (37).

Proof

Inside a chessboard region the inequality

$$\Delta \sin \phi \geqslant -\frac{H_c-H}{J}(\sin \phi)_{\max}, \tag{49}$$

where

$$\Delta \sin \phi = \sum_{\text{neighbors}} \sin \phi_k - 2D \sin \phi \tag{50}$$

is the lattice Laplacian, summation extends over nearest neighbors, and $(\sin \phi)_{\max}$ is the maximum value of the transverse magnetization inside the chessboard region, follows from Eqs. (8). To prove the inequality (49) Eq. (8) must be rewritten in the form

$$\sin \phi_k \leqslant \left(J \sum_l \sin \phi_k \right) / (H-2DJ),$$

and then transformed identically to the form

$$\Delta \sin \phi \geqslant -\frac{H_c-H}{J} \sin \phi.$$

The solutions of the inequalities (49) are majorized by solutions of the equation

$$\Delta u = -\frac{H_c-H}{J}(\sin \phi)_{\max} \tag{51}$$

with the same boundaries conditions as the inequality. We construct the solution of the last equation as follows. First, we find a particular solution u_1. It can be taken in the form of a convolution of the right-hand side of the equation and the Green's function of the lattice Laplace equation. It can be verified that because α (36) is small this solution is small compared with $(\sin \phi)_{\max}$

$$|u_1| < \beta (\sin \phi)_{\max}, \quad \beta = \text{const}\, \alpha^{2/D} \ll 1. \tag{52}$$

For $D=1$ and 2 there is no difficulty in establishing the estimate (52). The derivation for $D>2$ requires for a fixed volume of the region that $u_1(\mathbf{r})$ be maximum for a region in the shape of a sphere centered at the point $\mathbf{r}$. Next we find the solution u_2 of the equation with a zero right-hand side such that $u=u_1+u_2$ satisfies the boundary conditions. By virtue of the principal of the maximum,

$$u_2 < (\sin \phi_b)_{\max} + \beta (\sin \phi)_{\max}. \tag{53}$$

Recalling that u is the majorant of the desired function, and using the inequalities (52) and (53), we obtain the required estimate for $\sin \phi$.

*)E-mail: feldman@itp.ac.ru; dima@feb.uupc.chg.ru

---

[1] D. P. Belanger and A. P. Young, J. Magn. Magn. Mater. **100**, 272 (1991).
[2] V. S. Dotsenko, Usp. Fiz. Nauk **165**, 481 (1995).
[3] T. Nattermann, E-prints archive cond-mat/9705295 (1997).
[4] S. Fishman and A. Aharony, J. Phys. C **12**, L729 (1979).
[5] D. P. Belanger, A. R. King, V. Jaccarino, and J. L. Cardy, Phys. Rev. B **28**, 2522 (1983).
[6] Ch. Binek, S. Kuttler, and W. Kleemann, Phys. Rev. Lett. **75**, 2412 (1995).
[7] J. Villain, J. Physique Lett. **43**, 808 (1982).
[8] C. Bostoen and K. H. Michel, Z. Phys. B **71**, 369 (1988).

[9] T. Nattermann, Ferroelectrics **104**, 171 (1990).
[10] P. G. de Gennes, J. Phys. Chem. **88**, 6469 (1984).
[11] J. V. Porto and J. M. Parpia, Phys. Rev. Lett. **74**, 4667 (1995).
[12] J. Hook, Bull. Am. Phys. Soc. **42**, 799 (1997).
[13] K. Matsumoto, J. V. Porto, L. Pollak *et al.*, Phys. Rev. Lett. **79**, 253 (1997).
[14] N. A. Clark, T. Bellini, R. M. Malzbender *et al.*, Phys. Rev. Lett. **71**, 3505 (1993).
[15] H. Haga and C. W. Carland, Liq. Cryst. **22**, 275 (1997).
[16] G. Blatter, M. V. Feigelman, V. B. Geshkenbein, and A. I. Larkin, Rev. Mod. Phys. **66**, 1125 (1994).
[17] T. Giamarchi and P. Le Doussal, E-prints archive cond-mat/9705096 (1997).
[18] A. Aharony, Y. Imry, and S.-K. Ma, Phys. Rev. Lett. **37**, 1364 (1976).
[19] A. P. Young, J. Phys. C **10**, L257 (1977).
[20] G. Parisi and N. Sourlas, Phys. Rev. Lett. **43**, 744 (1979).
[21] J. Imbrie, Phys. Rev. Lett. **53**, 1747 (1984).
[22] J. Imbrie, Commun. Math. Phys. **98**, 145 (1985).
[23] J. Bricmont and A. Kupiainen, Phys. Rev. Lett. **59**, 817 (1987).
[24] M. Gofman, J. Adler, A. Aharony *et al.*, Phys. Rev. B **53**, 6362 (1996).
[25] G. Parisi, in *Proceedings of Les Houches 1982,Session* XXXIX, edited by J. B. Zuber and R. Stora (North Holland, Amsterdam, 1984).
[26] J. Villain, J. de Phys. **46**, 1843 (1985).
[27] M. Mezard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and Beyond* (World Scientific, Singapore, 1987).
[28] K. H. Fischer and J. A. Hertz, *Spin Glasses* (Cambridge University Press, Cambridge, 1993).
[29] G. Parisi, J. Phys. A **13**, 1887 (1980).
[30] V. S. Dotsenko, Usp. Fiz. Nauk **163**(6), 1 (1993).
[31] M. Mezard and A. P. Young, Europhys. Lett. **18**, 653 (1992).
[32] S. E. Korshunov, Phys. Rev. B **48**, 3969 (1993).
[33] M. Mezard and R. Monasson, Phys. Rev. B **50**, 7199 (1994).
[34] T. Giamarchi and P. Le Doussal, Phys. Rev. B **52**, 1242 (1995).
[35] D. E. Feldman, Phys. Rev. B **56**, 3167 (1997).
[36] E. Farhi and S. Gutman, Phys. Rev. B **48**, 7114 (1993).
[37] M. Schwartz, Phys. Lett. A **76**, 408 (1980).
[38] R. M. Hornreich and H. G. Schuster, Phys. Rev. B **26**, 3929 (1982).
[39] T. Vojta, J. Phys. A **26**, 2883 (1993).
[40] T. Vojta and M. Schreiber, Phys. Rev. B **50**, 1272 (1994).
[41] D. E. Feldman, J. Phys. A **31**, L177 (1998).
[42] A. Aharony, Phys. Rev. B **18**, 3328 (1978).
[43] V. S. Dotsenko and M. V. Feigelman, J. Phys. C **14**, L823 (1981).
[44] B. J. Minchau and R. A. Pelcovits, Phys. Rev. B **32**, 3081 (1985).
[45] D. C. Mattis, Phys. Lett. A **56**, 421 (1976).
[46] R. Bidraux, J. P. Carton, and G. Sarma, Phys. Lett. A **58**, 467 (1976).
[47] L. D. Landau and E. M. Lifshitz, *Electrodynamics of Continuous Media* (Pergamon Press, New York) [Russian original, Nauka, Moscow, 1982].
[48] R. B. Griffiths, Phys. Rev. Lett. **23**, 17 (1969).

Translated by M. E. Alferieff

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Generation, dynamics, and collisions of bending waves at domain boundaries in yttrium orthoferrite

M. V. Chetkin,*) Yu. N. Kurbatova, A. I. Akhutkina, and T. B. Shapaeva

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

(Submitted 2 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 2160–2169 (June 1999)

Solitary bending waves have been observed on domain boundaries of Néel type in wafers of yttrium orthoferrite, having a very sharp leading edge and an extended trailing edge and offset as a whole from the domain boundary and moving with high speeds close to the limiting velocity. Head-on collisions of two such waves of the same amplitude lead to their complete annihilation. Analogous collisions of two such waves, but of different amplitudes, lead to the appearance of a wave with the difference amplitude moving in the same direction as the wave of larger amplitude. The solitary bending waves investigated in this study appear to move under the action of gyroscopic forces acting on magnetic vortices on domain boundaries in yttrium orthoferrite, analogous to vertical Bloch lines with departure of the magnetization vector from the ac plane. From equality of the gyroscopic force with the friction force acting on the leading edge of the solitary bending wave we have estimated the amplitudes of these waves and the magnitudes of the topological charges of the magnetic vortices. © *1999 American Institute of Physics.* [S1063-7761(99)01906-X]

1. INTRODUCTION

Domain boundaries in orthoferrites move with high supersonic velocities[1,2] and with increasing magnetic field reach a record high limiting velocity of 20 km/s equal to the velocity of spin waves on the linear segment of the dispersion law.[3] The dependence of the velocity of a domain boundary on the magnetic field in all experiments performed to date is a nondecreasing function without any segments with negative differential mobility. Magnetic vortices of vertical-Bloch-line type on domain boundaries in orthoferrites have not been observed experimentally under either static or dynamic conditions. By virtue of Lorentz invariance of the dynamics of domain boundaries in orthoferrites[4,5] gyroscopic forces capable of moving the magnetic vortices should be equal to zero.[6] Experiments of the last two to three years have shown that solitary bending waves exist and move on domain boundaries in yttrium orthoferrite, offset from them as a whole.[7–9] These waves have a sharp leading edge and an extended trailing edge, in their shape resembling analogous waves accompanying magnetic vortices—vertical Bloch lines, which have been investigated in detail both experimentally[6,10] and theoretically[11–13] in ferrite–garnet films with uniaxial magnetic anisotropy. In the last few years a number of theoretical works have appeared addressing the dynamics of magnetic vortices in easy-plane antiferromagnets with Dzyaloshinskiĭ interaction.[14] These vortices contribute substantially to the thermodynamic characteristics of antiferromagnets and move with high velocities. What the above-mentioned works on solitary bending waves on domain boundaries in orthoferrites have in common is the requirement that the antiferromagnetism vector l and weak ferromagnetism vector m extend out from the easy magnetization plane in order to form three-dimensional magnetic vortices moving with high velocities, almost the limiting velocity, under the action of gyroscopic forces. These forces arise under the action of a magnetic field perpendicular to the ac plane.[15] In the present work we systematize the experimental results on generation, dynamics, and collisions of solitary bending waves on moving domain boundaries in yttrium orthoferrite. These data are compared with results on the dynamics and collisions of magnetic vortices in ferrite–garnet films, a qualitative analysis of the gyroscopic dynamics of solitary bending waves in orthoferrites is provided, and the amplitudes of these waves accompanying magnetic vortices are estimated. The possibility of experimentally observing limiting velocities of domain boundaries in yttrium orthoferrite in weak magnetic fields is linked with the record small value of their dimensionless decay parameter $\alpha \simeq 10^{-5}$ (Refs. 6 and 16). Such a small value of α makes it possible to experimentally observe motion of magnetic vortices at limiting velocities on domain boundaries in yttrium orthoferrite.

2. EXPERIMENTAL TECHNIQUE

The generation, dynamics, and results of pairwise collisions of solitary bending waves on domain boundaries in yttrium orthoferrite have been studied with the help of double-exposure high-speed photography in real time[6–10] using two light pulses with a time delay of 5–20 ns. The duration of light pulses from a dye laser pumped by a transverse-discharge pulsed nitrogen laser was 0.25 ns. A glass-ceramic wafer with a thin opening was placed between the electrodes of the spark-gap of the laser. This ensured extended stable operation of the laser. We used thin yttrium-orthoferrite wafers of thickness $30-40\ \mu$m, cut perpendicular to the optical axis. In such wafers in the absence of a magnetic field

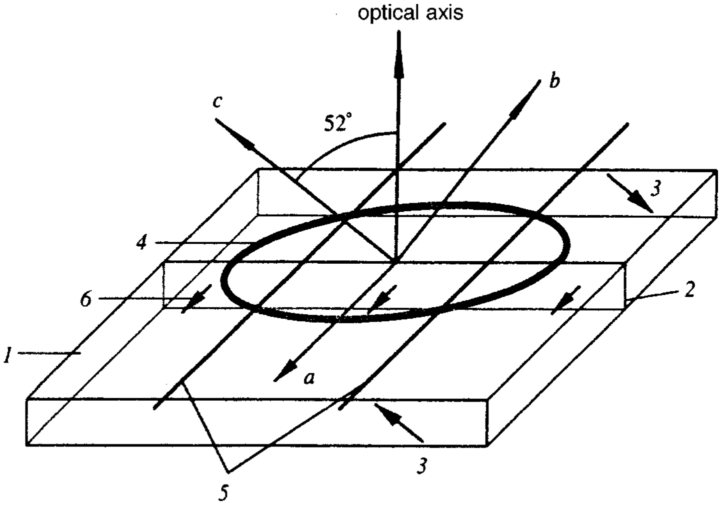

FIG. 1. Diagram of the arrangement of the main controlling elements and magnetization in a two-domain wafer of yttrium orthoferrite perpendicular to the optical axis for studying the generation, dynamics, and collisions of solitary bending waves at a domain boundary: *1* — orthoferrite wafer, *2* — Néel domain boundary, *3* — magnetizations of neighboring domains, *4* — coil for creating the magnetic field shifting the domain boundary, *5* — isolated leads for creating the magnetic field braking local segments of the domain boundary; and *6* — magnetization in the center of the domain boundary.

banded domain structures exist having domain boundaries of Néel type, perpendicular to the a axis lying in the plane of the sample. With the help of a magnetic field perpendicular to the sample surface with a gradient of 800 Oe/cm along the a axis, we set up an isolated domain boundary of Néel type (Fig. 1). In the center of this domain boundary the m vector was aligned or anti-aligned with the a axis and was oriented perpendicular to the plane of the boundary. On such an orthoferrite boundary the l and m vectors ordinarily rotate either clockwise or counterclockwise in the ac plane . Therefore, in principle, in such a geometry the existence of transitional regions is possible, with these regions being of magnetic vortex or vertical Bloch line type, separating segments with opposite directions of rotation of the l and m vectors. By rotating the gradient field about the optical axis of the sample, it is possible to set up domain boundaries of Néel, Bloch, and intermediate type in it. The isolated domain boundary in the sample was set in motion by the magnetic field from two coils with eight windings each and inner diameter 1.5 mm. To generate solitary bending waves, we used isolated leads 20 μm in diameter, perpendicular to the static orientation of the domain boundary. Short current pulses were sent through them, creating magnetic fields braking local segments of the domain boundary moving with supersonic speed. The field dependence of the speeds of the three different types of domain boundaries in the wafer is plotted in Fig. 2. The experiments primarily examined domain boundaries of Néel type. For this type of boundary the range of fields where the speed is constant and close to the speed of transverse sound was maximal and equal in width to about 100 Oe. In thicker wafers, for the Néel domain boundary and especially for the Bloch and intermediate domain boundaries the regions of constant velocity were substantially narrower (see Fig. 2) and in them it was not possible to observe generation of solitary bending waves. The question of the interval where the velocity of the domain boundary is constant and equal to the speed of transverse v_t or longitudinal v_l sound was considered theoretically only in the one-dimensional case.[6,17–19] As for the dependence of the width of the region $\Delta H(v_t)$ where the velocity of the domain boundary is constant and equal to the speed of transverse sound on the wafer thickness, this dependence is qualitatively linked with the curvature of the originally flat domain boundary that ensues upon the transition to supersonic velocity. Under these conditions, a braking force due to surface tension arises that is inversely proportional to the radius of curvature, and for sample thicknesses equal to 30−40 μm it is equivalent to a magnetic field of a few hundred oersteds, aligned opposite the direction of motion. This field retards the domain boundary and leads to an expansion of the region where the velocity is constant and equal approximately to the speed of sound. The first experimental observations of solitary bending waves in yttrium-orthoferrite wafers were made in the absence of an isolated lead.[7] In this case, what probably took place was a braking of the segment of the domain boundary in the upper part of the sample on the boundary with the windings of the coils creating the motive magnetic field. Generation of solitary waves in these first studies was significantly less reproducible than with the help of an isolated lead, as was done in subsequent works.

3. EXPERIMENTAL RESULTS

Experiments on the dynamics and collisions of solitary bending waves on a supersonic domain boundary in yttrium orthoferrite were preceded by experiments on the dynamics of the domain boundaries themselves. The dependence of the

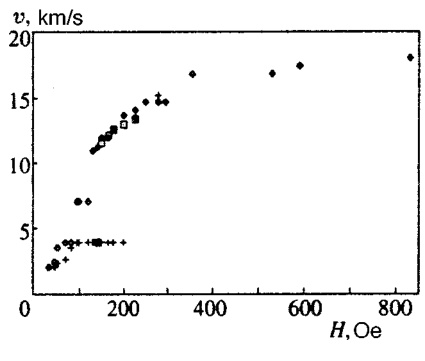

FIG. 2. Velocities of domain boundaries of Néel (+), Bloch (◇), and intermediate type (□), plotted as functions of the magnetic field in a $YFeO_3$ wafer of thickness 30 μm perpendicular to the optical axis.

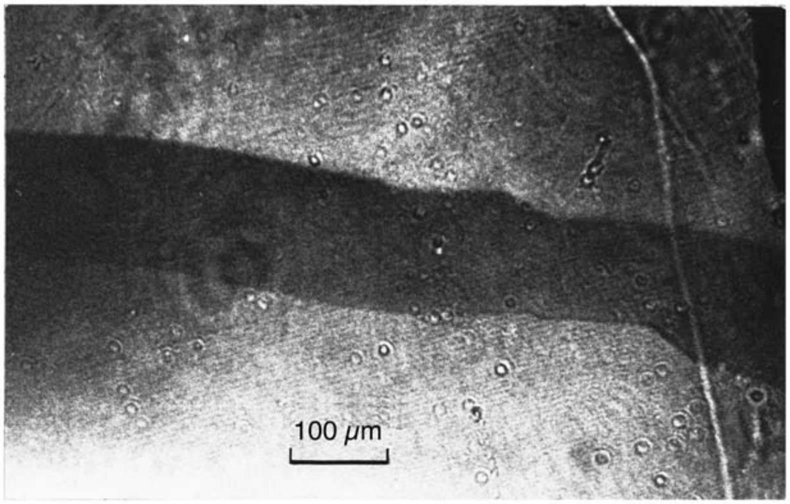

FIG. 3. Double-exposure high-speed photograph of two solitary bending waves moving from right to left on a domain boundary in $YFeO_3$, moving upward (in the figure). Delay time between light pulses 9 ns.

velocities of the domain boundaries in a $YFeO_3$ wafer of thickness 40 μm is plotted in Fig. 2 for different orientations of the plane of the boundary. For a domain boundary of Bloch type, for which in the central plane of its static orientation the weak ferromagnetism vector *m* lies in the plane of the domain boundary and is aligned with the *a* axis, the singularity in the speed of transverse sound is absent. On a domain boundary of Néel type, for which in the central plane of its static orientation the weak ferromagnetism vector *m* is oriented perpendicular to the plane of the domain boundary, a region of constant velocity exists having maximum width ΔH_t on the order of 100 Oe. In this interval the velocity of the Néel domain boundary is equal to the speed of transverse sound. Both of these results for Bloch and Néel domain boundaries are in agreement with theoretical analysis[6,17–19] and the results of earlier experiments.[1–3] For domain boundaries of intermediate type the widths of the regions of constant velocity near the speed of transverse sound fall monotonically from their maximum to zero. Note that the above characterization of Bloch and Néel domain boundaries in yttrium orthoferrite agrees with the data of earlier studies of the dynamics and mobility[6] and does not agree with their definition given in Ref. 20. This work defines Néel domain boundaries in terms of rotations of *l* in the *ac* plane and Bloch domain boundaries in terms of rotations in the *ab* plane. Experiments on the generation, dynamics, and collisions of solitary bending waves were performed on domain boundaries of Néel type. With further increase of the magnetic field above the constant velocity region the velocity of the domain boundary grew abruptly to 12 km/s. The subsequent local, rapidly relaxing decrease in the velocity of the domain boundary aided by a short current pulses in the isolated lead intersecting the domain boundary led to generation of stationary solitary bending waves, moving along the domain boundary during the entire observation time. An example of such waves is shown in Fig. 3, which shows two frames in which two isolated bending waves are clearly visible, moving from right to left along a domain boundary that is moving upward in the figure. The time delay between light pulses is 9 ns. The leading edges of these waves is very sharp, the trailing edges are extended, and both waves move along the domain boundary with velocity $u=16$ km/s. The velocity of the domain boundary is $v=12$ km/s, so that the velocities u and v and the limiting velocity of the domain boundary c obey the relation $v^2+u^2=c^2$, i.e., the total velocity of the solitary wave is close to the limiting velocity of the domain boundary. This means that the solitary waves move under the action of very strong forces, which can apparently only be gyroscopic forces. Indeed, the magnetic field moving the domain boundary upward (per the figure) can only brake the leading edge of the solitary wave. The field in the plane of the $YFeO_3$ wafer, which can, in principle, be created by magnets creating a gradient magnetic field, does not exceed a few tens of oersteds. It is clearly insufficient to move the leading edge of an isolated wave with a velocity near the limiting velocity, equal to 20 km/s. The amplitudes of the solitary bending waves in Fig. 3 are equal to 5 and 30 μm, and both these waves move with total velocities of 20 km/s. Bending waves in our experiment were generated by local braking of the domain boundary moving initially with a velocity of only 12 km/s. This velocity value ends up the interval of supersonic instability of the domain boundary, and it is found right in the middle between the limiting velocity and the speed of transverse sound.

Figure 4 shows a double-exposure high-speed photograph in which two solitary bending waves of identical amplitude are clearly visible, moving in opposite directions with the above-indicated speeds on a domain boundary moving upward (in the figure) with a velocity of 12 km/s. Motion in different directions is probably associated with different signs of the topological charges of magnetic vortices of the type vertical Bloch lines, which are accompanied by these waves. Different segments of the locally retarded part of the domain boundary move with various speeds less than or equal to 12 km/s, so that solitary bending waves can move along domain boundaries moving not only with a velocity of 12 km/s.

Figure 5 shows the results of head-on collisions of two solitary bending waves of the same amplitude, generated with the help of two leads intersecting the domain boundary.

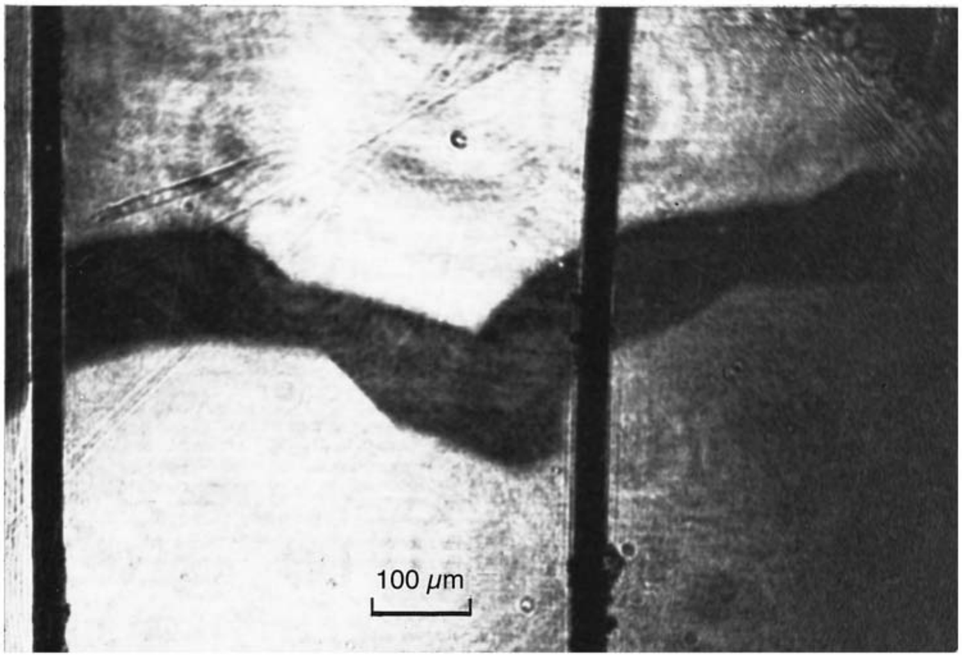

FIG. 4. Double-exposure high-speed photograph of two solitary bending waves moving in opposite directions on a domain boundary in $YFeO_3$, moving upward (in the figure). Delay time between light pulses 7 ns.

In the first frame, in Fig. 5a, they are quite far apart, in the second frame, after 6 ns they have almost annihilated. In Fig. 5b they have annihilated completely. This result can be explained by annihilation of topological charges of magnetic vortices accompanied by solitary bending waves and equal in absolute value but different in sign. Two counterpropagating solitary waves of different amplitudes, moving with velocities that are the same in absolute value form one solitary wave after they collide having amplitude equal to the difference of the amplitudes of the two colliding waves. This result is depicted in Fig. 6. The solitary wave formed after the collision continues to move in the same direction as the original wave of larger amplitude before the collision. On the whole, the dynamic profiles of solitary bending waves in yttrium orthoferrite are qualitatively very similar to the analogous profiles of such waves in ferrite–garnet films with uniaxial anisotropy. These waves in ferrite–garnets accompany magnetic vortices of vertical Bloch line type. The dynamics and collisions of magnetic vortices in ferrite–garnets were studied experimentally[6,10] and theoretically.[11–13] It was found that two vertical Bloch lines with topological charges of equal magnitude can, in a head-on collision, annihilate completely, annihilate partially, behave in a solitonlike manner, and even increase their topological charges near the critical velocity of the domain boundary. The maximum velocities of magnetic vortices on domain boundaries of ferrite–garnets do not exceed 150–200 m/s and they move under the action of gyroscopic forces. In yttrium orthoferrite the total velocity of the solitary bending waves is close to the limiting velocity of the domain boundary. This indicates that the motion of solitary bending waves takes place under the action of large gyroscopic forces, proportional to the large velocities of the domain boundaries in orthoferrites. These forces should be proportional to the velocity of the domain boundary v. Our experimental results confirm this conclusion. A decrease in the velocity of the domain boundary leads to a decrease in the velocity of the solitary bending wave. Extrapolation of the dependence $u(v)$ shows that in this case u also tends to zero.

4. DISCUSSION

The experimentally detected solitary bending waves on domain boundaries in yttrium orthoferrite have amplitudes of several microns to several tens of microns. They are offset as a whole from the domain boundary. Their rectangular leading edges resemble large-amplitude kinks on a domain boundary in $YFeO_3$, moving with the speed of transverse sound.[21] The motion of large-amplitude kinks takes place under the action of a magnetic field moving the domain boundary as a whole. According to the data of Fig. 2, the inclined leading edge of a kink, being a domain boundary of intermediate type, has a larger velocity in the same magnetic field than the entire Néel domain boundary. The leading edge of a solitary bending wave is also a boundary of intermediate type, but it cannot move under the action of an external magnetic field moving the entire domain boundary. Ignoring the external magnetic field moving the orthoferrite domain boundary, the equations of motion of this boundary are Lorentz-invariant.[4–6] The gyroscopic force in this case should be equal to zero.[6] In a magnetic field H moving the entire domain boundary as a whole, Lorentz invariance is violated and a gyroscopic force appears that is proportional to the ratio H/H_E, where H_E is the exchange interaction force. These forces in an antiferromagnet F_{AFM} and in a ferromagnet F_{FM} are related by the equation[14]

$$F_{AFM}=\frac{8H}{H_E}F_{GFM}\,.$$

The small factor H/H_E is present here. However, the speed of the domain boundary in orthoferrites reaches large values $\sim 10^6$ cm/s, which is much larger than the speed of the do-

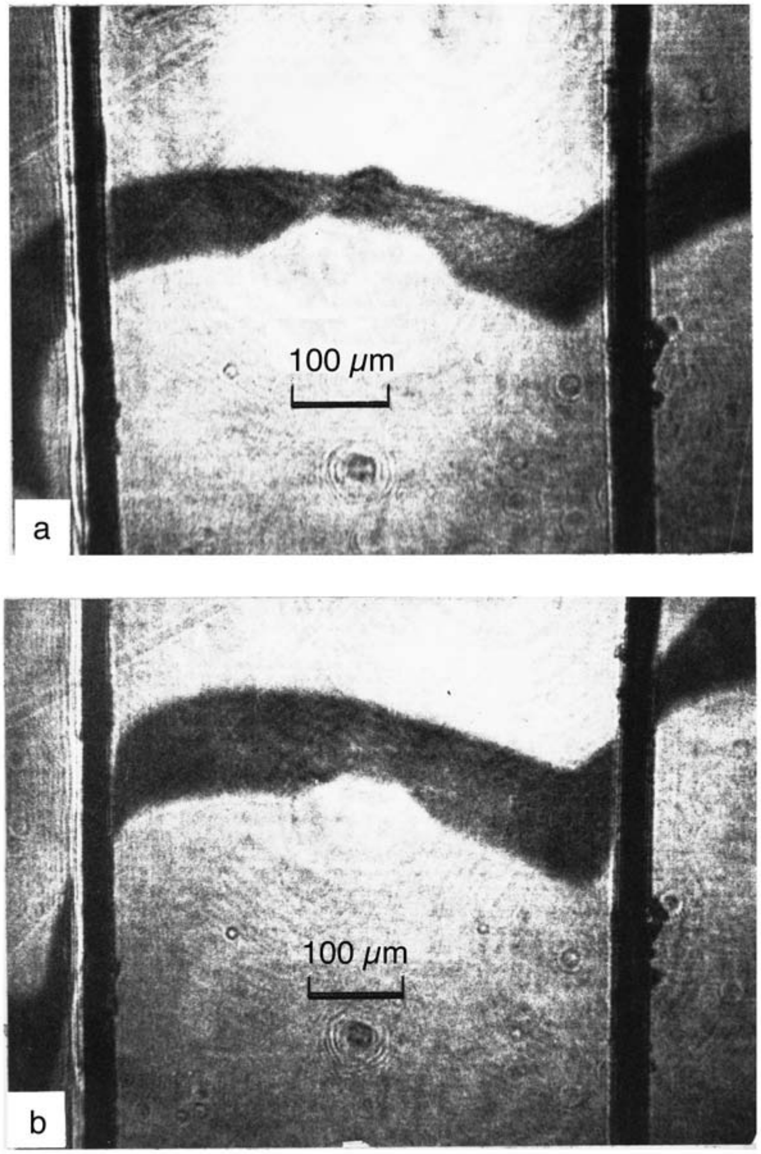

FIG. 5. Double-exposure high-speed photograph of two colliding solitary bending waves of the same amplitude on a domain boundary in $YFeO_3$, moving upward (in the figure): a — just before collision, b — annihilation of solitary waves. Delay time between light pulses 6 ns.

main boundary in ferromagnets, where it is on the order of 10^4 cm/s. The gyroscopic force acting on a magnetic vortex for steady-state motion should be balanced by the friction force, which is proportional to the dimensionless decay parameter $\alpha = 10^{-5}$. This value follows for yttrium orthoferrite from the mobility of the domain boundaries, which is equal to 10^4 cm/s$\cdot$Oe (see Fig. 2) and from the antiferromagnetic resonance data.[16] Equating the gyroscopic force to the friction force acting on the leading edge of a solitary bending wave, it is possible to estimate the amplitude of a solitary bending wave accompanying a magnetic vortex with definite topological charge. For a charge of 2π the wave amplitude is equal to $0.1\,\mu$m. The experimentally observed amplitudes correspond to topological charges from several units of π to several tens of units of π. In contrast to ferrite–garnets, the slope angles of the leading edges relative to the unperturbed domain boundary for waves of all amplitudes are identical, and their total velocities are large and close to the limiting velocity of the domain boundary. In uniaxial ferrite–garnet films, where the dynamics of the vertical Bloch lines and the results of their pairwise collisions have been investigated both experimentally and theoretically, these were either Néel segments of a Bloch domain boundary or Bloch segments of a Néel boundary. The lengths of these segments were much greater than the width of the domain boundary. In orthoferrites, besides rotation of l and m in the ac plane, only rotation of l in the ab plane without rotation of m but with variation in the magnitude and sign of m is known. Domain boundaries in orthoferrites without rotation of m were predicted in Refs. 22 and 23. Such domain boundaries were detected in $DyFeO_3$ from an analysis of the NMR spectrum[24] at temperatures much higher than the transition temperature to the antiferromagnetic state. Reference 25 calculated the ''fine'' structure of an immobile orthoferrite domain boundary with alternating segments with and without rotation of m. The dynamics of the segments without rotation of m was analyzed theoretically under the condition $v \ll c$ in Refs. 26 and 27. This analysis did not encompass the experimental results described above on the dynamics and collisions of

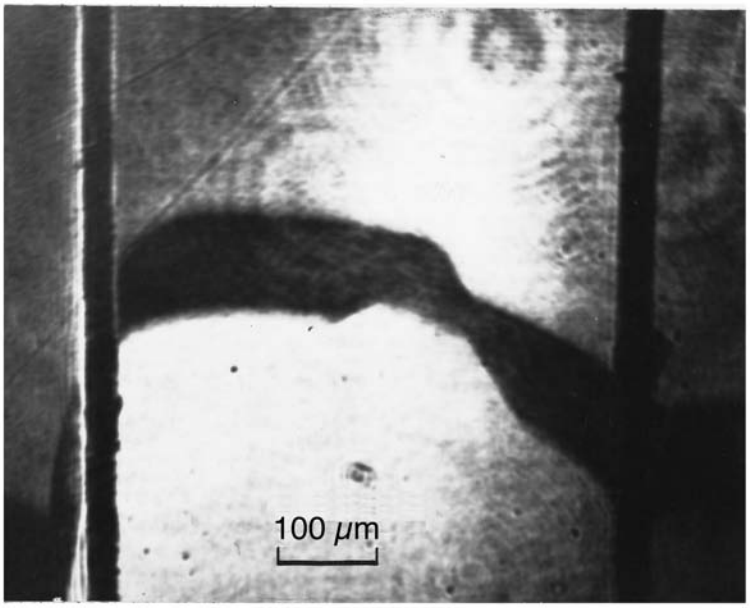

FIG. 6. Double-exposure high-speed photograph of two colliding solitary bending waves of different amplitudes on a domain boundary in $YFeO_3$, moving upward (in the figure): a — just before collision, b — annihilation of solitary waves. Delay time between light pulses 6 ns.

solitary bending waves in $YFeO_3$, where their velocities are comparable with or even equal to the limiting velocity c. For the case of a Néel domain boundary in our experiment rotation of l in the ab plane does not allow one to describe the structure of a magnetic vortex. Toward this end, it is necessary to analyze the possibility of rotation of the l and m vectors in other planes, different from the ones mentioned above. This requires that the orthorhombic anisotropy be overcome and leads to the result that the length of a magnetic vortex along a domain boundary will be less than its width, in contrast to the case for uniaxial ferrite–garnets. In principle, the large orthorhombic magnetic anisotropy can be decreased as a result of motion of a domain boundary with velocity close to the speed of transverse or longitudinal sound. For the appearance of the spin-reorientation transition $G_zF_x \leftrightarrows G_y$ inside a domain boundary, the surface of the sample can be significant.[25,28] Departure of the l and m vectors from the ac plane can be facilitated by the supersonic instability of an orthoferrite domain boundary, manifested, in particular, in the steep slope of its plane and by the absence of hysteresis in the field dependence of its velocity $v(H)$ in this region. For a complete theoretical analysis of the dynamics and collisions of magnetic vortices in weak ferromagnets, it is necessary to analyze these processes on the basis of a numerical simulation of the three-dimensional motion of the magnetic moment in analogy with the analysis of these processes in ferrite–garnets with uniaxial anisotropy performed in Ref. 13. It is of interest to extend this analysis to ferrite–garnet films with orthorhombic anisotropy, where magnetic vortices have not yet been experimentally observed.

5. CONCLUSIONS

1. On domain boundaries of Néel type in yttrium-orthoferrite wafers, we have detected and investigated solitary bending waves with sharp leading edges and extended trailing edges moving with very high velocities near the limiting velocities of domain boundaries in these weak ferromagnets, and set off as a whole from the domain boundary. These waves cannot move under the action of an external magnetic field moving the domain boundary as a whole. Such a field can only brake the leading edge of solitary bending waves. Nor can then move with velocities close to the limiting velocity under the action of small planar fields perpendicular to the plane of the domain boundary.

2. Head-on collisions of two solitary bending waves of the same amplitude on a domain boundary in yttrium orthoferrite, moving with a velocity of around 12 km/s, lead to their total annihilation. Analogous collisions of two solitary bending waves of different amplitude lead to the appearance of one solitary wave with the difference amplitude, moving in the same direction as the wave with the larger amplitude before the collision.

3. The experimental results obtained in yttrium orthoferrite and their comparison with the results of studies of the dynamics and collisions of magnetic vortices (vertical Bloch lines) in ferrite–garnets show that solitary bending waves in orthoferrites also accompany magnetic vortices moving under the action of gyroscopic forces with departure of the magnetization vector in them from the ac plane. These forces are proportional both to the small ratio of the magnetic field moving the domain boundary as a whole to the exchange field and to the very large speed of this boundary, and in the steady state are balanced by the friction force acting on the leading edge of the solitary bending wave and proportional to the very small dimensionless decay parameter in the Landau–Lifshitz equations of motion.

The authors thank A. K. Zvezdin and A. F. Popkov for discussion of the results, and A. M. Balbashov for providing single crystals of yttrium orthoferrite. This work was carried out with the financial support of the Russian Fund for Fundamental Research (Project No. 98-02-16440).

[*] E-mail: chetkin@magn145.phys.msu.su

---

[1] M. V. Chetkin, A. N. Shalygin, A. de la Campa, Fiz. Tverd. Tela **19**, 3470 (1977) [Sov. Phys. Solid State **19**, 2029 (1977)].

[2] C. H. Tsang, R. L. White, and R. M. White, J. Appl. Phys. **49**, 6052 (1978).

[3] M. V. Chetkin, A. De La Kampa, JETP Lett. **27**, 157 (1978).

[4] A. K. Zvezdin, JETP Lett. **29**, 553 (1978).

[5] V. G. Bar'yakhtar, B. A. Ivanov, A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **78**, 1509 (1980) [Sov. Phys. JETP **51**, 757 (1980)].

[6] V. G. Bar'jakhtar, M. V. Chetkin, B. A. Ivanov, and S. N. Gadetskiy, *Dynamics of Topological Magnetic Solitons*, Springer Tracts in Modern Physics, Vol. 129 (Springer, Berlin, 1994), p. 182.

[7] M. V. Chetkin, Yu. N. Kurbatova, and A. I. Akhutkina, Phys. Lett. A **215**, 211 (1996).

[8] M. V. Chetkin, Yu. N. Kurbatova, V. N. Filatov, JETP Lett. **66**, 797 (1997)].

[9] M. V. Chetkin and Yu. N. Kurbatova, IEEE Trans. Magn. **34**, 1075 (1998).

[10] M. V. Chetkin, I. V. Parygina, V. B. Smirnov *et al.*, Phys. Lett. A **140**, 428 (1989).

[11] A. K. Zvezdin, A. F. Popkov, I. P. Yarema, Zh. Éksp. Teor. Fiz. **98**, 1070 (1990) [Sov. Phys. JETP **71**, 597 (1990)].

[12] E. E. Kotova, V. M. Chetverikov, Zh. Éksp. Teor. Fiz. **98**, 2011 (1990) [Sov. Phys. JETP **71**, 1131 (1990)].

[13] L. L. Savchenko, M. V. Chetkin, and V. B. Bondarenko, J. Magn. Magn. Mater. **183**, 313 (1998).

[14] B. A. Ivanov and D. D. Sheka, Phys. Rev. Lett. **72**, 404 (1994).

[15] A. F. Andreev, V. I. Marchenko, Usp. Fiz. Nauk **130**, 39 (1980) [Sov. Phys. Usp. **23**, 21 (1980)].

[16] A. A. Mukhin, V. D. Travkin, S. P. Lebedev *et al.*, J. Phys. (Paris) **7**, S-1-713 (1997).

[17] V. G. Bar'yakhtar, B. A. Ivanov, A. L. Sukstanskiĭ, Zh. Éksp. Teor. Fiz. **75**, 2183 (1978) [Sov. Phys. JETP **48**, 1100 (1978)].

[18] A. K. Zvezdin, A. F. Popkov, Fiz. Tverd. Tela **21**, 1334 (1979) [Sov. Phys. Solid State **21**, 771 (1979)].

[19] A. K. Zvezdin, A. A. Mukhin, Zh. Éksp. Teor. Fiz. **102**, 577 (1992) [Sov. Phys. JETP **75**, 306 (1992)].

[20] N. Papanicolaou, Phys. Rev. B **55**, 12290 (1997).

[21] M. V. Chetkin, S. N. Gadetskiĭ, JETP Lett. **38**, 308 (1983).

[22] M. M. Farztdinov, S. D. Mal'ginova, Fiz. Tverd. Tela **12**, 2955 (1970) [Sov. Phys. Solid State **12**, 2385 (1970)].

[23] L. N. Bulaevskiĭ, V. L. Ginzburg, JETP Lett. **11**, 272 (1970).

[24] A. V. Zalesskiĭ, A. M. Savvinov, I. S. Zheludev, Zh. Éksp. Teor. Fiz. **68**, 1449 (1975) [Sov. Phys. JETP **41**, 723 (1975)].

[25] M. M. Farztdinov, M. A. Shamsutdinov, A. A. Khalfina, Fiz. Tverd. Tela **21**, 1522 (1979) [Sov. Phys. Solid State **21**, 878 (1979)].

[26] Yu. V. Melekhov, O. A. Perekhod, Fiz. Tverd. Tela **25**, 713 (1983) [*sic*].

[27] M. M. Farztdinov, M. A. Shamsutdinov, E. G. Ekomasov, Fiz. Tverd. Tela **30**, 1866 (1988), Fiz. Tverd. Tela **32**, 1542 (1990) [Sov. Phys. Solid State **30**, 1076 (1988); **32**, 902 (1990)].

[28] C. Micheletti, R. B. Griffiths, and J. H. Yeomans, J. Phys. A **30**, L233 (1997).

Translated by Paul F. Schippnick

Lattice vibrations of α'-NaV$_2$O$_5$

M. N. Popova,*) A. B. Sushkov, S. A. Golubchik, B. N. Mavrin, and V. N. Denisov

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

B. Z. Malkin and A. I. Iskhakova

Kazan State University, 420008 Kazan, Russia

M. Isobe and Y. Ueda

Institute for Solid State Physics, University of Tokyo, 7-22-1 Roppongi, Minato-ku, Tokyo 106, Japan

(Submitted 10 September 1998)
Zh. Éksp. Teor. Fiz. **115**, 2170–2189 (June 1999)

We have measured far infrared reflectance and transmittance spectra as well as Raman scattering spectra of α'-NaV$_2$O$_5$ single crystals for all the principal polarizations. The temperature range above the phase transition temperature $T_c = 35$ K was investigated, mainly. On the basis of this experimental study and of the lattice dynamics calculations we conclude that the symmetry of NaV$_2$O$_5$ in the high-temperature phase is described by the centrosymmetric D_{2h}^{13} space group. This conclusion leads to important physical consequences concerning the interpretation of one-dimensional magnetic properties of NaV$_2$O$_5$ and of the phase transition at 35 K considered earlier to be an ordinary spin-Peierls transition. The assignment of the observed phonons is given. Values of dielectric constants are obtained from the infrared data. Asymmetric shapes of several infrared lines and higher-order infrared vibrational spectra are discussed. The crystal field energy levels of the 3d electron localized at the V^{4+} site have been calculated in the framework of the exchange charge model using the values of effective charges obtained from the lattice dynamics calculations. According to the results of these calculations, the broad optical bands observed earlier in the vinicity of 1 eV can be interpreted as phonon assisted $d-d$ transitions.
[S1063-7761(99)02006-5]

1. INTRODUCTION

The vanadate α'-NaV$_2$O$_5$ has attracted considerable interest recently as the second inorganic compound undergoing the spin-Peierls transition (at the highest known temperature for the spin-Peierls compounds, $T_c = 35$ K[1]). The spin-Peierls transition is expected to occur within a system of linear spin-1/2 Heisenberg antiferromagnetic chains coupled to a three-dimensional phonon field. As a result of such a coupling, magnetic atoms of the chain dimerize and a spin gap opens.[2] One-dimensional magnetic properties of NaV$_2$O$_5$ above 35 K follow from magnetic susceptibility,[1] ESR[3] and angle-resolved photoemission[4] measurements.

Below 35 K the lattice dimerizes, as observed by x-ray[5] and Raman[6,7] scattering, infrared transmission[8] and reflection[9] measurements, while the magnetic susceptibility decreases isotropically, thus showing a spin gap formation.[6] The size of the gap $\Delta = 10$ meV follows from inelastic neutron scattering study of NaV$_2$O$_5$ single crystals.[5,10]

The structure of NaV$_2$O$_5$ contains double chains of edge-sharing distorted VO$_5$ pyramids running along the orthorhombic *b*-axis (Fig. 1). These double chains are linked via common corners of the pyramids to form the *ab*-layers. Na atoms lie between the layers.[1,11] The structure of NaV$_2$O$_5$ looks like the structure of V$_2$O$_5$[12] intercalated with sodium. In an early x-ray room temperature investigation on polycrystalline samples of NaV$_2$O$_5$, Carpy and Galy[11] suggested the noncentrosymmetric space group $C_{2v}^7 - P2_1mn$ with two nonequivalent vanadium positions in the unit cell. The picture of magnetic chains of V^{4+}O$_5$ ($S = 1/2$) pyramids isolated by nonmagnetic chains of V^{5+}O$_5$ ($S = 0$) pyramids proposed to account for one-dimensional magnetic properties of this mixed valence (V$^{4.5+}$) compound is compatible with this space group.[1]

However, the recent redetermination of the structure by single crystal x-ray diffraction at room temperature was in favor of the centrosymmetric D_{2h}^{13}-$Pmmn$ group with only one vanadium position in the structure.[13,14] Though the topology of the structure remains essentially the same as in the previously proposed noncentrosymmetric space group,[11] the possibility for charge ordering is, however, lost in the new higher symmetry group. Smolinski *et al.*,[13] and Horsch and Mack[15] suggested a quarter-filled ladder model for NaV$_2$O$_5$, with the spins carried by V–O–V molecular orbitals on the rungs of the ladder. They argued that the exchange interaction along the ladder is much greater than that between the neighboring ladders which would explain the one-dimensional magnetic properties of the high temperature phase of NaV$_2$O$_5$. The transition at 35 K was supposed to be an ordinary spin-Peierls transition. Quite recent ^{51}V-NMR experiment on a single-crystalline sample of NaV$_2$O$_5$ also

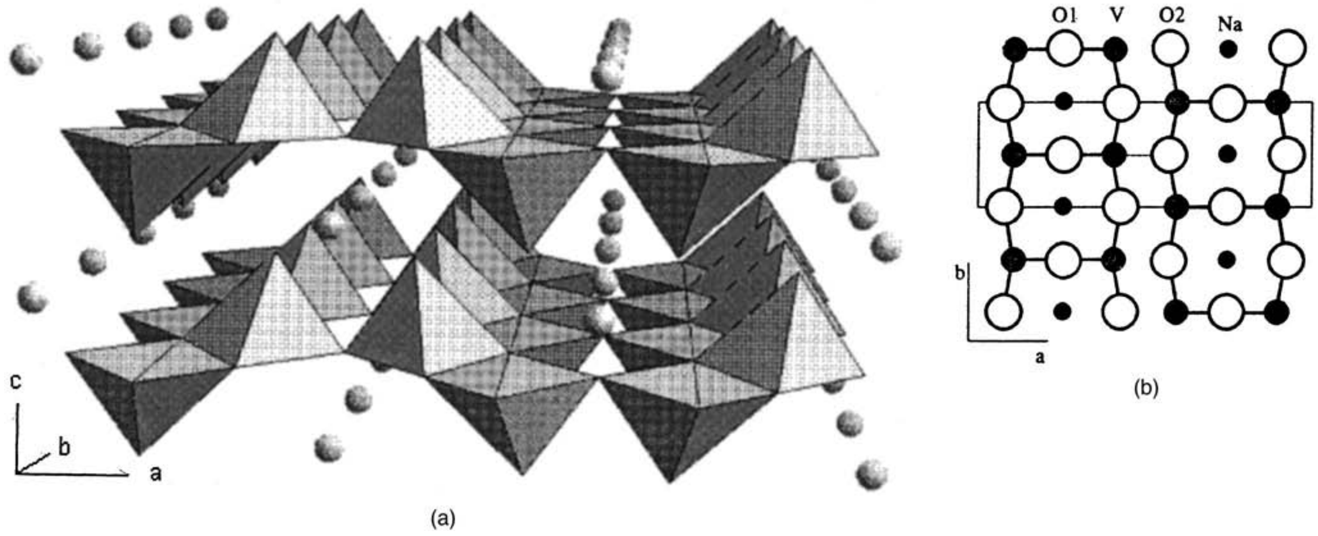

FIG. 1. Structure of NaV_2O_5. a) Stereometric projection. Oxygen and vanadium atoms are at the corners of and inside the pyramids, respectively. Sodium atoms are represented by balls. b) *ab* projection. Apical oxygen O3 atoms (situated above or below the corresponding *V* atoms) are not shown. A dashed line indicates the longest V–O2 bond (0.199 nm). *ab*-projection of the crystal unit cell is shown by a thin solid line.

revealed only one vanadium position in the high-temperature phase but pointed unambiguously to the existence of two different vanadium sites occupied by V^{4+} and V^{5+} at liquid-helium temperatures.[16] Thus, the transition at 35 K is connected with a structure and charge ordering processes. Very recently. Seo and Fukuyama[17] and Mostovoy and Khomskii[18] proposed a zigzag scheme of V^{4+}–V^{5+} ordering. Seo and Fukuyama argued that, as a result, two-dimensional lattice of antiferromagnetic dimers is formed,[17] while Mostovoy and Khomskii gave reasons in support of a system of alternating chains.[18] Thalmeier and Fulde[19] have presented some theoretical reasons for the primary charge ordering which provides neighboring linear V^{4+} and V^{5+} chains with a subsequent spin-Peierls transition. Two close transitions near 35 K in NaV_2O_5 were detected by Köppen *et al.*, via thermal expansion measurements.[20]

In view of these recent works, the symmetry problem of the high-temperature phase seems to be a matter of great urgency. Raman and infrared measurements could give additional information to clarify whether the space group is centrosymmetric or not, because of totally different selection rules in these two cases. We reanalyzed our earlier infrared and Raman spectra of NaV_2O_5[21] and found that they are in a better agreement with the centrosymmetric D_{2h}^{13} space group than with the noncentrosymmetric C_{2v}^{7}. However, in our work[21] we did not measure infrared spectra in the $\mathbf{E}\|\mathbf{c}$ polarization. Also, the signal-to-noise ratio of Raman spectra was rather low. In the present work, we reinvestigate vibrational spectra of the high-temperature phase of NaV_2O_5 using different single crystals, including extraordinary big ones. We present far-infrared reflectivity as well as Raman-scattering spectra for all principal polarizations. In addition, transmittance spectra were studied. We show that our results are in a much better agreement with the centrosymmetric D_{2h}^{13} group than with the noncentrosymmetric C_{2v}^{7}. The assignment of vibrational modes is given. It is based on a comparison with results of the previously studied of V_2O_5[22] and on lattice dynamics calculations of this work performed in the framework of the rigid-ion model.

2. EXPERIMENTAL

Single crystals of stoichiometric α'-NaV_2O_5 used in this study were grown by a melt growth method using $NaVO_3$ as a flux.[23] Samples from different batches were used. One sample was $1.3\times 8\times 1$ mm, another one was $3\times 17.3\times 1.6$ mm along *a*-, *b*-, and *c*-axes, respectively. For transmission measurements we have prepared four thin samples cleaved perpendicular to the *c*-axis. Their thicknesses were 110 ± 1, 45 ± 5, 14 ± 1.5 and 6 ± 1 μm. The samples were checked with x-ray diffraction, magnetization, and ESR measurements. They exhibited a sharp transition at about 35 K.

Reflection and transmission measurements were performed with a BOMEM DA3.002 Fourier transform spectrometer at nearly normal incidence of polarized infrared radiation. The following geometries of the experiment were used: 1) $\mathbf{k}\|\mathbf{c}$, $\mathbf{E}\|\mathbf{a}$ and $\mathbf{E}\|\mathbf{b}$; 2) $\mathbf{k}\|\mathbf{a}$, $\mathbf{E}\|\mathbf{c}$ and $\mathbf{E}\|\mathbf{b}$. Room-temperature reflectance and transmittance spectra were measured in a spectral range 30–5000 cm^{-1} with a resolution 0.5–2.0 cm^{-1}. Using both reflectance and transmittance spectra, the absorption coefficient α was calculated. Low-temperature (down to 6 K) transmittance spectra were measured with a He vapor cryostat in the spectral range 30–1000 cm^{-1} with a resolution 0.05–1.0 cm^{-1}.

Raman spectra were excited at room temperature by the 514-nm and 488-nm lines of an Ar-ion laser in backscattering geometries, dispersed by a home-made triple spectrograph, and recorded using a multichannel system consisting of an image intensifier tube with a multichannel plate and a vidicon.

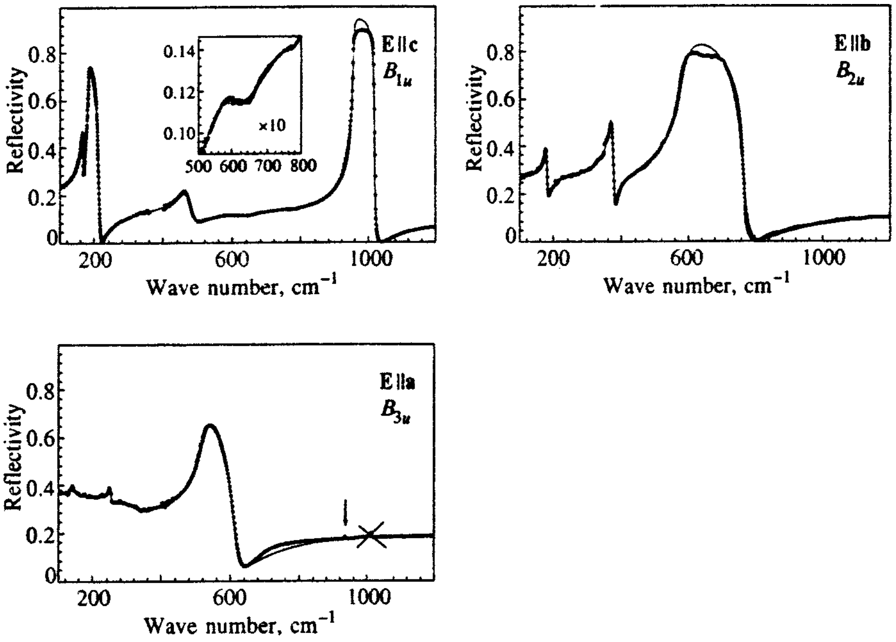

FIG. 2. Room-temperature far-infrared reflectivity spectra of NaV_2O_5. Open circles represent experimental data. Solid lines are fit results (see the text).

3. RESULTS

3.1. Factor-group analysis

There are two formula units and, hence, 16 atoms in the NaV_2O_5 orthorhombic unit cell with lattice constants $a = 1.1316\,\mathrm{nm}$, $b = 0.3611\,\mathrm{nm}$, $c = 0.4797\,\mathrm{nm}$.[11,13,14] Below, we present the results of factor-group analysis for both centrosymmetric D_{2h}^{13}[13,14] and noncentrosymmetric C_{2v}^{7}[11] space groups.

a. Space group D_{2h}^{13}*-Pmmn*. The notation *Pmmn* refers to the standard axis setting, such that **x**||**a**, **y**||**b**, **z**||**c**. It follows from x-ray diffraction data[13,14] that Na atoms occupy $2b$ positions (the corresponding fractional atomic coordinates are defined by the basis vectors $\mathbf{r}_1(\mathrm{Na}) = -\mathbf{r}_2(\mathrm{Na}) = (1/4, -1/4, z_1), z_1 = 0.8592$) and oxygen O1 atoms occupy $2a$ positions $(\mathbf{r}_1(\mathrm{O1}) = -\mathbf{r}_2(\mathrm{O1}) = (1/4, 1/4, z_2), z_2 = 0.5195)$, both these positions having C_{2v}^{z} local symmetry. V, O2 and O3 atoms reside in different $4f$ positions $(\mathbf{r}_1(\mathbf{A}) = -\mathbf{r}_3(\mathbf{A}) = (x_A, 1/4, z_A)$, $\mathbf{r}_2(\mathbf{A}) = -\mathbf{r}_4(\mathbf{A}) = (1/2 - x_A, 1/4, z_A)$; $\mathbf{A} = \mathrm{V}$, O2, O3; $x_{\mathrm{V}} = 0.40212$, $z_{\mathrm{V}} = 0.39219$, $x_{\mathrm{O2}} = 0.57302$, $z_{\mathrm{O2}} = 0.48769$, $x_{\mathrm{O3}} = 0.38548$, $z_{\mathrm{O3}} = 0.05803)$ with the local symmetry C_s^{xz}. These positions yield the following irreducible representations:[24,25]

$$C_{2v}^{z}:\Gamma = A_g + B_{2g} + B_{3g} + B_{1u} + B_{2u} + B_{3u},$$

$$C_s^{xz}:\Gamma = 2A_g + B_{1g} + 2B_{2g} + B_{3g} + A_u + 2B_{1u} + B_{2u} + 2B_{3u}.$$

Multiplying the representations given above by the number of different positions of the appropriate symmetry, summarizing them, and subtracting acoustic modes $(B_{1u} + B_{2u} + B_{3u})$, we obtain the following NaV_2O_5 optical vibrational modes:

$$\begin{aligned}\Gamma^{\mathrm{vib}}_{\mathrm{NaV_2O_5}}(Pmmn) = {} & 8A_g(aa,bb,cc) + 3B_{1g}(ab) \\ & + 8B_{2g}(ac) + 5B_{3g}(bc) + 3A_u \\ & + 7B_{1u}(\mathbf{E}\|\mathbf{c}) + 4B_{2u}(\mathbf{E}\|\mathbf{b}) \\ & + 7B_{3u}(\mathbf{E}\|\mathbf{a}). \end{aligned} \tag{1}$$

There are 45 vibrational modes in total. A_u modes being silent, 24 Raman $(A_g, B_{1g}, B_{2g}, B_{3g})$ and 18 infrared (B_{1u}, B_{2u}, B_{3u}) active modes are expected to be found in the spectra of NaV_2O_5, provided the crystal space group is D_{2h}^{13}.

b. Space group $C_{2v}^{7} - P2_1mn$. In their original work,[11] Carpy and Galy adopted the axis setting for the $P2_1mn$ space group. Below, we use the standard setting for the $Pmn2_1$ space group: $x\|b, y\|c, z\|a$. There are two nonequivalent V positions, five nonequivalent O positions and one Na position in this group, all of them being $2a$ positions with C_s^{yz} local symmetry. In the same way as in the previous case, using tables[25] and subtracting acoustic modes $(A_1 + B_2 + B_1)$, we find the following vibrational modes:

$$\begin{aligned}\Gamma^{\mathrm{vib}}_{\mathrm{NaV_2O_5}}(Pmn2_1) = {} & 15A_1(aa,bb,cc;\mathbf{E}\|\mathbf{a}) + 8A_2(bc) \\ & + 7B_1(ab;\mathbf{E}\|\mathbf{b}) + 15B_2(ac;\mathbf{E}\|\mathbf{c}). \end{aligned} \tag{2}$$

There are 45 optical modes again. But in the case of this noncentrosymmetric space group all of them are Raman active, 37 of them are also infrared active.

3.2. Infrared spectra

Figure 2 shows the room temperature far-infrared reflectivity spectra of NaV_2O_5, for different polarizations of the incident light. Experimental data are presented by open circles. Measured spectra were least-squares fitted by the spectra computed according to the expression

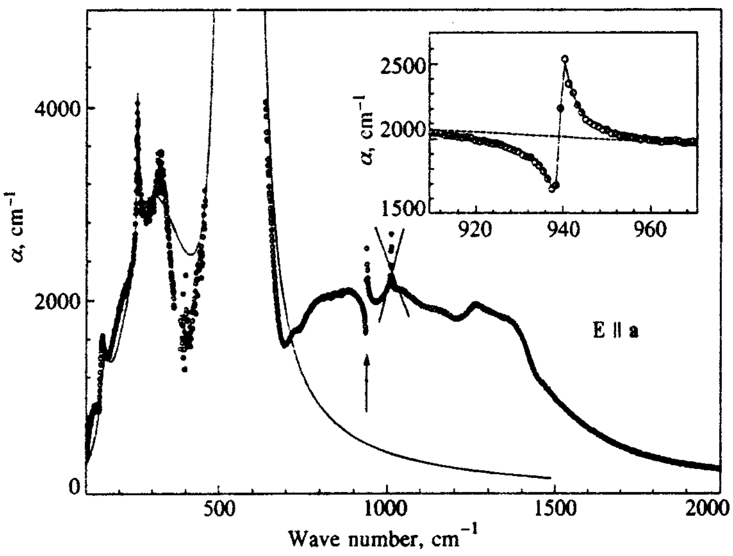

FIG. 3. Absorption coefficient α in the region of low-frequency absorption bands at room temperature. The arrow indicates a Fano-type resonance, shown separately in the inset. Open circles represent experimental data. Solid line in the main figure was calculated using the parameters obtained by fitting the reflectance spectrum. Solid line in the inset is a result of fitting the expression (5) with $\alpha_B(\omega)$, shown as a dashed line.

$$\mathscr{R}=\left|\frac{\sqrt{\varepsilon}-1}{\sqrt{\varepsilon}+1}\right|^2. \tag{3}$$

The classical dispersion formula for N independent damped oscillators was used:

$$\varepsilon=\varepsilon_\infty+\sum_{i=1}^{N}\frac{4\pi f_i\omega_i^2}{\omega_i^2-\omega^2-i\gamma_i\omega}. \tag{4}$$

For **E**∥**b** and **E**∥**a** polarizations the number of oscillators and initial values of parameters were taken from the transmittance spectra ([21] and the present work). The anomaly crossed out in Figs. 2 and 3 at 1014 cm^{-1} in **E**∥**a** polarization and also observed in **E**∥**b** polarization for some samples depends on a particular sample, and is evidently not an intrinsic property of NaV$_2$O$_5$. It was not taken into account in the fitting procedure. In addition to weakly damped phonon oscillators, an overdamped oscillator centered at about 300 cm^{-1} ($\omega_i=291\,\mathrm{cm}^{-1}$, $\gamma_i=260\,\mathrm{cm}^{-1}$, $f_i=0.38$) was introduced in **E**∥**a** polarization to account for a low-frequency part of a broad absorption band of a complex two-humped shape found in our previous study[21] (see also Fig. 3). We failed to model the high-frequency hump of this band centered at about 1000 cm^{-1} with a similar oscillator, and did not try to use a more complicated model. This results in only a fair fit to the hight-frequency part of the reflectance spectrum. In **E**∥**a** polarization the phonon at about 150 cm^{-1} could not be fitted well. This line is strongly asymmetric in transmittance spectra, obviously due to interaction with the underlying broad band.

The small bump in reflection at 939 cm^{-1} shown by the arrow in Fig. 2 corresponds to the Fano-type resonance[26] well seen in the absorbance spectrum (Fig. 3). One more such resonance becomes visible below 200 K at about 91 cm^{-1} (see Fig. 4 and also Ref. 8). We fitted the absorption coefficient in the vicinity of these two strongly asymmetric lines by the expression:[26]

$$\alpha(\omega)=\alpha_B(\omega)+\alpha_0\frac{q^2+2\xi q-1}{1+\xi^2}. \tag{5}$$

where $\xi=(\omega-\omega_r)/\gamma$, $\alpha_B(\omega)$ is a slowly varying broad band absorption (it is shown as a dashed line in the vicinity of the 939-cm^{-1} sharp resonance in Fig. 3), and α_0, ω_r, γ, and q are variable parameters. Such an expression describes various physical situations for a sharp transitions overlapped by a broad continuum. The line shape of the sharp transition is altered by interference with a continuum, and depends heavily on the strength of the interaction between discrete and continuum states. The parameter q being inversely proportional to the matrix element of an interaction, the case $|q|=\infty$ corresponds to zero interaction and results in a normal Lorenzian resonance, $|q|=1$ yields a dispersion-like curve, while $|q|=0$ gives an inverted Lorenzian (antireso-

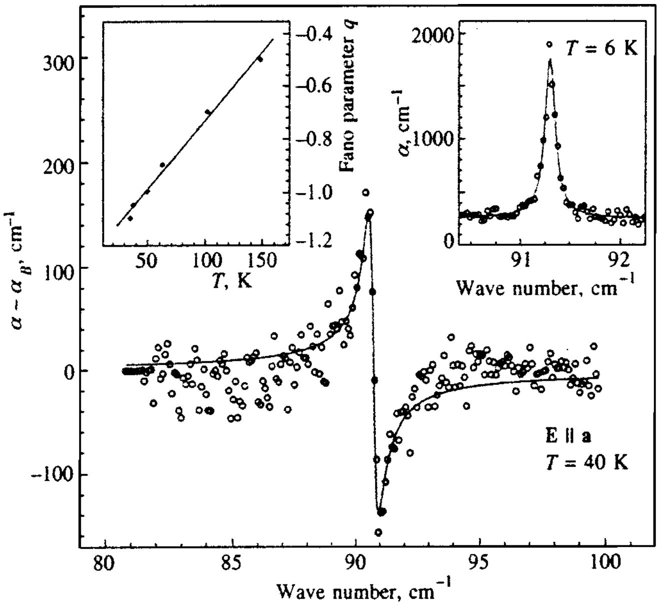

FIG. 4. Fano resonance near 91 cm^{-1} at 40 K (open circles) and its fit using Eq. (5) with parameters $\omega_r=90.7\,\mathrm{cm}^{-1}$, $\gamma=0.2\,\mathrm{cm}^{-1}$, $q=-1.0$, $\alpha_B(\omega_2)=270\,\mathrm{cm}^{-1}$, and $\alpha_0/\alpha_B=0.3$ (solid line). The temperature dependence of the Fano parameter q is given in the left inset. The right inset presents the absorbance spectrum in the vicinity of 91 cm^{-1} at 6 K with resolution 0.05 cm^{-1} (open circles), and Lorenzian fit with FWHM$=0.10\,\mathrm{cm}^{-1}$.

TABLE I. Infrared active vibrational modes (cm^{-1}) and dielectric constants of NaV_2O_5.

Polarization, mode symmetry	Observed													Calculated ($Pmmn$)	
	Transmission						Reflection								
	T = 40 K		T = 300 K				T = 300 K								
	ω_{TO}	γ_{TO}	ω_{TO}	γ_{TO}	$\varepsilon^a_{\omega_1}$	$\varepsilon^a_{\omega_2}$	ω_{TO}	γ_{TO}	ω_{LO}	γ_{LO}	$10^3 f$	ε_∞	ε_0	ω_{TO}	ω_{LO}
E||**c**						7.5 ±0.2	162	5.2	165	5.7	45	3.9	7.7	216	219
$B_{1u}(Pmmn)$							179	8.4	212	8.3	130			232	256
or							—							298	298
$B_2(Pmn2_1)$							468	38.0	483	38.0	23			430	430
							591	119.9	597	119.0	6.8			589	690
							760	59.4	762	59.3	1.6			691	716
							955	2.5	1017	3.0	39			961	1036
E||**b**	178	4	175	12	5.2 ±0.2	10.2 ±0.2	175	8.3	180	8.4	39	4.9	9.5	141	173
$B_{2u}(Pmmn)$	225	1	—											240	266
or	371		367	16			365	12.8	378	13.3	60			388	483
$B_1(Pmn2_1)$	594	13	582				584	29.5	769	29.0	271			578	747
E||**a**	91[c]		—		9.6 ±0.3	15.0 ±0.6						7.7	15.8	111	126
$B_{3u}(Pmmn)$	140[b]		145[b]				153	33.0	155	34.1	46			130	177
or	254		251				251	7.3	252	8.2	9			227	276
$A_1(Pmn2_1)$	—	—					—							493	534
	531	18	526	53			525	39.5	622	52.4	208			538	653
	—						—							742	808
			939[d]											955	957

[a] $\omega_1 = 3200\ \mathrm{cm}^{-1}$, $\omega_2 = 40\ \mathrm{cm}^{-1}$.

[b] Asymmetric line.

[c] Fano-type resonance: $\omega_\tau = 90.7\ \mathrm{cm}^{-1}$, $\gamma = 0.2\ \mathrm{cm}^{-1}$, $q = -1.0$, $\alpha_0/\alpha_B = 0.3$.

[d] Fano-type resonance: $\omega_\tau = 939\ \mathrm{cm}^{-1}$, $\gamma = 1.0\ \mathrm{cm}^{-1}$, $q = 1.1$, $\alpha_0/\alpha_B = 0.2$.

nance). The ratio α_0/α_B shows what fraction of the continuum states interacts with a sharp excited state. The results of fitting are displayed in the inset of Fig. 3 and in Fig. 4. A similar fit should be performed for the resonance at about 150 cm^{-1} but we failed to construct $\alpha_B(\omega)$ in this case. The fit parameters obtained are listed in Table I; ω_{TO} and γ_{TO} denote ω_i and γ_i of Eq. (4) or ω_r and γ of Eq. (5). LO frequencies and damping constants were calculated as complex roots of the equation $\varepsilon(\omega) = 0$.

The left inset of Fig. 4 presents the temperature dependence of the Fano parameter q for the spectral line near 91 cm^{-1} at temperatures higher than $T_c = 35\,\mathrm{K}$. It should be mentioned that below 35 K, the shape of this line changes to an ordinary Lorenzian (see the right inset of Fig. 4 and also[8]). Simultaneously, continuum absorption diminishes markedly in this spectral region while it is essentially unchanged at the maximum of the low-frequency hump at 320 cm^{-1}.

With decreasing the temperature, besides the asymmetric resonance at 91 cm^{-1} in **E**‖**a** transmittance, two lines at 215 and 225 cm^{-1} appear in **E**‖**b** transmittance spectra as well.[8,21] We have studied the resonances at 91 and 939 cm^{-1} (**E**‖**a**); 215 and 225 cm^{-1} (**E**‖**b**) for the samples of different thicknesses, and found that while the intensities of 91-, 939- and 225-cm^{-1} lines are proportional to the sample thickness d (that is, $\alpha = \mathrm{const}$), the intensity of the 215-cm^{-1} line is essentially independent of the thickness ($\alpha d \simeq \mathrm{const}$). Consequently, while the frequencies 91 and 939 cm^{-1} (**E**‖**a**) and 225 cm^{-1} (**E**‖**b**) correspond to intrinsic resonances, $\omega = 215\,\mathrm{cm}^{-1}$ must refer to a surface excitation. All the observed infrared phonon frequencies together with the calculated ones are displayed in Table I.

NaV_2O_5 crystals are well transparent in the frequency region between 2500 and 4500 cm^{-1} and below 100 cm^{-1}. In these regions, an interference pattern was observed in **E**‖**a** and **E**‖**b** transmittance spectra of the samples of good quality. We also managed to observe the interference pattern below 100 cm^{-1} in **E**‖**c** transmittance of 1.3-mm thick sample. By measuring the distances Δ between the interference maxima, we found the refractive indexes n according to the relation

$$\Delta = \frac{1}{2dn}. \tag{6}$$

Appropriate values of $\varepsilon = n^2$ are listed in Table I.

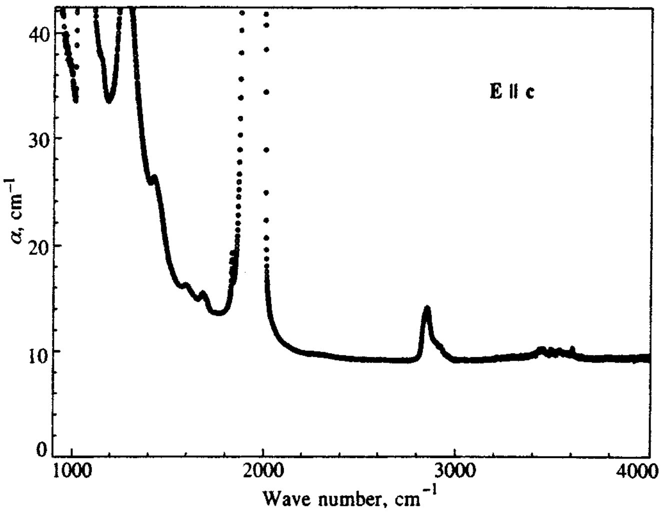

FIG. 5. Absorbance spectrum of NaV_2O_5 in the region of multiphonon bands at room temperature.

We also looked for the higher-order vibrational spectra by measuring the transmittance of thick ($d=0.4$–3.0 nm) samples in the frequency range 1000–4000 cm^{-1}. While no pronounced features were found in **E**∥**a** and **E**∥**b** polarizations, sharp resonances were observed in **E**∥**c**(**k**∥**a**) polarization at 1930, 2858 and possibly 1072 and 1270 cm^{-1}, the latter two lines being somewhat masked by the edge of a strong phonon at 955 cm^{-1} (Fig. 5).

3.3. Raman spectra

Polarized room-temperature Raman spectra of NaV_2O_5 in the spectral range 80–1000 cm^{-1} are shown in Fig. 6. One can see immediately that the three diagonal components *aa,bb,cc* of the Raman scattering tensor differ markedly one from another, which points to considerable anisotropy of the structure. The most intense spectra were observed in the A_g geometry $a(cc)\bar{a}$. The intensity of the lines marked by asterisks in B_{ig} ($i=1,2,3$) spectra depended strongly on slight variations in the sample orientation. Probably, these lines are due to a leakage of strong lines from A_g geometries. We failed to assign for certain a weak feature near 100 cm^{-1} in the $b(ac)\bar{b}$ spectrum overlapped by a strong unshifted line that is present in this geometry. It might-possibly come from the leakage of a very strong line 90 cm^{-1} from the (cc) polarization. Frequencies of the observed Raman modes together with the calculated ones are collected in Table II.

As we have already reported,[21] besides relatively narrow lines, a broad band with a maximum near 600 cm^{-1} is observed in the $c(aa)\bar{c}$ spectrum (see Fig. 6). Since this band appears under both 514.5-nm and 488-nm excitation, we conclude that it originates from the Raman scattering process. However, a large width of this band (213 cm^{-1}), which is essentially independent of the temperature, means that it is not attributable to fundamental modes.

We also studied Raman spectra of Na-deficient samples $Na_{1-x}V_2O_5$ ($x=0$, 0.05, 0.10, 0.15). The most prominent changes occur in the $A_g(aa)$ spectrum (see Fig. 7). The 447-cm^{-1} Raman line moves to higher frequencies as x increases. Its position shown by the vertical dashed lines in Fig. 7 is 477 cm^{-1} for the sample with $x=0.15$. The maximum of the broad band moves in the opposite direction, namely, from 632 cm^{-1} for $x=0$ to 562 cm^{-1} for $x=0.15$. This change of the frequency difference between these two Raman bands is, probably, due to a change in intermode interaction. The shape of the broad band can be approximated well by a Gaussian for all the values of x, its width growing from 213 cm^{-1} at $x=0$ to 290 cm^{-1} at $x=0.15$. As for phonon Raman lines, their shape is almost Lorenzian, and their width grows too. For example, the lines at 177, 301 and 531 cm^{-1} broaden from 11, 18 and 20 cm^{-1} at $x=0$ to 16, 27 and 34 cm^{-1} at $x=0.15$. The broadening of Raman bands with increasing x is probably associated with an increase in lattice disorder.

It is difficult to compare absolute intensities of the spectra at different x. However, certain conclusions concerning relative spectral intensities can be drawn. The most prominent features are a rise in intensity of the 301 cm^{-1} line, and the amergence of a new line at 988 cm^{-1} at $x=0.15$. All these results were obtained by expanding the observed spectrum into individual spectral profiles. An example of such expansion is shown in Fig. 8.

3.4. Calculations of vibrational spectra

To obtain an information about the phonon spectrum of NaV_2O_5 throughout the Brillouin zone, which is necessary for the analysis of the spin-phonon interaction effects, we

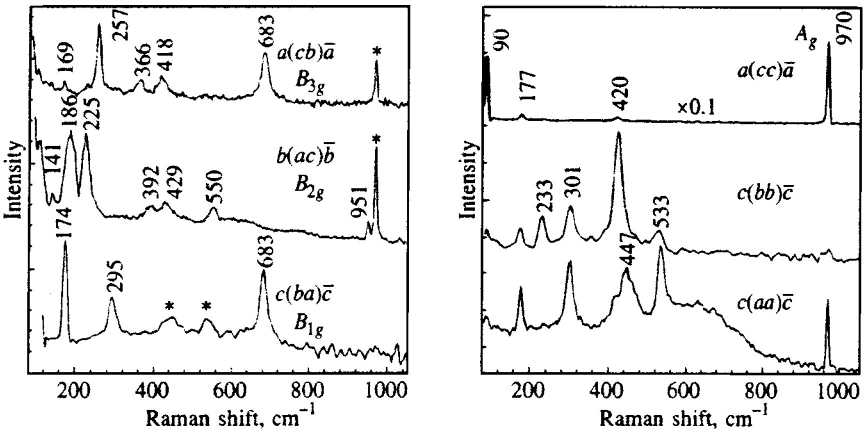

FIG. 6. Room-temperature Raman spectra of NaV_2O_5. Asterisks label A_g lines seen in B_{ig} ($i=1,2,3$) spectra.

TABLE II. Room-temperature Raman frequencies (cm^{-1}) for NaV$_2$O$_5$.

aa, *bb*, *cc* $A_g(Pmmn)$ or $A_1(Pmn2_1)$		*ab* $B_{1g}(Pmmn)$ or $B_1(Pmn2_1)$		*ac* $B_{2g}(Pmmn)$ or $B_2(Pmn2_1)$		*bc* $B_{3g}(Pmmn)$ or $A_2(Pmn2_1)$	
Observed	Calc. ($Pmmn$)	Observed	Calc. ($Pmmn$)	Observed	Calc. ($Pmmn$)	Observed	Calc. ($Pmmn$)
90 (*cc*, *aa*)	91	174	191	141	129	169	149
177 (*aa*, *bb*, *cc*)	226	295	288	186	193	257	239
233 (*bb*)	319	683	679	225	296	366	262
301 (*aa*, *bb*)	362			392	332	418	396
420 (*bb*, *cc*)	439			429	410	683	685
447 (*aa*)	518			550	619		
533 (*aa*, *bb*)	626			–	798		
970 (*cc*, *aa*)	964			951	961		

have considered the lattice dynamics of this crystal in the framework of the rigid ion model. The goal of this study is to display the basic pairwise interionic interactions that determine the main features of the measured Raman and infrared transmittance and reflection spectra.

A theoretical analysis of the vibrational spectra has been carried out for both lattice structures proposed in the literature. We did not obtain any physically well-grounded set of parameters which might provide the stable C_{2v}^7 lattice structure. We therefore discuss in this section only vibrations of the centrosymmetric lattice with the D_{2h}^{13} space group.

From the large measured large TO–LO splittings of some normal modes at the center of the Brillouin zone (Γ point), it is clear that long-range Coulomb forces play a crucial role in formation of the vibrational spectrum of NaV$_2$O$_5$. The potential energy of the lattice was represented by a sum of Coulomb and non-Coulomb interactions. The Coulomb terms in the dynamical matrix were calculated exactly using the Ewald method. Non-Coulomb interactions in the form of the Born–Mayer potentials with the exponential dependence on the interionic distance $r(\varphi_{ij}(r)=C_{ij}\exp(-r/\rho_{ij}))$ were introduced between V–O (five bonds per vanadium ion), Na–O (eight bonds per sodium ion) and O–O neighboring ions at interionic distances less than 0.325 nm. Because of the nonequivalence of the oxygen O1, O2 and O3 ions we have to introduce different potentials for different types of bonding. In the initial step we confined ourselves to just ten fitting parameters (instead of the 34 independent force constants for the V$_2$O$_5$ lattice in Ref. 22) including ion charges $Z(\mathrm{A})$ (a condition of lattice neutrality brings the relation $Z(\mathrm{Na})+Z(\mathrm{O1})+2Z(\mathrm{V})+2Z(\mathrm{O2})+2Z(\mathrm{O3})=0$ about) and C_{ij},ρ_{ij} constants for V–O, Na–O and O–O pairs of ions. Starting values of the parameters were taken from the lattice dynamics simulations of TmVO$_4$,[27] LuPO$_4$[28] and NaNO$_3$.[29]

An orthogonal transformation of the atomic displacements to symmetrized and normalized linear combinations, namely,

$$u_\alpha(\Gamma_{1u},\mathrm{A})=\frac{1}{2}[u_{1\alpha}(\mathrm{A})+u_{2\alpha}(\mathrm{A})+u_{3\alpha}(\mathrm{A})+u_{4\alpha}(\mathrm{A})],$$

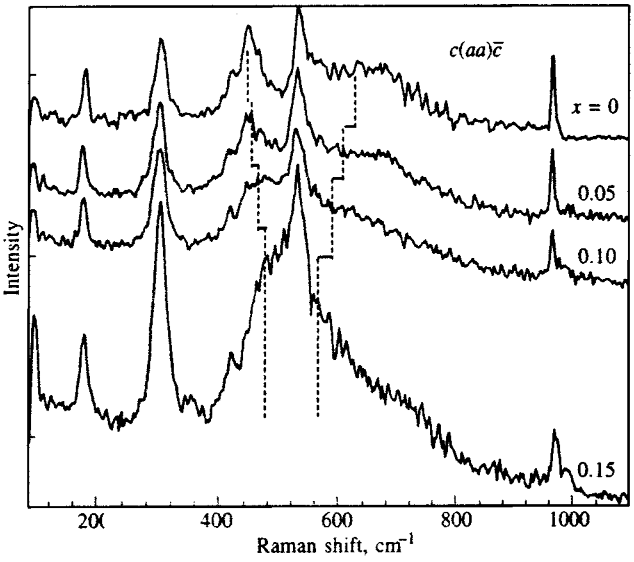

FIG. 7. Room-temperature Raman spectra of Na$_{1-x}$V$_2$O$_5$ for various x.

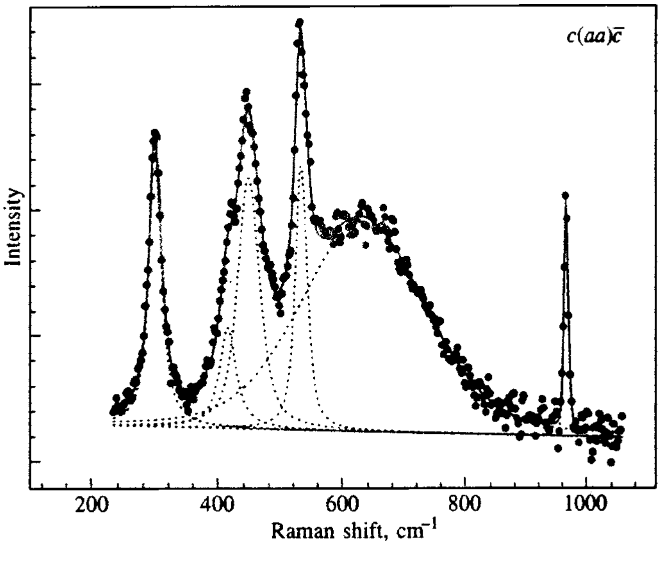

FIG. 8. Expansion of the room temperature $c(aa)\tilde{c}$ Raman spectrum of NaV$_2$O$_5$ into individual profiles (dashed lines). The sum of these profiles, shown by a solid line, approximates the experimental spectrum (circles) well.

$$u_\alpha(\Gamma_{1g},\mathrm{A})=\frac{1}{2}[-u_{1\alpha}(\mathrm{A})-u_{2\alpha}(\mathrm{A})+u_{3\alpha}(\mathrm{A})+u_{4\alpha}(\mathrm{A})],$$

$$u_\alpha(\Gamma_{2u},\mathrm{A})=\frac{1}{2}[-u_{1\alpha}(\mathrm{A})+u_{2\alpha}(\mathrm{A})-u_{3\alpha}(\mathrm{A})+u_{4\alpha}(\mathrm{A})],$$

$$u_\alpha(\Gamma_{2g},\mathrm{A})=\frac{1}{2}[u_{1\alpha}(\mathrm{A})-u_{2\alpha}(\mathrm{A})-u_{3\alpha}(\mathrm{A})+u_{4\alpha}(\mathrm{A})],$$

$$u_\alpha(\Gamma_{1u},\mathrm{B})=\frac{1}{\sqrt{2}}[u_{1\alpha}(\mathrm{B})+u_{2\alpha}(\mathrm{B})],$$

$$u_\alpha(\Gamma_{1g},\mathrm{B})=\frac{1}{\sqrt{2}}[-u_{1\alpha}(\mathrm{B})+u_{2\alpha}(\mathrm{B})],$$

where A, denotes V, O2, or O3 ions, and B denotes Na or O1 ions, divides the dynamical matrix at the Γ-point into blocks corresponding to irreducible representations of the crystal factor-group. Here,

$$\Gamma_{1u}=B_{3u},\quad \Gamma_{1g}=B_{2g},\quad \Gamma_{2u}=B_{1u},$$
$$\Gamma_{2g}=A_g \quad \text{for } \alpha=x,$$
$$\Gamma_{1u}=B_{2u},\quad \Gamma_{1g}=B_{3g},\quad \Gamma_{2u}=A_u,$$
$$\Gamma_{2g}=B_{1g} \quad \text{for } \alpha=y,$$
$$\Gamma_{1u}=B_{1u},\quad \Gamma_{1g}=A_g,\quad \Gamma_{2u}=B_{3u},$$
$$\Gamma_{2g}=B_{2g} \quad \text{for } \alpha=z.$$

Comparing the calculated eigenvalues of the dynamical matrix with the measured frequencies of the lattice normal modes, we varied the parameters Z, C_{ij}, ρ_{ij} step-by-step in a physically motivated direction (e.g., absolute values of Z were diminished to account for the observed maximum LO–TO splittings) with the aim of achieving a better agreement with the measured frequencies at the Brillouin zone center. Simultaneously, frequencies of normal modes at zone boundaries and with the wave vectors **k** close to the Γ point (acoustical modes) were controlled.

The rigid-ion model presents a very crude approximation to the charge distribution in covalent compounds, and the optimization of the model parameters was terminated when achieving real values for the lattice normal modes frequencies throughout the Brillouin zone. The final values of the effective ionic charges were $Z(\mathrm{V})=2.405$, $Z(\mathrm{Na})=0.83$, $Z(\mathrm{O1})=-1.22$, $Z(\mathrm{O2})=-1.23$, $Z(\mathrm{O3})=-0.98$ (in units of the proton charge), which are close to corresponding effective charges for vanadium $Z(\mathrm{V})=3$ and oxygen $Z(\mathrm{O})=-1.5$ in $\mathrm{TmVO_4}$,[27] phosphorus $Z(\mathrm{P})=2.33$ and oxygen $Z(\mathrm{O})=-1.19$ in $\mathrm{LuPO_4}$,[28] and sodium $Z(\mathrm{Na})=0.87$ in $\mathrm{NaNO_3}$.[29] The apical oxygen ion O3, closest to the vanadium ion, has the lowest charge due to its having the strongest covalent binding.

The calculated frequencies of the lattice normal modes at the Brillouin zone center are presented in Table I and II for infrared and Raman active modes. The calculated frequencies of silent A_u modes are 120, 167 and 572 cm^{-1}. The acoustic properties of the lattice are defined by nine elastic constants, the predicted values of $C_{11}=17.7$, $C_{12}=9.7$, $C_{22}=23.6$ (in units of 10^{10} N/m^2) are less dependent on variations of the model parameters.

The measured components of the high-frequency dielectric tensor ε_∞ differ appreciably from unity (see Table I), so neglect of electronic polarization is a very crude approximation in this case—in particular, when estimating LO–TO splittings at the Γ point. However, for most of the infrared active normal modes, our model yields a satisfactory description of the longitudinal macroscopic electric field induced by the vibrations of ions. Very strong damping of the B_{1u} TO mode at 591 cm^{-1} may be the reason for the large difference between the calculated and measured LO–TO splitting (see Table I); in the case of the B_{2u} TO mode at 365 cm^{-1}, our model yields a greatly overestimated frequency of the corresponding LO mode. Large discrepancies between several calculated frequencies of the Raman active modes and the experimental data (see Table II) clearly demonstrate that some significant interactions—in particular, three-body forces, which strongly affect the frequencies of bending vibrations—are to be included in more thorough study of the lattice dynamics of this system.

Perhaps, the most interesting result of this analysis of the $\mathrm{NaV_2O_5}$ lattice dynamics is the predicted soft mode behavior of the transverse acoustic mode at the Brillouin zone boundary [with the wave vector $\mathbf{k}_0=\pi(0,0,1/c)$], polarized in the *ac*-plane. Due to the competition between long-range Coulomb and short-range non-Coulomb forces, the corresponding branch of the vibrational spectrum moves to the range of imaginary frequencies when approaching the $\mathbf{k}_0$ point thus making it possible to consider the $\mathrm{NaV_2O_5}$ crystal an improper virtual ferroelastic. To stabilize the lattice against $\mathbf{k}_0$-excitations, we had to introduce an attractive interaction between neighboring $\mathrm{V_1}$ and $\mathrm{V_2}$ ($\mathrm{V_3}$ and $\mathrm{V_4}$) ions along the *a*-axis with the significant bending force constant of approximately 5 N/m. Charge ordering in the subsystem of V ions can destroy the balance between forces of opposite sign and induce freezing of the soft-mode atomic displacements (the unit cell doubles in the *c* direction, the neighboring layers shift in opposite directions, and in each layer the right and left legs of the vanadium ladders become nonequivalent, due to shifts of $\mathrm{V_1}$–$\mathrm{V_2}$ and $\mathrm{V_3}$–$\mathrm{V_4}$ rungs along opposite directions in the *ac*-plane) as a precursor of the subsequent magnetic ordering with doubling of a unit cell in *a*-, *b*- and *c*-directions.

4. DISCUSSION

4.1. Symmetry group of $\mathrm{NaV_2O_5}$

Table III summarizes the observed vibrational modes together with their interpretation both in centrosymmetric D_{2h}^{13} and noncentrosymmetric C_{2v}^{7} groups. While the former group explains naturally the experimental data provided one Raman and three infrared frequencies remain undetected, the latter group leads to an assumption that 22 of 45 expected Raman and 23 of 37 expected infrared modes were not detected. Moreover, only three frequencies (90, 174, and 951

TABLE III. Comparison of experimentally observed Raman and infrared modes with the expected ones within centrosymmetric D_{2h}^{13} and noncentrosymmetric c_{2v}^{7} space groups (mode frequencies are in cm^{-1}).

$Pmmn(D_{2h}^{13})$	Observed modes	$Pmn2_1(C_{2v}^{7})$
$8A_g(aa,bb,cc)$	90 177 233 301 420 447 533 970	$15A_1(aa,bb,cc;\mathbf{E}\|\|\mathbf{a})$
$7B_{3u}(\mathbf{E}\|\|\mathbf{a})$	91[a] 145 251 526 939	
$3B_{1g}(ab)$	174 295 683	$7B_1\ (ab;\mathbf{E}\|\|\mathbf{b})$
$4B_{2u}(\mathbf{E}\|\|\mathbf{b})$	175 225[a] 367 582	
$8B_{2g}(ac)$	141 186 225 392 429 550 951	$15B_2\ (ac;\mathbf{E}\|\|\mathbf{c})$
$7B_{1u}(\mathbf{E}\|\|\mathbf{c})$	162 179 468 591 760 955	
$5B_{3g}(bc)$	169 257 366 418 683	$8A_2\ (bc)$
$3A_u$		

[a] Observed below 200 K.

cm^{-1}) coincide to within the experimental accuracy (±4 cm^{-1}) in the sets of Raman and infrared modes corresponding to a given irreducible representation of the noncentrosymmetric group, whereas all the modes should be both Raman and infrared active in that case. We also note once more that we failed to obtain a realistic set of force constants when carrying out the lattice dynamics calculations in the assumption C_{2v}^{7} noncentrosymmetric space group.

We consider our Raman and infrared data, and the results of lattice dynamics calculations, to support strongly the conclusion of the previous structural studies[13,14] that the space group of NaV_2O_5 above $T_c = 35\,K$ is the centrosymmetric D_{2h}^{13} rather than noncentrosymmetric C_{2v}^{7} group. From the point of view of D_{2h}^{13} group it is also easy to explain the results of a recent ^{51}V-NMR study[16] that revealed only one vanadium position at elevated temperatures.

4.2. Atomic displacements

As we have mentioned in the Introduction, the strucutre of NaV_2O_5 looks like the structure of V_2O_5 intercalated with Na. The V–O bond lengths within the vanadium-oxygen layers are close in these two compounds (see Table IV). The longest bond within the layer interconnects two V_2O_5 units in the crystal unit cell (see Fig. 1b where this bond is indicated by a dashed line). Thus, it makes a certain amount of sense to classify the $k=0$ crystal vibrations on the basis of internal vibrations of the V_2O_5 «molecule» (C_{2v} point symmetry group), split into Davydov doublets of the D_{2h} factor group by an interaction between two «molecules» in the crystal unit cell. $(A_g + B_{1u})$, $(B_{2g} + B_{3u})$, $(B_{3g} + B_{2u})$, and $(B_{1g} + A_u)$ Davydov doublets come, respectively, from A_1, B_1, B_2, and A_2 vibrations of the V_2O_5 ''molecule.'' Splittings of these doublets can be as great as 100 cm^{-1} due to Coulomb interactions and, in particular, due to interactions between adjacent V ions via common neighbors (O2 ions) along the chains (see Fig. 1). Many of the vibrational frequencies of NaV_2O_5 are close to those of V_2O_5.[22]

A comparison of our observed vibrational frequencies with those of V_2O_5[22] and with the results of our calculations leads to the following assignment of the vibrational modes of NaV_2O_5. The V–O3 stretching modes are manifested by two Davydov doublets: $951\,cm^{-1}(B_{2g}) + 939\,cm^{-1}(B_{3u})$ and $970\,cm^{-1}(A_g) + 955\,cm^{-1}(B_{1u})$. The Davydov splittings are relatively small in this case, indicating that these vibrations associated with the strongest bond V–O3 are really well localized. The mode frequencies are somewhat lower than the corresponding frequencies in V_2O_5 (976, 982, 994, and 975 cm^{-1}), which is consistent with longer V–O3 bonds in NaV_2O_5 in comparison with V_2O_5. The following vibrations are associated with the O1–V–O3 bending modes: $177\,cm^{-1}(A_g)$, $162\,cm^{-1}(B_{1u})$, $392\,cm^{-1}(B_{2g})$, $366\,cm^{-1}(B_{3g})$.

The bridging oxygens O1 participate in V–O1–V bending vibrations $418\,cm^{-1}(B_{3g})$, $367\,cm^{-1}(B_{2u})$,

TABLE IV. The bond lengths (nm) in NaV_2O_5 and V_2O_5.

Bond	NaV_2O_5 Ref. 13	V_2O_5 Ref. 22
V–O3	0.161	0.158
V–O1	0.183	0.177
(V–O2)×2	0.192	0.188
V–O2′	0.199	0.202
V–O3′	0.318	0.278

447 $\mathrm{cm}^{-1}(A_g)$, and 468 $\mathrm{cm}^{-1}(B_{1u})$. The V–O1–V stretching vibration is located at 420 $\mathrm{cm}^{-1}(A_g)$ and mainly involves the motion of the V atoms along the *a*-axis.

The modes at 683 cm^{-1} (B_{3g} and B_{1g}) and 582 $\mathrm{cm}^{-1}(B_{2u})$ correspond to V–O2 stretching vibrations along the *b*-axis, while those at 550 $\mathrm{cm}^{-1}(B_{2g})$, 533 $\mathrm{cm}^{-1}(A_g)$, and 526 $\mathrm{cm}^{-1}(B_{3u})$ correspond to the bending vibrations.

Most of the remaining modes can be described in terms of external modes of the V_2O_5 units. Thus, the modes at 186 $\mathrm{cm}^{-1}(B_{2g})$, 169 $\mathrm{cm}^{-1}(B_{3g})$ and 90 $\mathrm{cm}^{-1}(A_g)$, correspond to the relative translations of the two V_2O_5 units within the crystal unit cell along the *a*-, *b*-, and *c*-axes, respectively. As the V_2O_5 units are bind along the *b*-axis, these modes can be considered relative translations of neighboring $(VO_5)_n$ chains. The B_{1g} mode at 174 cm^{-1} (O3 ions move along the *b*-axis) and the A_g mode at 301 cm^{-1} (O2 ions move along the *c*-axis) correlate with in-plane and out-of-plane chain bending vibrations, respectively. The B_{2g} mode at 141 cm^{-1} and B_{3u} mode at 91 cm^{-1} are associated with rotation of the chains around the *b*-axis.

Modes that involve mainly displacements of Na atoms are at 225 $\mathrm{cm}^{-1}(B_{2u})$, 251 $\mathrm{cm}^{-1}(B_{3u})$, and 179 $\mathrm{cm}^{-1}(B_{1u})$.

4.3. Spectra of electron excitations

With our derived values of the effective charges, we estimated the crystal field energies of the $3d$ electron localized at a V^{4+} ion site. The crystal field parameters $B_2^0 = 1360 + 2090G$, $B_2^1 = 2020 - 1590G$, $B_2^2 = 820 + 640G$, $B_4^0 = 610 + 1430G$, $B_4^1 = -1810 - 3690G$, $B_4^2 33 + 144G$, $B_4^3 = 3650 + 8180G$, $B_4^4 = 4070 + 7420G\ \mathrm{cm}^{-1}$ for the V_1 and V_3 sites were calculated in the framework of the exchange charge model[30] (for the V_2 and V_4 sites B_p^1 and B_p^3 parameters change signs; the first terms correspond to point charge contributions, and Stevens normalization is used). The scaling factor G determines the strength of the exchange charge field. We estimated this phenomenological parameter of the model ($G = 4$) by fitting the total crystal field splitting to the width of the V-$3d$ bands presented in Ref. 13. In this case, the effective crystal field provides the following energy level pattern of the V^{4+} ion: $0(A'')$, $1.10(A')$, $1.18(A'')$, $3.39(A')$ and $4.78(A')$ eV (irreducible representations of the C_s point group, corresponding to the space symmetry of the electron wave function, are given in brackets, additional shifts of the crystal field levels due to the spin-orbit interaction and the electrostatic field of a hole at the neighboring vanadium site are less than 0.025 eV). The ground state wave function is the d_{xy} orbital with small admixture of the d_{yz} orbital, as previously pointed out,[13] and the sequence of the excited states is in agreement with band structure calculations as well.[13]

Strong absorption of light (**k**$\|$**c**) with **E**$\|$**a** as well as with **E**$\|$**b** was observed in the region 1.0–1.2 eV.[21] Both magnetic dipole and induced electric dipole $d-d$ transitions in the odd crystal field are allowed between the A$''$ states for **E**$\|$**a**, and between the A$''$ and A$'$ states for **E**$\|$**b**. Thus, in accordance with the results on crystal field energies given above, the observed broad optical bands in the vicinity of 1 eV can be interpreted as phonon assisted $d-d$ transitions without any additional suppositions about the broken symmetry between the legs of vanadium ladders.[31]

The next step towards the detailed description of the spectra of electron excitations is to construct molecular orbitals for the $[V_2O]^{7+}$ ''molecule,'' which has C_{2v} point symmetry, using vanadium *d*-orbitals and oxygen *p*-orbitals. The vanadium ground state wave function d_{xy} yields the nonbonding a_2 orbital as well as bonding B_2 and antibonding b_2^* molecular orbitals, namely,

$$a_2:[d_{xy}(\mathrm{L}) + d_{xy}(\mathrm{R})], \quad b_2:[d_{xy}(\mathrm{L}) - d_{xy}(\mathrm{R})] + p_y,$$

$$b_2^*:[d_{xy}(\mathrm{L}) - d_{xy}(\mathrm{R})] - p_y.$$

Here, a_2, b_2 denote irreducible representations of the C_{2v} point group and L and R denote vanadium sites on the left and right sides of a ladder rung. The highest filled orbital being a_2, the $a_2 \to b_2^*$ electronic transition allowed in **E**$\|$**a** polarization can account for the low-frequency absorption band observed only in this polarization. Quantum-chemical calculations are necessary to varify this qualitative interpretation.

4.4. Fano resonances with a continuum

The asymmetric line shapes of the infrared active modes at 91, 150, 939 cm^{-1} in **E**$\|$**a** polarization highlight the strong interference between these modes and a continuum, observed just in this polarization. This interpretation is supported by the fact that the spectral line near 91 cm^{-1} becomes perfectly symmetric when the continuum absorption vanishes in this spectral range below the phase transition temperature $T_c = 35\,\mathrm{K}$. In our earlier work[8] we argued that these changes are related to the opening of a gap in the magnetic excitation spectrum at T_c, the observed continuum being due to two-magnon absorption.

However, such a straightforward interpretation is no longer valid in the case of the space group D_{2h}^{13}. It must be revised, taking into account possible electronic excitations in this frequency range, as discussed in the previous section, and charge ordering at the transition temperature.

4.5. Higher order infrared vibrational spectra

Two- and three-phonon absorption results from anharmonicity of crystal vibrations. It is continuous, displaying singularities corresponding to critical points of the Brillouin zone. Leaving the detailed analysis of multiphonon bands to the future, we discuss here only sharp lines observed in **E**$\|$**c** absorbance spectrum (Fig. 5). They are listed in Table V together with their tentative assignment, using symmetry-allowed combinations of Γ-point phonons observed in our first-order spectra. The coincidence of the observed and combinational frequencies lies within the accuracy of our measurements.

The strongest narrow peak at 1930 cm^{-1} corresponds, according to this assignment, to sum of the components of the Davydov doublet originating from the V–O3 stretching vibration. This stretching mode is well localized, which re-

TABLE V. Multiphonon bands in NaV_2O_5 observed in $E \| c$ polarization.

Observed bands, cm^{-1}	Combination of phonons, cm^{-1}
1072 (B_{1u})	550 (B_{2g}) + 526(B_{3u}) = 1076(B_{1u})
1270 (B_{1u})	683 (B_{3g}) + 582(B_{2u}) = 1265(B_{1u})
1930 (B_{1u})	970 (A_g) + 955(B_{1u}) = 1925(B_{1u})
2858 (B_{1u})	3 × 955(B_{1u}) = 2865(B_{1u})
2901 (B_{1u})	2 × 970(A_g) + 955(B_{1u}) = 2895(B_{1u})

sults in its small dispersion over the Brillouin zone, thus delivering a narrow two-phonon band, in accordance with the experimental observation.

5. SUMMARY

We have performed a thorough spectroscopic study of far infrared reflectance and transmittance, along with Raman scattering of α'-NaV_2O_5 single crystals in the high-temperature phase (above $T_c = 35\,K$). Far infrared spectra were obtained for **E**‖**a**, **E**‖**b**, and **E**‖**c** polarizations of incident light. Diagonal (*aa*), (*bb*), (*cc*) and off-diagonal (*ab*), (*bc*), (*ac*) components of the Raman scattering tensor were investigated. We report five infrared active modes in **E**‖**a** polarization, four in **E**‖**b** polarization, and six in **E**‖**c** polarization. Eight Raman active modes have been detected for parallel polarizations of incident and scattered light (*aa*), (*bb*), (*cc*). The (*ab*), (*ac*) and (*bc*) Raman geometries delivered three, seven, and five modes, respectively. These results are in much better agreement with the recently proposed centrosymmetric space group $D_{2h}^{13}(Pmmn)$ for the high-temperature phase of NaV_2O_5 than with the previously adopted noncentrosymmetric space group $C_{2v}^{7}(Pmn2_1)$. We have also carried out the lattice dynamics calculations, based on the rigid ion model for both structures of NaV_2O_5 proposed in the literature. We failed to obtain any physically well-grounded set of parameters providing a stable C_{2v}^{7} lattice structure. Thus, our infrared and Raman experimental data, along with the results of lattice dynamics calculations, strongly support the conclusion of the previous structural study[13,14] that the space group of NaV_2O_5 above $T_c = 35\,K$ is the centrosymmetric D_{2h}^{13} rather than noncentrosymmetric C_{2v}^{7} group.

This conclusion leads to important physical consequences. In particular, it requires a revised interpretation of one-dimensional magnetic properties of NaV_2O_5 and of the phase transition at 35 K, previously considered as an ordinary spin-Peierls transition. The interpretation of the previously observed broad bands in near and far infrared absorption[8,21] needs to be reconsidered as well.

Using the effective charges derived via lattice dynamics calculations, and fitting the total crystal field splitting to the width of the V-3*d* bands,[12] we estimated the crystal field energies of the 3*d* electron localized at the vanadium site. It follows from this estimate that the observed[21] near infrared broad band absorption of NaV_2O_5 can be interpreted as phonon-assisted *d*–*d* transitions. We speculate that the far infrared **E**‖**a** polarized absorption continuum might be associated with electron excitations of $[V_2O]^{7+}$ rungs in a crystal field of C_{2v} symmetry.

Strongly asymmetric spectral lines observed in **E**‖**a** absorbance spectra of NaV_2O_5 highlight a strong interference between relatively narrow phonon lines and the underlying continuum. This suggest an interaction between crystal vibrations and magnetic or electronic excitations. The detailed physical interpretation of the observed phenomenon depends on the nature of the far infrared **E**‖**a** polarized continuum, which requires special investigation.

In conclusion, we reported also some preliminary results on higher-order vibrational spectra of NaV_2O_5 resulting from anharmonicity of lattice vibrations.

After having submitted this paper for publication (see Ref. 32) we became aware of a similar investigation.[33,34] Experimental data presented in these papers are in good agreement with our results.

We are grateful to A. I. Smirnov for checking the samples by ESR measurements, to A. N. Vasil'ev for stimulating discussions, and to G. N. Zhizhin for sustained support of this research. This work was made possible in part by Grant No. 98-02-17620 from the Russian Fund for Fundamental Research.

*)E-mail: popova@isan.troitsk.ru

---

[1] I. M. Isobe and Y. Ueda, J. Phys. Soc. Jpn. **65**, 1178 (1996).
[2] A. I. Buzdin and L. N. Bulaevskii, Usp. Fiz. Nauk **131**, 495 (1980) [Sov. Phys. Usp. **23**, 409 (1980)].
[3] J. Hemberger, M. Lohmann, N. Nickloas, A. Loidl, M. Klemm, G. Obermeier, and S. Horn, Europhys. Lett. **42**, 661 (1998); T. Yamada, private communication.
[4] K. Kobayashi, T. Mizokawa, A. Fujimori, M. Isobe, and Y. Ueda, Phys. Rev. Lett. **80**, 3121 (1998).
[5] Y. Fujii, H. Nakao, T. Yoshihama, M. Nishi, K. Nakajima, K. Kakurai, M. Isobe, Y. Ueda, and H. Sawa, J. Phys. Soc. Jpn. **66**, 326 (1997).
[6] M. Weiden, R. Hauptmann, C. Geibel, F. Steglich, M. Fisher, P. Lemmens, and G. Guntherodt, Z. Phys. B **103**, 1 (1997).
[7] H. Curoe, H. Seto, J. Sasaki, T. Sekine, M. Isobe, and Y. Ueda, submitted to J. Phys. Soc. Jap.; E-prints archive cond-mat/9805251.
[8] M. N. Popova, A. B. Sushkov, A. N. Vasil'ev, M. Isobe, and Y. Ueda, JETP Lett. **65**, 743 (1997); E-prints archive cond-mat/9711052.
[9] D. Smirnov, P. Millet, J. Leotin, D. Poilblanc, J. Riera, D. Augier, and P. Hansen, Phys. Rev. B **57**, R11035 (1998).
[10] T. Yoshihama, M. Nishi, K. Nakajima, Y. Fujii, M. Isobe, and Y. Ueda, Physica B **234-236**, 539 (1997).
[11] A. Carpy and J. Galy, Acta Crystallogr. B **31**, 1481 (1975).
[12] H. G. Backman, F. R. Ahmed, and W. H. Barnes, Z. Krist. **115**, 110 (1961).
[13] H. Smolinski, C. Gros, W. Weber, U. Peuchert, G. Roth, M. Weiden, and C. Geibel, Phys. Rev. Lett. **80**, 5164 (1998).

[14] A. Meetsma, J. L. de Boer, A. Damascelli, T. T. M. Palstra, J. Jegoudez, and A. Revcolevschi, Acta Cryst. C. **54**, 1558 (1998).
[15] P. Horsch and F. Mack, E-prints archive cond-mat/9801316.
[16] T. Ohama, H. Yasuoka, M. Isobe, and Y. Ueda, Phys. Rev. B **59**, 3299 (1999).
[17] H. Seo and H. Fukuyama, E-prints archive cond-mat/9805185.
[18] M. V. Mostovoy and D. I. Khomskii, E-prints archive cond-mat/9806215.
[19] P. Thalmeier and P. Fulde, E-prints archive cond-mat/9805230.
[20] M. Köppen, D. Pankert, R. Hauptmann, M. Lang, M. Weiden, C. Geibel, and F. Steglich, Phys. Rev. B **57**, 8466 (1998).
[21] S. A. Golubchik, M. Isobe, A. N. Ivlev, B. N. Mavrin, M. N. Popova, A. B. Sushkov, Y. Ueda, and A. N. Vasil'ev, J. Phys. Soc. Jpn. **66**, 4042 (1997).
[22] L. Abello, E. Husson, Y. Repelin, and G. Lucazeau, Spectrochem. Acta **39A**, 641 (1983).
[23] M. Isobe, C. Kagami, and Y. Ueda, J. Cryst. Growth **181**, 314 (1997).
[24] B. N. Mavrin, Optika i Spectroscopiya **49**, 79 (1980).
[25] D. L. Rousseau, R. P. Bauman, and S. P. S. Porto, J. Raman Spectrosc. **10**, 253 (1981).
[26] U. Fano, Phys. Rev. **124**, 1866 (1961).
[27] V. R. Pekurovskii, Izvestiya AN USSR, Ser. Fiz. **50**, 324 (1986).
[28] J. C. Nipko, C.-K. Loong, M. Loewenhaupt, M. Braden, W. Reichardt, and L. A. Boatner, Phys. Rev. B **56**, 11584 (1997).
[29] A. Yamamoto, T. Utida, H. Murata, and Y. Shiro, J. Phys. Chem. Solids **37**, 693 (1976).
[30] B. Z. Malkin, in Spectroscopy of solids containing rare-earth ions, ed. by A. A. Kaplyanskii and R. M. Macfarlane, Elsevier Science PB, Amsterdam (1987), pp. 13–49.
[31] A. Damascelli, D. van der Marel, J. Jegoudez, G. Dhalenne, and A. Revcolevschi, E-prints archive cond-mat/9806222.
[32] M. N. Popova, A. B. Sushkov, S. A. Golubchik, B. N. Mavrin, V. N. Denisov, B. Z. Malkin, A. I. Iskhakova, M. Isobe, and Y. Ueda, E-prints archive cond-mat/9807369.
[33] Z. V. Popović, M. J. Konstantinović, R. Gajić, V. Popov, Y. S. Raptis, A. N. Vasil'ev, M. Isobe, and Y. Ueda, J. Phys.: Condens. Matter **10**, L513 (1998).
[34] D. Smirnov, J. Leotin, P. Millet, J. Jegoudez, and A. Revcolevschi, E-prints archive cond-mat/9808006.

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

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Behavior of the $p-T$ phase diagram of the organic conductor $(ET)_4Hg_3I_8$

A. V. Kornilov and V. M. Pudalov

Institute of High-Pressure Physics, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia;
P. N. Lebedev Institute of Physics, Russian Academy of Sciences, 117924 Moscow, Russia

A. P. Kochkin

Institute of High-Pressure Physics, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

R. N. Lyubovskaya and R. B. Lyubovskii*)

Institute for Problems of Chemical Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 21 September 1998)
Zh. Éksp. Teor. Fiz. **115**, 2190–2196 (June 1999)

The conductivity of single crystals of the organic conductor $(ET)_4Hg_3I_8$ [ET-bis(ethylendithio)tetrathiafulvalene] has been investigated at temperatures from 4.2 to 360 K and pressures of up to 75 kbar. Two first-order phase transitions have been detected at room temperature at pressures of 2.75 and 6.7 kbar. On the basis of the experimental data, the $p-T$ phase diagram for the first-order phase transitions has been plotted. The unusual shape of the phase diagram (a slow monotonic growth of the transition temperature with a slope $dT/dp = 4$ deg/kbar followed by a sharp drop around the point $p_0 = 6.5$ kbar, $T_0 = 324$ K) has been analyzed using the Landau theory of second-order phase transitions. Our analysis supports the hypothesis of a second-order phase transition around this point and also exhibits satisfactory agreement between calculations and the experimental curves of the first-order phase transitions. © *1999 American Institute of Physics.* [S1063-7761(99)02106-X]

1. INTRODUCTION

Most low-dimensional organic conductors (salts) are of interest to researchers because small changes in external parameters (temperature, pressure, and magnetic field) can have a considerable effect on their properties and behavior. These changes result in phase transitions leading to new states of materials. As a rule, these are second-order transitions, where the first derivatives of entropy or volume are discontinuous. These transitions give rise to charge-density or spin-density waves in the system of charge carriers, order–disorder or metal–insulator transformations, which often compete with superconducting transitions.[1,2] Much less often low-dimensional conductors undergo first-order transitions, in which first derivatives of the basic thermodynamic functions (thermodynamic potential, free energy, etc.) have jumps. The features of the first-order transitions in the materials under discussion are hysteresis loops on curves of thermodynamic functions plotted against pressure or temperature, and these transitions are structural metal–insulator transformations[3] [for example, in the $MEM(TCNQ)_2$ salt, where MEM is methylethylmorpholinium and TCNQ is tetracyanchinodimethane], semiconductor–semiconductor transitions [in the $MTPP(TCNQ)_2$ salt, where MTPP is methyltriphenylphosphonium], or metal–metal transitions [in $(BEDO\text{-}TTF)_2ReO_4\cdot H_2O$, where BEDO-TTF is bis-(ethylendioxy)tetrathiafulvalene]. Synthesis of organic conductors based on ET and TMTSeF (thetramethyteraseleniumfulvalene) allows one to produce salts in which all these states can be observed.

One of the most interesting types of salts for researchers of organic conductors is $ET_4Hg_{3-\delta}X_8$, where X=Cl, Br, and I and $\delta = 0.22$, 0.11, and 0, respectively.[6] The salt with composition $(ET)_4Hg_{2.78}Cl_8$ is an organic metal at atmospheric pressure and all temperatures. At higher pressures and low temperatures, it is a semiconductor, and at $p = 12$ kbar it goes over to a superconducting state with $T_c = 1.8$ K.[7] The salt with composition $ET_4Hg_{2.89}Br_8$ is a superconductor with $T_c = 4.3$ K at atmospheric pressure. Its T_c changes with pressure in an unusual manner ($dT_c/dP > 0$), and it becomes an insulator at $p > 25$ kbar.[8] All the salts listed above are isostructural, and their anion and cation sublattices are incommensurate. The third salt of this group, $(ET)_4Hg_3I_8$, has the structure and properties notably different from those of the former two. At atmospheric pressure and room temperature it is a semiconductor which undergoes a first-order transition to the insulator state at $T = 260$ K.[9]

The present investigation is concerned with the phase diagram of $(ET)_4Hg_3I_8$ at pressures of up to 26 kbar and temperatures ranging between 4.2 K and 360 K, which was studied by measuring electric resistivity. At room temperature measurements have been conducted at pressures of up to 75 kbar.

2. EXPERIMENT

$(ET)_4Hg_3I_8$ crystals were fabricated by electrochemical oxidation of ET in tetrahydrofuran in the presence of the electrolyte $(Bu_4N)_2Hg_3I_8$. The crystal lattice of $(ET)_4Hg_3I_8$

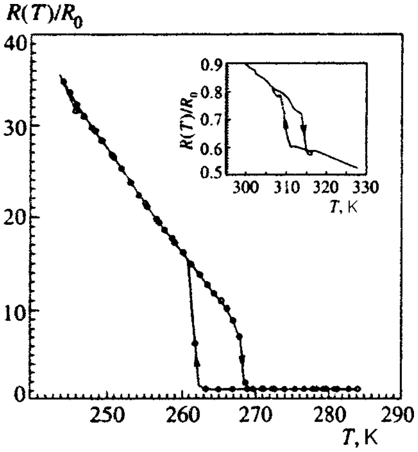

FIG. 1. Resistance of a $(ET)_4Hg_3I_8$ single crystal versus temperature at atmospheric pressure. The inset shows the hystersis loop for phase transition I↔II at $p \approx 3.9$ kbar and $T = 310$ K.

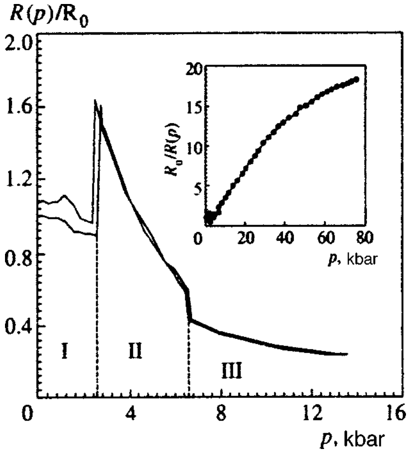

FIG. 2. Resistance and conductivity (inset) of $(ET)_4Hg_3I_8$ single crystal versus pressure at room temperature.

is composed of cation–radical layers formed by piles of two independent ET molecules and $Hg_3I_8^{2-}$ anion layers. The mercury atoms are inside I_4 tetrahedra, and the occupancy of mercury sites is 0.5.

The conductivity measured at room temperature in the **ab** plane of the conducting layer is 0.3–2 $(\Omega \cdot cm)^{-1}$. The conductivity measured in the normal direction demonstrates an anisotropy factor of $\sim 10^3$, which is typical of low-dimensional organic conductors. In our experiments, we tested crystals of the same batch; measurements of samples from other batches have not revealed notable differences in the phase diagrams. In our measurements of conductivity, we have used several devices of different types, depending on the pressure. At pressures up to 75 kbar and at room temperature we used a toroidal cell made of limestone.[10] In measurements at a variable temperature, the toroid was placed in a squirrel-wheel apparatus,[10] where a pressure of up to 26 kbar could be fixed to within ± 2 kbar and the temperature lowered to 4.2 K. Measurements at temperatures of 4.2–360 K and pressures of up to 15 kbar were conducted in a high-pressure chamber formed by a cylinder and a piston.[11] The pressure in this chamber was measured using a manganin wire with an precision of 0.03 kbar, and the temperature was measured by a Cu–CuFe thermocouple to within 0.5 K. The phase diagram was plotted using the following procedure: the pressure was chosen at room temperature, after that the resistivity was measured as a function of temperature. The phase transition point was determined by detecting a jump in the resistivity in the process of cooling (Fig. 1), where the pressure was measured by the manganin probe.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The temperature dependence of the resistance of the $(ET)_4Hg_3I_8$ single crystal measured at atmospheric pressure in the direction of the *b*-axis is plotted in Fig. 1. One can see that the salt undergoes a semiconductor–insulator phase transition at $T = 260$ K. At this temperature, the resistance abruptly increases by 1–1.5 orders of magnitude. The activation energy above the transition point corresponds to $E_1 = 500$ K, and below the transition point we have $E_2 = 5000$ K. This is a first-order transition with a hysteresis loop ~ 9 K wide. Figure 2 shows the resistance of the $(ET)_4Hg_3I_8$ single crystal as a function of pressure measured in the cylinder–piston chamber. It is clear that the crystal resistance displays two jumps as the pressure increases, at 2.75 and 6.7 kbar, which indicates that there can be three phases in this material. Measurements performed at decreasing pressure show two hystersis loops of widths < 0.1 and 0.25 kbar. The inset to Fig. 2 plots the single crystal conductivity measured at pressures of up to 75 kbar at room temperature in the toroidal apparatus made from limestone. The conductivity gradually saturates at higher pressures, which is typical of most organic conductors, and at 75 kbar it is a 18 times higher than under normal conditions. This moderate increase in the conductivity of the organic conductor at so high a pressure indicates that the molecules of this salt are packed fairly densely. It is noteworthy that there is another modification of this salt which displays no phase transition under increasing pressure and whose conductivity at 75 kbar is 300 times as high as under normal conditions.

The resistance plotted as a function of temperature (Fig. 3) indicates that the band-gap width in the phase existing at pressures $p > 6.7$ kbar, where the conductivity is of the semiconducting type, drops with the increasing pressure, therefore, it is possible that this phase undergoes a transition to a metallic state at a pressure of about 30 kbar.

The experimentally detected transition points are plotted in Fig. 4. One can see that the phase boundary monotonically shifts with the pressure up to the point with coordinates $p_0 = 6.5$ kbar and $T_0 = 324$ K, where it has a sharp bend, and then the phase boundary is an almost vertical line parallel to the temperature axis.

At normal pressure and at temperature $T = 260$ K a first-order phase transition with a conventional hysteresis loop of width $T \approx 9$ K and a sharp change in the resistance (Fig. 1) occurs in the $(ET)_4Hg_3I_8$ salt. The transition detected on the $R(T)$ curves rapidly shifts upwards with the temperature, whereas the resistance jump decreases monotonically (inset to Fig. 1) and is smeared out at pressure $p_0 = 6.5$ kbar and temperature $T_0 = 324$ K, i.e., at the point where dT/dp goes

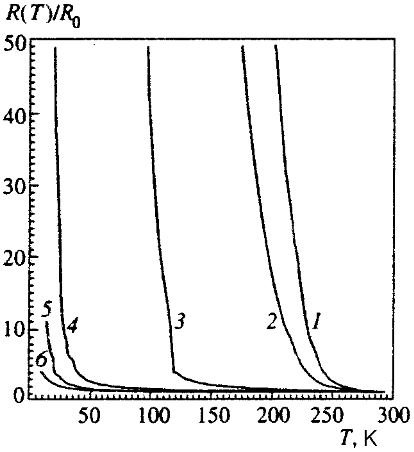

FIG. 3. Resistance measured at (curve *1*) 4, (*2*) 6, (*3*) 12, (*4*) 16, (*5*) 22, and (*6*) 26 kbar versus temperature.

to zero. The width of the hysteresis loop decreases at this point, which may be associated with a decrease in the change in volume.

One can see in Fig. 2 that in addition to the transition between phases I and II, which occurs at $p=2.75$ kbar, there is another transition at $p=6.7$ kbar between phase II and the hypothetical phase III in this salt at room temperature. The investigation of the boundary between phases II and III using curves of $R(T)$ in the pressure range 6.5–9 kbar has demonstrated that this boundary rapidly moves towatd lower temperatures with the pressure up to $p=6.7$ kbar, where $dT/dp \rightarrow -\infty$, and then returns to the region of lower pressures and temperatures. The jump in the curve of $R(T)$ in this region also has a hysteresis loop, which indicates that the transition is of first order.

This unusual shape of the $p-T$ phase diagram leads us to conclude that phases I and III are not identical. In this connection, we suppose that there should be a second-order phase transition between them, and its boundary is assumed to originate from the triple point with coordinates $p_0=6.5$ kbar and $T_0=324$ K and is directed toward higher temperatures. In measuring $R(T)$ in the range of high temperatures, we could not detect a difference between these two phases, since no resistivity jumps were observed on these curves.

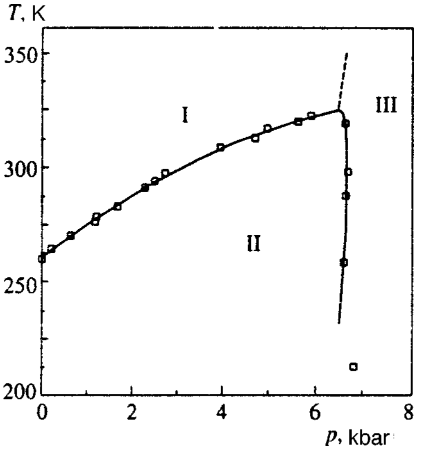

FIG. 4. Phase diagram in the $p-T$ plane for the $(ET)_4Hg_3I_8$ salt. Experimental data are plotted by symbols, solid curves show approximations described in the text. The dashed line is the boundary between phases I and III derived from calculations.

4. ANALYSIS OF THE PHASE DIAGRAM

Usually the effect of a second-order phase transition on the phase equilibrium near a first-order transition line can be neglected. This is not so, however, in the case when the jumps in the entropy and volume due to the first-order transition are small, which is consistent with the slope of dT/dp at $p<6.5$ kbar.

Hence there is good reason to try to apply the Landau theory of second-order phase transitions to interpret the curves of the first-order phase transitions I↔II and (tentatively) II↔III. This approach produced not only qualitatively adequate results, but also satisfactory quantitative agreement with the experimental data, which allows us to assert with a high degree of certainty that a line of second-order phase transitions between phases I and III really exists.

For definiteness, suppose (assuming this is consistent with shape of the curve) that phase I is symmetrical in the sense of the Landau theory (Fig. 4), and phase III is its asymmetrical modification (with a nonvanishing order parameter η).

The chemical potential of the asymmetrical phase is[12]

$$\mu_3=\mu_1+A\eta^2+B\eta^4+\ \dots, \qquad (1)$$

where $B>0$, and A near the second-order transition line can be expressed by the expansion

$$A(p,T)=a(T-T_0)+\alpha(p-p_0), \qquad (1')$$

so that we can derive from the equation $A(p,T)=0$ the temperature of the transition I↔III near the triple point as a function of pressure:

$$T_c(p)=T_0-(\alpha/a)(p-p_0). \qquad (1'')$$

After minimizing μ_3 in Eq. (1) with respect to η (Ref. 13) we have

$$\eta^2=-A/2B, \quad \mu_3=\mu_1-A^2/4B.$$

The equilibrium between phases I and II is determined by equating the chemical potentials: $\mu_1(p,T)=\mu_2(p,T)$. Differentiation yields the Clapeyron–Clausius equation and the slope of the phase boundary:

$$\frac{dT}{dp}=\frac{v_{12}}{s_{12}}.$$

Here $s_{12}=s_1-s_2$ and $v_{12}=v_1-v_2$ are the differences between the specific entropies and volumes, respectively. For the II↔III transition we obtain

$$\mu_2(p,T)=\mu_3=\mu_1(p,T)-A^2/4B.$$

In this case, the Clapeyron–Clausius equation takes the form

$$\frac{dT}{dp}=\frac{v_{12}-\alpha A/2B}{s_{12}+aA/2B}.$$

This equation can be solved, given the slopes of curves for the transitions I↔II and I↔III around T_0:

$$t_0'=\frac{v_{12}}{s_{12}},\quad t_c'=\frac{dT_c(p)}{dp}=-\frac{\alpha}{a}.$$

By introducing the variables $T-T_0=t$ and $p-p_0=x$, we transform this equation to

$$t'\equiv\frac{dt}{dx}=\frac{t_0'+t_c'(t-t_c'x)/T_m}{1+(t-t_c'x)/T_m},\tag{2}$$

where $T_m=2Bs_{12}/a^2$.

Since, as is well known,[12] the jump in the specific heat at the point of a second-order transition is

$$\Delta c_{31}=c_3-c_1=T_0\alpha^2/2B$$

(it is clear that near the triple point $T_c\rightarrow T_0$), one can easily find that

$$T_m=\frac{T_c s_{12}}{\Delta c_{31}}=\frac{\Delta q_{12}}{\Delta c_{31}},$$

where Δq_{12} is the heat of transition I↔II. Thus, T_m is effectively the ratio between the heat of the first-order transition and the specific heat jump at the neighboring second-order transition. By substituting in Eq. (2) (which is valid only for $x>0$) $y=t-t_c'x$, we obtain its solution:

$$y+y^2/2T_m=(t_0'-t_c')x+\text{const.}\tag{3}$$

Given that at $x=0$ we have $t=0$, hence $y=0$, we derive from Eq. (3) that const$=0$ and

$$t=\begin{cases}t_0'x, & x<0,\\ t_c'x-T_m+\sqrt{T_m^2-2T_m(t_c'-t_0')x}, & 0<x<x_m,\end{cases}\tag{4}$$

where $x_m=T_m/2(t_c'-t_0')$.

This solution describes the branch of the II↔III transition above (with respect to the temperature) the point where $dT/dp\rightarrow-\infty$. The second solution of Eq. (3) describes the branch of transition II↔III that tends to higher temperatures at higher pressures (below the point where $dT/dp\rightarrow-\infty$):

$$t=t_c'x-T_m-\sqrt{T_m^2-2T_m(t_c'-t_0')x},\quad x<x_m.\tag{4'}$$

This solution is meaningful if the Landau theory also applies to the region about the point where the line of transition II ↔III is vertical: $(p_0+x_m,\ T_0+t_c'x_m-T_m)$. This requires that the term $C\eta^6$ in expansion (1) for μ_3 be small, i.e., $A\ll B^2/C\sim B$ should hold (since usually $C\sim B$), and since $A=\alpha y=-\alpha T_m$, this is equivalent to the condition $T_m\ll B/a$.

Figure 4 shows the phase diagram of the material under investigation, where the solid trace for the transition I↔II is a parabola obtained by a least-square fit to the experimental data (plotted as points), and for the transition II↔III the trace shows calculations using Eqs. (4) and (4′) for $p_0=6.5$ kbar, $T_0=324$ K, $t_0'=4$ K/kbar, $t_c'=125$ K/kbar, and $T_m=47$ K.

Thus, if one accepts our interpretation of the phase diagram, it follows that the curve of the I↔III second-order phase transition must be present in the high-temperature region, and its shape is determined by Eq. (1″) with the reduced parameter $t_c'=-\alpha/a$.

As for the microscopic nature of the transition, i.e., the change in the structural symmetry, it seems natural to assume that this change is associated with the deviation of the Hg site occupancy in the anion chain from 0.5 (according to the X-ray diffraction data[9] obtained at room temperature in phase I, the anion sublattice is composed of adjacent tetrahedra with iodine atoms at their apices and central points occupied by mercury atoms with a probability of 0.5). This assumption is supported by the fact that the curves of resistance versus pressure in phases I and III (Fig. 2) almost coincide if the presence of phase II is neglected. The reason is that the conductivity is due to the charge transfer via ET organic molecules, which are not affected by the symmetry change under discussion.

A full analysis of symmetry breaking would be possible only with full X-ray diffraction data on this transition. Nonetheless, we can assert that the dimensionality of the order parameter cannot be higher than unity,[13] and this is the only condition for the validity of our interpretation.

The work was supported by the Russian Fund for Fundamental Research (Grants Nos. 96-02-18957, 97-02-17387, and 97-03-33686a), Russian Ministry of Science and Technology, NWO, and INTAS.

*⁾E-mail: rustem@icp.ac.ru

---

[1] T. Ishiguro and K. Yamaji, *Organic Superconductors*, Springer, Berlin–Heidelberg (1990).

[2] J. P. Pouget and S. Ravy, J. de Phys. I **6**, 1501 (1996).

[3] R. C. Lacoe, G. Gruner, and P. M. Chaikin, Solid State Commun. **36**, 599 (1980).

[4] A. Graja, G. Sekretarczyk, and M. Krupski, J. de Phys. **46**, 1743 (1985).

[5] S. Kalich, D. Shweitzer, J. Heinen *et al.*, Solid State Commun. **80**, 191 (1991).

[6] R. B. Lyubovskii, R. N. Lyubovskaya, and O. A. Dyachenko, J. de Phys. I **6**, 1609 (1996).

[7] R. B. Lyubovskiĭ, R. N. Lyubovskaya, R. P. Shibaeva *et al.*, JETP Lett. **42**, 468 (1985)].

[8] S. I. Bud'ko, A. G. Gapotchenko, A. E. Luppov *et al.*, Zh. Éksp. Teor. Fiz. **101**, 1841 (1992) [Sov. Phys. JETP **74**, 983 (1992)].

[9] T. Takhirov, O. N. Krasochka, O. A. Dyachenko *et al.*, Mol. Cryst. Liq. Cryst. **185**, 215 (1990).

[10] R. B. Lyubovskiĭ, R. N. Lyubovskaya, and N. V. Kapustin, JETP Lett. **66**, 1063 (1987).

[11] A. V. Kornilov, V. A. Sukhoparov, and V. M. Pudalov, in *Proc. High Pressure Science and Technology*, ed. by W. Trzeciakowski, World Sci. Publ. (1996), p. 63.

[12] L. D. Landau and M. Lifshitz, *Statistical Physics*, Part I, Pergamon Press, Oxford (1980).

[13] Yu. A. Iziumov and V. N. Syromyatnikov, *Phase Transitions and Crystal Symmetry*, Kluwer Academic, Boston (1990) [Russian orig., Nauka, Moscow (1984)].

Translation provided by the Russian editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Peculiarities of crystal field effects in $CeInCu_2$ based heavy-fermion compounds

V. N. Lazukov,*) P. A. Alekseev, E. S. Klement'ev, and I. P. Sadikov

Kurchatov Institute, 123182 Moscow, Russia

N. B. Kol'chugina and O. D. Chistyakov

A. A. Baĭkov Institute of Metallurgy

(Submitted 18 November 1998)
Zh. Éksp. Teor. Fiz. **115**, 2197–2206 (June 1999)

We have studied the evolution of the inelastic neutron magnetic scattering spectra of a compound with cubic symmetry, $CeInCu_2$, in the temperature range 10–130 K, and also their transformation with variation of the Kondo temperature T_K due to substitution of cerium ions in the system $Ce_{1-x}(La,Y)_xInCu_2$ at $T=10$ K. It turns out that the energy of the transition between the ground state and excited state of the $4f$ electrons (Δ_{CF}) in the crystal electric field in $CeInCu_2$ increases with growth of the population of the ground state as the temperature is reduced, with a slight change in its intensity. Such behavior is inconsistent with the notion of classical one-ion effects of the crystal electric field. We have found that the scale of the observed variations in the excitation spectra of the $4f$ electrons depends on the Kondo temperature T_K and is insensitive to disorder in the rare-earth sublattice. Thus, despite the fact that $T_K \ll \Delta_{CF}$, hybridization with states in the conduction band has a substantial effect on all parameters of the excitation spectrum of the ground multiplet of the $4f$ electrons at low temperatures. © *1999 American Institute of Physics.* [S1063-7761(99)02206-4]

1. INTRODUCTION

One of the main interactions responsible for the physical properties of rare-earth (RE) based compounds at low temperatures is the interaction of $4f$ electrons with the potential of the crystal electric field, which removes the degeneracy of the ground multiplet of the rare-earth ions. In heavy-fermion compounds, in addition to this interaction, hybridization of the $4f$-electron states with electron states in the conduction band, which is usually described in terms of the Kondo interaction, also plays an important role. The characteristic energy (Δ_{CF}) of the interaction of the $4f$ electrons with the potential of the crystal electric field, as a rule, substantially exceeds the energy scale of the Kondo interaction (T_K). This is probably the reason why only the removal of degeneracy of the J multiplet by the crystal field is taken into account when determining the initial wave function of the ground state in various models of heavy-fermion systems.

The influence of the Kondo interaction on the excitation spectrum of the $4f$ electrons reduces to a renormalization of the ground state and to appreciable broadening of the transitions between levels of the crystal electric field (see Refs. 1 and 2 and the references therein). At the same time, the interaction with the crystal electric field determines for the ground state and excited states of the $4f$ electrons not only the magnitude of the splitting and multiplicity of the degeneracy, but also the symmetry of the wavefunctions of the states that hybridize with the states in the conduction band, which can be manifested in certain temperature effects.

The first experimental indications of a possible mechanism by which hybridization influences the formation of excitations of $4f$ electrons in the classical heavy-fermion compound with hexagonal symmetry $CeAl_3$ appeared in Ref. 3. In particular, in this study a strong modification of the neutron inelastic magnetic scattering spectra (changes in the position and intensity of the inelastic transitions associated with the energy and wave functions of the f-electron states) was observed when the temperature was reduced. In a recent study[4] on $Ce_{0.5}La_{0.5}Ni$ (a compound with lower, orthorhombic crystal symmetry), in which the Ce ions are in the heavy-fermion state, qualitatively similar effects were also observed.

In the systems $CeAl_3$ and $Ce_{0.5}La_{0.5}Ni$ the ground state multiplet of the cerium ions is sextuply degenerate and splits in the crystal electric field into three doublets, and, depending on the symmetry of the ground state, one or two transitions from the ground state can be observed in the neutron inelastic magnetic scattering spectra. Since the widths of the transitions are comparable to the excitation energy, separating out and analyzing the temperature variations presents real difficulties. For a higher symmetry of the environment, namely cubic, the ground-state $4f$ multiplet of the cerium ions splits only into two groups of states: a Γ_7 doublet and a Γ_8 quartet, i.e., in the neutron inelastic magnetic scattering spectra only one transition between levels of the crystal electric field is observed. Such a spectrum is simpler and more convenient for experimental and theoretical study of the renormalization of splitting of the crystal electric field and of the wave functions in the presence of hybridization. On the other hand, a comparison of the excitation spectra of compounds with a variety of crystal symmetry would make it possible to establish general trends and identify specific details in the formation of the ground state of heavy-fermion compounds. Therefore it would be useful to investigate experimentally the temperature evolution of the excitation spectra of $4f$ electrons in a heavy-fermion compound with cubic symmetry at temperatures ranging from less than T_K to values above Δ_{CF}, and to compare with analogous results for compounds with lower crystal symmetry.[3,4]

2. SAMPLES AND MEASUREMENT TECHNIQUE

A suitable object of study for examining the role of hybridization in the formation of the excitation spectrum of $4f$ electrons is the cerium-based heavy-fermion compound $\mathrm{CeInCu_2}$. This compound has cubic symmetry of the local environment of the rare-earth ion and belongs to the structural type $L2_1$ (Ref. 5). The coefficient of the electronic term in the specific heat, found by extrapolating C/T from the range 10–20 K to $T=0$ for this compound, is $\gamma=247\,\mathrm{mJ/(mole\cdot K^2)}$ (Ref. 6). It increases as the temperature is reduced, and reaches $\gamma=1200\,\mathrm{mJ/(mole\cdot K^2)}$ at $T<1.4\,\mathrm{K}$ (Refs. 6–8). Such a large value of γ at low temperatures is probably the result of the emergence of magnetic order at $T<2$ K (Refs. 6 and 9). The magnetic susceptibility deviates from the Curie–Weiss law at $T<30\,\mathrm{K}$, and extrapolation of its value χ as $T\rightarrow 0\,\mathrm{K}$ yields $\chi(0)=37\times 10^{-3}$ emu/mole (Ref. 7), which allows us to consider $\mathrm{CeInCu_2}$ to be a heavy-fermion compound. On the basis of low-temperature magnetic and thermodynamic data,[6–8,10] the Kondo temperature T_K for $\mathrm{CeInCu_2}$ was estimated to be approximately 4 K. On the other hand, the emergence of magnetic order at low temperatures[6] can lead to an increase in γ and thereby yield an underestimate of T_K (starting at $\gamma=247\,\mathrm{mJ/(mole\cdot K^2)}$ we have $T_K\approx 20\,\mathrm{K}$).

A strong dependence of the position of the maximum $T_{\max}$ (i.e., the maximum associated with T_K) on the applied hydrostatic pressure ($\Delta T_{\max}/\Delta P\approx 24\,\mathrm{K/GPa}$ to 1.55 GPa) was detected in measurements of the temperature dependence of the magnetic component of the resistance of $\mathrm{CeInCu_2}$ (Ref. 11), which indicates that T_K is very sensitive to pressure. This peculiarity of the compound makes it possible to study the influence of the magnitude of the Kondo interaction on the formation and properties of the ground state.

In neutron measurements it is more convenient for the purpose of measuring T_K to use the ''chemical pressure'' instead of the external hydrostatic pressure, i.e., replacing cerium ions with the smaller Y ions (which corresponds to applying a positive pressure with a resultant increase in T_K) or with the larger La ions (a ''negative'' pressure with a corresponding decrease in T_K). It should be noted that during chemical substitution in the rare-earth sublattice of the system $\mathrm{CeInCu_2}$, no change takes place in the type of ions in the immediate environment of the cerium ions, which consists of eight copper ions. This suggests that to first order the ionic component of the crystal, which as a rule[12] is determined mainly by its immediate environment, remains unchanged upon Ce ion substitution. The electronic component of the potential of the crystal electric field, which is substantial for metals, can also be considered to be invariant upon covalent substitution, as has been confirmed by experimental studies.[12]

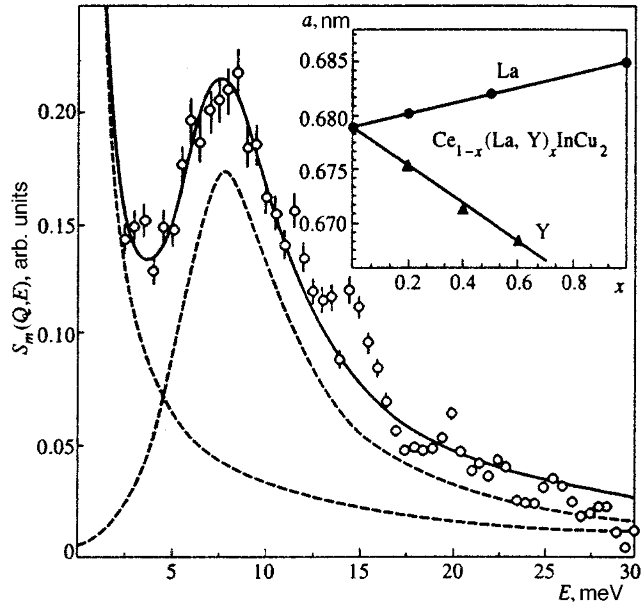

FIG. 1. Magnetic component of the spectral function $S_m(Q,E)$ for $\mathrm{CeInCu_2}$, obtained from inelastic neutron scattering spectra at $T=10\,\mathrm{K}$. Points — experiment (see text). Lines — fits to data using Lorentzian spectral functions: dashed lines — inelastic ($E_{in}=6.98\pm 0.09$ meV, $\Gamma_{in}=7\pm 1$ meV) and quasielastic ($\Gamma_{qe}=1.4$ meV) components, solid line — sum of the two. Inset plots the concentration dependence of the cubic lattice parameter for $\mathrm{Ce_{1-x}(La,Y)_xInCu_2}$. Error in the lattice parameter is less than the width of the symbol.

Neutron studies of the magnetic excitation spectrum of $\mathrm{CeInCu_2}$ were performed in Refs. 8 and 10. These studies revealed that at $T=5\,\mathrm{K}$ there is one inelastic peak in the spectrum with maximum around 8 meV, associated with the excitation from the ground state, and a quasielastic peak with a half-width at half-maximum (HWHM) of approximately 0.35 meV. In Refs. 8 and 10 it was concluded that when the temperature is raised, the inelastic peak broadens, and at $T>50\,\mathrm{K}$ it transforms into a quasielastic peak. The authors associated the observed evolution of the inelastic peak with the relaxation processes observed in Ref. 13. Such a conclusion is inconsistent with recent measurements on other heavy-fermion compounds, in particular $\mathrm{CeAl_3}$ (Ref. 3), and requires additional experimental verification.

Polycrystalline samples of $\mathrm{Ce_{1-x}(La,Y)_xInCu_2}$ ($x=0$ and 0.2) and the structural nonmagnetic analog $\mathrm{LaInCu_2}$ were prepared in an arc furnace in an argon atmosphere. X-ray analysis showed that all the samples correspond to the structural type of $\mathrm{CeInCu_2}$. Within the limits of experimental error ($\approx 3\%$) no other phases were detected. The lattice parameters for $\mathrm{CeInCu_2}$ ($a=0.67894\pm 0.00003\,\mathrm{nm}$) and $\mathrm{LaInCu_2}$ ($a=0.68512\pm 0.00004\,\mathrm{nm}$) are in good agreement with published data.

Measured values of the lattice parameter as a function of cerium ion concentration are shown in the inset to Fig. 1. Within the limits of the examined concentrations the lattice constant varies linearly (and significantly) as the Ce content is reduced. The lack of salient points in the dependence of the lattice parameter on the concentration suggests that no isostructural phase transition to a state with intermediate valence takes place (a transition of the type $\gamma \rightarrow \alpha$ in Ce), i.e., the cerium ions remain in the heavy-fermion state. For the neutron experiments we chose two samples with the same cerium ion concentration but opposite ''chemical'' pressure-effect: $Ce_{0.8}Y_{0.2}InCu_2$ and $Ce_{0.8}La_{0.2}InCu_2$. Comparison of the derived lattice constants upon substitution by Y atoms and application of hydrostatic pressure[14] allowed us to estimate that substitution of 20% of the cerium atoms is equivalent to application of a pressure of ~ 1 GPa. On the basis of the results of Ref. 14, we believe that the Kondo temperature T_K increases in proportion to the increase in the pressure. In this case the application of a pressure of 1 GPa should lead to roughly a doubling of T_K. Extrapolating this dependence to ''negative'' pressures, we obtain the value $\approx 0.7 T_K$ for the samples with lanthanum.

We measured the neutron inelastic magnetic scattering spectra of the indicated samples with a KDSOG-M time-of-flight spectrometer (reactor IBR-2, LNF, OIYaI) with a fixed final neutron energy $E_f = 4.9$ meV (we used a pyrographite analyzer) in the temperature range 10–130 K. The energy resolution at the elastic peak was 0.5 meV. The measurements at $T = 100$ K on $CeInCu_2$ and $Ce_{0.8}Y_{0.2}InCu_2$ samples had a resolution of around 1.5 meV. Powdered samples were prepared in the form of a set of wafers around 1 mm thick in thin aluminum foil, which made it possible to achieve a neutron transparency ≈ 80 for $E = 25$ meV. The range of neutron scattering angles was 30°–90°.

To obtain the scattering function $S(Q,E)$ we introduced a background correction into each of the derived neutron inelastic scattering spectra, normalized to the spectrum of incident neutrons, and summed over scattering angles in the range $30° - 90°$. A possible nonuniformity in the thicknesses of the samples precludes taking accurate account of the correction for absorption (self-screening) of the sample, which decreases with increasing transferred energy. However, a quantitative estimate shows that this self-screening, taking into account the actual, relatively high transmittance of the samples, is compensated by a correction for the magnetic form factor $F^2(Q)$ [variation of the scattering intensity with neutron momentum (energy) transfer at a fixed scattering angle], which has an inverse dependence on the energy transfer. Therefore we ignored these corrections.

To estimate the background components of the neutron inelastic scattering spectra for $CeInCu_2$ we used the neutron inelastic scattering spectrum of the structural nonmagnetic analog $LaInCu_2$. The background components were determined from the $LaInCu_2$ spectra measured under the same experimental conditions. An example of the neutron inelastic magnetic scattering spectrum $S_m(Q,E)$ of $CeInCu_2$ at $T = 10$ K derived in this way is shown in Fig. 1. All spectra taken on other samples had a qualitatively similar form, differing only in the parameters of the inelastic peak. Some differences in the energy positions of the maxima and relative intensities of the peaks in the generalized density of phonon states for the lanthanum-based samples and the investigated samples led to the onset of small oscillations in the $S_m(Q,E)$ spectra for the $CeInCu_2$-based compounds (Fig. 1) which were not associated with magnetic scattering.

To determine the parameters characterizing the excitation spectrum of the $4f$ electrons, we fit the dependence $S_m(Q,E)$ with the help of two Lorentzian spectral functions[15] with energy position $E_{qe} = 0$ meV (quasielastic scattering) and with $E_{\mathrm{in}} \neq 0$ meV (inelastic scattering):

$$S_m(Q,E) \propto \frac{F^2(Q)E}{1-\exp(-E/kT)} \left[\frac{\Gamma_{qe}/2}{(\Gamma_{qe}/2)^2 + E^2} + \frac{\Gamma_{\mathrm{in}}/2}{(\Gamma_{\mathrm{in}}/2)^2 + (E - E_{\mathrm{in}})^2} \right], \tag{1}$$

where Γ_{qe} and Γ_{in} are the widths of the quasielastic peak and the inelastic peak, and $F(Q)$ is the magnetic form factor.

Since the energy resolution in the given measurements was insufficient to determine the parameters of the quasielastic component, in the fitting of the spectra for $CeInCu_2$ we used the temperature dependence of Γ_{qe} derived in Ref. 10. In fitting the $S_m(Q,E)$ spectra of the samples with partial substitution of cerium we used data obtained for the $CeInCu_2$ spectra as the initial parameters.

3. MEASUREMENT RESULTS AND DISCUSSION

The derived energy positions of the maxima of the Lorentz peaks and their total intensities are plotted in Fig. 2. For $CeInCu_2$ the maximum of $S(Q,E)$ corresponds to $E_{\mathrm{in}} = 6.98 \pm 0.09$ meV at $T = 10$ K. The peak is preserved as the temperature is increased, but its energy in the magnetic spectrum drops to $E_{\mathrm{in}} = 6.4 \pm 0.1$ meV at $T = 130$ K. In the compound with 20% substitution of yttrium for cerium the inelastic peak has a large value ($E_{\mathrm{in}} = 7.63 \pm 0.08$ meV), which decreases noticeably($E_{\mathrm{in}} = 6.7 \pm 0.2$ meV) as the temperature is increased to $T = 100$ K. Substituting lanthanum for cerium leads to a negligible decrease in the energy of the maximum of the magnetic peak ($E_{\mathrm{in}} = 6.76 \pm 0.08$ meV) [Fig. 2a].

The magnetic excitation spectrum of $CeInCu_2$ at $T = 10$ K obtained in the present study is in good agreement with the results of Ref. 10. Differences at high temperatures are probably due to ignoring background scattering in Ref. 10. According to measurements of the magnetic susceptibility and specific heat of $CeInCu_2$ (Ref. 7), the ground state is a doublet, and the excited state is a quartet. It is noteworthy that as the temperature is reduced from 130 K the energy of the inelastic peak experiences a considerable increase (of approximately 0.6 meV). Here we are talking about an increase in the energy of the maximum of the spectral function E_{in}, not the experimentally observed peak, which is shifted considerably more by virtue of the influence of the thermal factor [see formula (1)]. The greatest change takes place in the temperature range 30–10 K (Fig. 2). In this temperature range the lattice constant varies by less than 6×10^{-5} nm (Ref. 11), which can lead only to a negligible shift in the levels, much less than the experimental error in the position of the peak obtained in the present work. This prevents us from relating the observed effects to changes in the potential of the crystal electric field due to changes in the distance of the rare-earth ion to its nearest neighbors.

Note that in the temperature range investigated here, the total intensity of inelastic magnetic scattering was also found

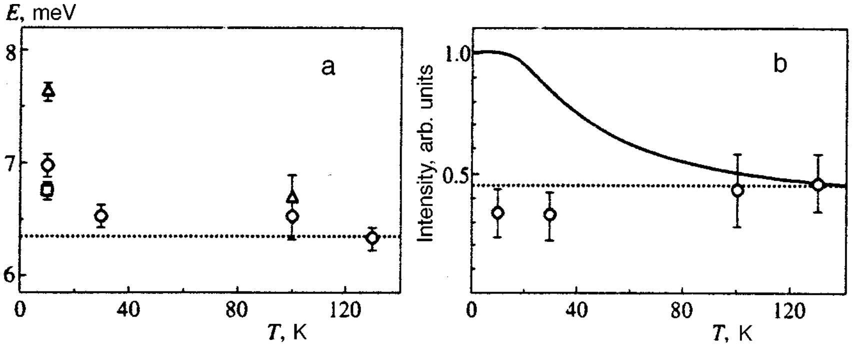

FIG. 2. a — Temperature dependence of the energy position E of the inelastic peak in the inelastic neutron magnetic scattering spectrum: ○ — $CeInCu_2$, Δ — $Ce_{0.8}Y_{0.2}InCu_2$; □ — $Ce_{0.8}La_{0.2}InCu_2$. The dotted line corresponds to $E=6.45$ meV. b — Temperature dependence of the total intensity of inelastic neutron magnetic scattering in the $CeInCu_2$ spectra, normalized to the intensity at $T=130$ K. Solid line — calculation of the intensity of the inelastic transition between levels of the $4f$ multiplet of cerium in the cubic crystal electric field for asplitting diagram with Γ_7 doublet as the ground state and splitting energy 7 meV. Dotted line — relative intensity 0.45. Experimental and calculated data ''referenced'' to the experimental data at $T=130$ K.

to have an anomalous temperature dependence. Figure 2b plots experimental data on the temperature dependence of the total intensity of inelastic scattering together with results of a calculation of the intensity of the transition between levels of the crystal electric field $\Gamma_7 \rightarrow \Gamma_8$ with energy $E_{\text{in}} = 6.98$ meV, ''referenced'' to $T=130$ K. It can be seen that as the temperature is reduced, not only does the observed intensity not grow, it even decreases somewhat. Thus, the inelastic component of the magnetic excitation spectrum of the given compound cannot be considered to correspond to the ordinary transition between levels of the crystal electric field, for which reducing the temperature leads to an increase in the population of the ground state, and consequently to an increase in the intensity of the inelastic peak with almost no change in its energy position. The width of the inelastic peak in the $S_m(Q,E)$ spectra is $\Gamma \approx 7$ meV at low temperatures, which is much greater than the energy resolution of the KDSOG-M spectrometer. Comparison of the parameters of the $S_m(Q,E)$ spectra at $T=10$ K obtained for all the samples with the parameters of the spectrum of $CeInCu_2$ at $T=130$ K shows that the observed changes in the position of the maximum in the investigated temperature range is approximately proportional to the change in T_K: an increase in T_K ($Ce_{0.8}Y_{0.2}InCu_2$) leads roughly to double the shift of the maximum toward higher energies ($\Delta E = 1.2$ meV), while a decrease in T_K ($Ce_{0.8}La_{0.2}InCu_2$) reduces the energy shift of the maximum to $\Delta E = 0.4$ meV.

The observed features in the $S_m(Q,E)$ spectra for $CeInCu_2$ are in qualitative agreement with the temperature variations observed in the heavy-fermion compound $CeAl_3$ (Ref. 3) and in $Ce_{0.5}La_{0.5}Ni$ (Ref. 4). A parameter of the system that varies appreciably at these temperatures is the population of the levels of the crystal electricfield, which for the ground state grows from ≈45% at $T=130$ K to ≈100 at $T=10$ K. Figure 3 plots the energies of the peaks and their total intensities normalized to the energy and total intensity at the maximum temperature of the measurements versus the population of the ground state. Figure 3b also shows results derived from a calculation based on one-ion notions. Alignment of the experimental and calculated data was accomplished at the maximum measurement temperatures. The figure presents results for $CeInCu_2$ obtained in the present work, and also for $CeAl_3$ (1.5–90 K, Ref. 3) and $Ce_{0.5}La_{0.5}Ni$(12–150 K, Ref. 4).Consideration of all the data taken together enables us to tentatively identify three temperature ranges for all these compounds. For temperatures at which the population of the ground state is less than about 0.7, the position of the peaks within the limits of experimental accuracy does not change. As the population of the ground state increases from 0.7 to 1, the inelastic peak begins to move toward higher energies while the total intensity decreases only insignificantly. In the third range the population of the ground state is almost constant and equal to unity, which on the temperature scale corresponds to $T \leqslant T_K$. In

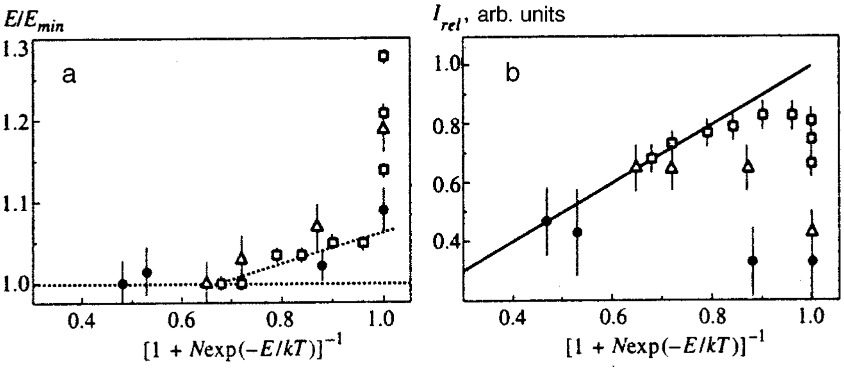

FIG. 3. Dependence of the energy position of the inelastic peak (Lorentzian, see text) of the transition between the ground state and the first excited level of the crystal electric field (a) and its intensity $S_m(Q,E)$ (b), normalized to the corresponding high-temperature values, on the population of the ground state of the $4f$ multiplet for $CeInCu_2$ (●, present work), $Ce_{0.5}La_{0.5}Ni$ (Δ, Ref. 4), $CeAl_3$ (□, Ref. 3); N — coefficient associated with the degree of degeneracy of the excited state ($N=1$ for $Ce_{0.5}La_{0.5}Ni$ and $CeAl_3$, $N=2$ for $CeInCu_2$). Solid line in b — calculation of the relative intensity of the transition between the ground state and the excited level in of the crystal electric field in the one-ion picture. The dashed lines a are fitted by eye (see text).

this range the energy of the peak is significantly greater, and its intensity is somewhat less. The derived temperature dependences tell us that as the temperature is reduced, changes in the excitation spectra of the $4f$ electrons for heavy-fermion systems start up at that point at which significant growth begins in the population of the ground state, and the greatest changes occur at temperatures on the scale of T_K and for an unchanging population of the ground state. The observed effects cannot be a consequence of ordinary relaxation processes since according to the results of Ref. 13 the effect of relaxation on the characteristics of the spectrum increases when the temperature is raised, and not the other way around, as was found to be the case in the present work.

At the same time it should be noted that the magnitudes of the observed effects differ substantially from one compound to another. It is clear from looking at the temperature dependences of the total intensity of the inelastic peak [Fig. 3b] that the largest deviation of the experimental intensity from the result of the one-ion calculation at lower temperatures is observed for $CeInCu_2$. This possibly has to do with the different values of the Kondo temperature in these compounds: T_K $(CeAl_3)\approx 5$ K, T_K $(Ce_{0.5}La_{0.5}Ni)\approx 15$ K (Ref. 4), and T_K $(CeInCu_2)\approx 20$ K (Ref. 6). As follows from the data for $Ce_{1-x}(La,Y)_xInCu_2$ at $T=10$ K (Fig. 2), the energies of the maxima of the excitation spectra of the $4f$ electrons increase with increasing T_K. But a comparison of the positions of the maxima of the excitation spectra for compounds with different structures and Kondo temperatures shows that the Kondo temperature is not the only thing that determines the scale of the observed effects. The form of the ground-state wave function probably also has a substantial influence on the formation of the excitation spectrum. Thus, the greatest variation in the energy of the maximum of $S(Q,E)$ at $T\approx 0.5T_K$ is observed for systems with crystal symmetry of the local environment of the rare-earth ions lower than cubic ($E/E_{\min}=1.19$ for $Ce_{0.5}La_{0.5}Ni$, $E/E_{\min}=1.28$ for $CeAl_3$), and the least variation—for cubic symmetry ($E/E_{\min}=1.09$ for $CeInCu_2$) (Fig. 3). For systems with high crystal symmetry ($CeAl_3$ and $CeInCu_2$) it is not possible to explain the observed variation of the intensity simply on the basis of ''classical'' one-ion notions of crystal electric field effects since the wave functions of the $4f$-electron levels do not depend on the parameters of the crystal electric field. Thus, the detected effect is evidence of a change in the symmetry properties of the f-electron states in a certain temperature range.

Recently, the authors of Ref. 16 suggested an explanation for the observed changes in the neutron inelastic magnetic scattering spectra based on the temperature dependence of the component of the total Hamiltonian for the $4f$ electrons due to the anisotropic exchange interaction in Kondo systems. The proposed model considers the influence of the population of the levels of the crystal electric field on the constants of the exchange interaction. But at temperatures comparable to T_K, the marked increase in the temperature dependence is related to Kondo-scattering induced renormalization of the exchange integrals. In compounds with high crystal symmetry, in which the wave functions of states in the crystal electric field do not depend on the values of its parameters, two possible mechanisms for reducing the symmetry of the Hamiltonian have been considered to explain the emergence of temperature dependence in the intensity of transitions between states: one is based on reducing the symmetry of the exchange interaction due to defects in the immediate environment of the rare-earth ion, and the other is related to dynamic distortion of the environment of the rare-earth ion and, accordingly, to a change in the symmetry of the exchange interaction due to magnetoelastic interaction. The temperature dependence of each of these mechanisms comes about just as it does for the energy level of the crystal electric field, i.e., it is substantially enhanced at $T\sim T_K$ due to the Kondo effect. The emergence of ranges with different temperature dependence of the energies of transitions between levels of the crystal electric field [Fig. 3a] is in good agreement with predictions based on this model.[16] The changes we have observed in the energy of the peak in the $S(Q,E)$ spectrum in $CeInCu_2$ based compounds with substitution of 20% of their cerium ions but with different degrees of hybridization obtained in this research (Fig. 2) indicate that the introduction of defects into the rare-earth sublattice most probably does not have a substantial effect on the scale of the effect, i.e., the main role is played by the magnetoelastic interaction. Thus, the nature of the temperature effects in pure $CeInCu_2$ is the same as in $CeAl_3$ and $Ce_{0.5}La_{0.5}Ni$, and the observed renormalization of the spectrum reflects the formation of a new ground state of the f shell as a result of the interaction of the effects of the crystal electricfield and the Kondo effect.

4. CONCLUSION

Our results lead us to believe that in a heavy-fermion system with cubic symmetry of the crystal lattice, the excitation spectrum of the $4f$ electrons definitely evolves as the temperature is reduced, with the population of the ground state tending toward 100%. In particular, the energy increases and the total intensity of the peaks of the neutron inelastic magnetic scattering peaks decreases. No influence of disorder in the rare-earth sublattice on the transformation of the neutron inelastic magnetic scattering was detected, but the ongoing transformation is not a consequence of relaxation processes. The scale of changes in the spectrum is probably related mainly to the degree of hybridization (i.e., with T_K). It is important to stress that significant and qualitatively similar deviations from the temperature dependences for classical one-ion effects of the crystal electric field were observed in some heavy-fermion compounds with differing crystal symmetry.

In conclusion, we wish to thank A. S. Mishchenko for fruitful and stimulating discussions, A. Yu. Muzychok and Ch. Gantulga for assistance with the neutron measurements. This work was carried out with the support of the Russian Fund for Fundamental Research (Project No. 98-02-16229) and the State Science and Engineering Program ''Current Topics in Condensed-Matter Physics.''

*)E-mail: paval@sftt.kiae.su, lvn@isssph.kiae.ru

---

[1] E. Holland-Moritz and G. H. Lander, in *Handbook of the Physics and Chemistry of Rare Earths*, Vol. 19 (1994), Ch. 130.
[2] M. Loewenhaupt and K. H. Fisher, in *Handbook of the Physics and Chemistry of Rare Earths*, Vol. 16 (1993), Ch. 105.
[3] P. A. Alekseev, W. Bührer, V. N. Lazukov, E. V. Nefeodova, I. P. Sadikov, O. D. Chistyakov, and M. Zolliker, Physica B **217**, 241 (1996).
[4] V. N. Lazukov, P. A. Alekseev, E. S. Klement'ev, *et al.*, Zh. Éksp. Teor. Fiz. **113**, 1731 (1998) [JETP **86**, 943 (1998)].
[5] I. Felner, Solid State Commun. **56**, 315 (1985).
[6] S. Takagi, T. Kimura, N. Sato *et al.*, J. Phys. Soc. Jpn. **57**, 1562 (1988).
[7] Y. Onuki, T. Yamazaki, A. Kobori *et al.*, J. Phys. Soc. Jpn. **56**, 4251 (1987).
[8] R. Lahiouel, J. Pierre, E. Siaud *et al.*, Z. Phys. B **67**, 185 (1987).
[9] H. Kadowaki, S. Mitsuda, H. Yoshizawa *et al.*, J. Phys. Soc. Jpn. **58**, 4292 (1989).
[10] A. P. Murani, A. Najib, and J. Pierre, Ann. ILL Exp. Report 4-03-478 (1988), p. 74.
[11] A. Najib, J. Beille, R. Lahiouel *et al.*, J. Phys. F.: Met. Phys. **17**, 2395 (1987).
[12] P. A. Alekseev, V. N. Lazukov, A. Yu. Rumyantsev, and I. P. Sadikov, J. Magn. Magn. Mater. **75**, 323 (1988).
[13] K. W. Becker, P. Fulde, and J. Keller, Z. Phys. B **28**, 9 (1977).
[14] T. Kagayama, K. Suenaga, G. Oomi *et al.*, J. Magn. Magn. Mater. **90–91**, 451 (1990).
[15] E. Holland-Moritz, D. Wohlleben, and M. Loewenhaupt, Phys. Rev. B **25**, 7482 (1982).
[16] A. S. Mishchenko, JETP Lett. **66**, 487 (1997).

Translated by Paul F. Schippnick

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

''Symmetrical'' phase and collective excitations in the proton system of ice

I. A. Ryzhkin*)

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 18 November 1998)
Zh. Éksp. Teor. Fiz. **115**, 2207–2213 (June 1999)

A model of the ice proton system taking into account quantum-mechanical tunneling of protons along hydrogen bonds has been formulated and investigated. When the tunneling amplitude is small the quantum ground state of the proton system is degenerate, like the classical ground state. At higher tunneling amplitudes, however, a transition to a nondegenerate state with a symmetrical distribution of protons on hydrogen bonds (a symmetrical phase of ice) is possible. Collective excitations of protons in the nonsymmetrical phase have been considered, and an equation that determines their spectrum has been derived. © *1999 American Institute of Physics.* [S1063-7761(99)02306-9]

1. INTRODUCTION

Ice is one of the most common materials on the earth. Its unique mechanical , electrical, and thermodynamic properties are of great importance for various applications. From the fundamental viewpoint, the specific physical properties of ice can be interpreted in terms of the unusual structure of its crystal lattice, which contains two sublattices with radically different properties. One of them is formed by oxygen ions and is a regular lattice similar to crystal lattices of ordinary solid materials. Figure 1 shows the unit cell of the most common hexagonal modification of ice, which exists under natural conditions (oxygen ions are shown by open circles). Note that the oxygen sublattice has the wurtzite structure and is identical to the lattices of II–VI semiconducting compounds if oxygen ions are substituted for those of groups II and VI.

The second sublattice formed by protons, however, is disordered (it is shown by closed circles in Fig. 1). Specifically, protons can occupy one of two possible positions on each hydrogen bond, which connect oxygen ions, at distances of 0.1 nm from the oxygens (the whole length of the hydrogen bond is 0.27 nm).[1] Disorder implies that N protons are distributed randomly of among $2N$ positions in accordance with the two ice rules (also called Bernal–Fowler rules): there are two protons near each oxygen ion, and there is one proton on each bond. It is clear that these ice rules rule out proton conductivity, because any motion of a proton would break the ice rules and cause an increase in the energy. At finite temperatures the ice rules can be broken, giving rise to configurations with one or three protons near the oxygen ion (H_3O^+ and OH^- are ionic defects) and with two or zero protons on hydrogen bonds (D and L are bond defects). As a result of the motion of protons along bonds or from one bond to another, these defects can move through the oxygen lattice without a further increase in the proton system energy. This configuration of the proton lattice ground state, excitations, and charge transfer in ice find their natural interpretation in terms of the classical microscopic model, which takes into account the Coulomb interaction between protons and specific properties of the oxygen lattice.[2]

In the classical model, the motion of protons across potential barriers is treated as a classical thermally activated process. However, the proton motion along hydrogen bonds is assumed to be quantum-mechanical tunneling. The parameter measuring the tunneling efficiency is the splitting of the proton level in the symmetrical potential with two minima (the corresponding coordinate is measured along the hydrogen bond). In ice under natural conditions the splitting is nearly 3×10^{-3} eV, which is much smaller than the characteristic defect energy ≈ 1 eV (see Ref. 2). But high pressure applied to ice can reduce both the length of the hydrogen bond and the separation between potential energy minima.[3,4] The probability of tunneling under these conditions increases while the energy of the ion defects decreases. Thus, a situation can arise in which the quantum-mechanical tunneling becomes important. The aim of the present research is to study proton tunneling effects under such conditions. We will determine the ground state of a strongly correlated proton system, investigate its variation with the tunneling amplitude, and show that excitations in the nonsymmetrical state are the collective tunneling modes first introduced to describe collective motion in ferroelectrics.[5] Section 2 briefly describes the model, the basic approximations, and the transition to a symmetrical phase. A detailed description of the model can be found elsewhere.[2] Section 3 treats collective excitations of protons in the presence of tunneling along hydrogen bonds and derives an equation for the excitation spectrum.

2. MODEL AND PROPERTIES OF THE GROUND STATE

In describing protons as quantum-mechanical particles moving in the potential of the oxygen lattice and interacting with each other, it is natural to apply the Hubbard model.[6] The derivation of the basic equations of the Hubbard model version used in this study is described in detail elsewhere;[2,7]

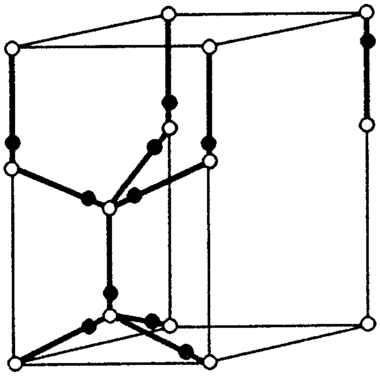

FIG. 1. Unit cell of hexagonal ice (the modification existing under natural conditions). Oxygen ions are shown by open circles, protons by closed circles. There is one proton on each hydrogen bond (bold lines connecting oxygen ions) and two protons near each oxygen ion.

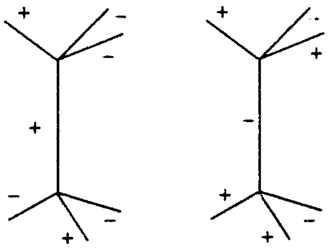

FIG. 2. Fragment of ice lattice illustrating division of bonds into two groups (positive and negative), in accordance with the ice rules. The effective field acting on the central bond pseudospin is $\mp\cos\theta$ for all states of neighboring bonds that satisfy the ice rules.

here we give only a brief description. Of especial importance, note that the localization radius of proton wave functions is much smaller than that of wave functions in electronic models, so the effect of intrasite correlations between protons is unusually strong. For this reason we first eliminate all states with two (or zero) protons at one potential minimum. Secondly, we treat a hydrogen bond as a single lattice site and two proton positions at two different minima of the potential energy as states characterized by the pseudospin variable. Thereby we describe the system in terms of the conventional Hubbard model for particles with spin 1/2. Remember that this parameter is the pseudospin which describes the proton position on a hydrogen bond, whereas the real proton spin is insignificant and will be ignored. Thirdly, in view of available experimental data, it is reasonable to keep in the model only proton tunneling along hydrogen bonds and neglect tunneling between different bonds. The presence of tunneling along the bond leads to a ''fictional'' magnetic field acting on the pseudospin variable. Fourthly, and finally, we neglect all configurations with zero or two protons on one bond (Bjerrum defects), since the mainapplication of the model is to ice under high pressure (where the distance between proton positions is smaller and the Coulomb interaction between protons at two minima on one bond is stronger). The projection of the whole Hamiltonian on the corresponding subspace of quantum states can be described in terms of the Pauli matrices:

$$H=-\Omega\sum_{i}\sigma_{xi}+\frac{J}{2}\sum_{ij}\sigma_{zi}\sigma_{zj}, \tag{1}$$

where the sum is performed over all bonds and pairs of the nearest neighbors, Ω is the matrix element of proton tunneling along a bond, and σ_x and σ_z are Pauli matrices. The spin-up and spin-down states correspond to different proton positions on a bond, and J is a positive constant of Coulomb interaction between pseudospins on the nearest hydrogen bonds. The sign of this constant is of great significance and means that this is an antiferromagnetic model. The point is that the pseudospin lattice (i.e., the lattice formed by centers of hydrogen bonds) is frustrated and contains cycles of nearest neighbors with odd numbers of sites. For this lattice one cannot satisfy all rules of antiferromagnetic ordering, which makes its ground state degenerate. With good accuracy the degeneracy of this lattice is given by the expression $(3/2)^N$ (Refs. 1 and 2). Neither can one transform the antiferromagnetic model to a ferromagnetic one by a gauge transformation (by changing signs of some spins and interaction constants), since a frustrated lattice cannot be divided into two mutually penetrating sublattices 1 and 2 such that all the nearest neighbors of sites of sublattice 1 (2) belong to sublattice 2 (1).

Although the Hamiltonian has a simple form, it cannot be diagonalized analytically. For this reason, we restrict our analysis to a version of the mean-field approximation. Before discussing this approximation, recall that in the approximation of the mean (or self-consistent) field, spin products are replaced by the following expressions:

$$\sigma_{zi}\sigma_{zj}=\sigma_{zi}\langle\sigma_{zj}\rangle+\langle\sigma_{zi}\rangle\sigma_{zj}-\langle\sigma_{zi}\rangle\langle\sigma_{zj}\rangle, \tag{2}$$

where the mean values $\langle\sigma_{zi}\rangle$ are determined in a self-consistent manner. In the ferromagnetic model the mean values of spins are constant, independent of the bond number i, and can be considered as an order parameter.^{5} For an ordinary nonfrustrated antiferromagnet the mean values of spins can take two values: $\pm\sigma$, where σ is the order parameter. In the case of ice, however, we deal with a frustrated antiferromagnetic model, whose ground state is degenerate and has a more complicated structure. To describe this state, we assume that the mean values of the spin variables, $\langle\sigma_{zi}\rangle$, depend on the index i in the following manner. Let us divide all hydrogen bonds into two equal groups (positive and negative ones) with the only condition that two positive and two negative bonds have to be attached at each oxygen site. It is clear that this separation is in fact equivalent to the first ice rule, whereas the second ice rule is incorporated in the model. The mean values of spins for the two groups of bonds are determined by the expressions $\langle\sigma_{zi}\rangle=\pm\cos(2\theta)$, where θ is defined from the self-consistency condition (or from energy minimum at $T=0$, or from the free energy minimum at $T\neq0$). An important point is that, although the bonds are divided randomly into two groups, we can consider the mean spins $\langle\sigma_{zi}\rangle$ taking only two values: $\pm\cos(2\theta)$. This follows from the analysis of effective fields, which can take in this case only two values with opposite signs for arbitrary bonds (Fig. 2). In fact, a stronger disorder cannot take place because of the short-range nature of the interaction (only nearest neighbors interact) and the ice rule (two positive and two negative bonds are attached at each oxygen lattice site). It is clear that this version of the mean-field approximation is identical to the variational approach with a wave function

$$\Psi=\prod_{i=1}^{N/2}\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}\prod_{j=1}^{N/2}\begin{pmatrix}\sin(\theta)\\ \cos(\theta)\end{pmatrix}, \tag{3}$$

where the products are performed over bond groups described above.

Using Eqs. (1) and (3) and definitions of the Pauli matrices, we easily obtain the following expression for the energy per bond:

$$E/N=-\Omega\sin(2\theta)-J\cos^2(2\theta). \tag{4}$$

This equation applies to both positive and negative bonds in any environment (Fig. 2). An elementary analysis of Eq. (4) shows that for $\Omega<2J$ the energy has a minimum at

$$\theta=\frac{1}{2}\arcsin\frac{\Omega}{2J}, \tag{5}$$

where the extremum at $\theta=\pi/4$ is a maximum. But at $\Omega=2J$ this second extremum becomes a minimum, and for all $\Omega>2J$ this is the only minimum on the energy curve. Since the wave functions of all hydrogen bonds are symmetrical with respect to the pseudospin variable or the proton distribution on the bond, the phase transition at $\Omega=2J$ can be associated with the conversion of the ice to the symmetrical phase. The phase characterized by θ given by Eq. (5) is the conventional disordered phase of ice, but of course with quantum corrections for $\Omega\neq0$. The degeneracy of the nonsymmetrical ground state is determined by the number of divisions of the hydrogen bonds into two groups in the manner describe above. Obviously, it equals the number of proton configurations that satisfy the ice rule, i.e., the number of the Bernal–Fowler configurations.

Note that, since the variational procedure discussed above is approximate, the resulting states are only approximately orthogonal. Let us estimate the largest overlap between two wave functions of the nearest ground states. It clearly follows from the shapes of wave functions (3) and the solution procedure that the nearest states are those with the smallest difference between the sets of positive and negative bonds. Given the correspondence between the division of bonds into two groups and the Bernal–Fowler classical configurations, it is obvious that the difference between the two nearest states is in the configuration of six hydrogen bonds. Figure 3 illustrates a transition between two such states: an H_3O^+ ion defect is generated, it moves along the shortest cycle and annihilates with an OH^- defect on the last step. Using Eqs. (3) and (5), one can easily show that the overlap between the nearest states is $(\Omega/2J)^6$, and the largest matrix element of the Hamiltonian is of order $\Omega(\Omega/2J)^6$. These quantities characterize the accuracy of the approximation treating barriers between the ground states as infinite and considering elementary excitations in one separate degenerate state.

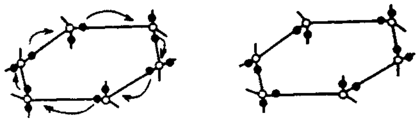

FIG. 3. Fragment of ice lattice containing the shortest cycle of six hydrogen bonds. Protons are shown by closed circles. The arrows depict six sequential proton jumps (creation of a pair of ionic defects on the first jump, subsequent motion of the positive defect along the closed circuit, and its annihilation on the sixth jump). As a result of these six sequential jumps, the left-hand proton configuration transforms to the right-hand one, which is the one closest to it.

3. TUNNELING MODES

The quantum tunneling of protons results in the existence of a new kind of excitations in the proton system. If hydrogen bonds were absolutely uncorrelated, these excitations would be tunneling modes similar to excitations in two-level systems in glasses. But the Coulomb interaction between protons leads to correlations between events of proton tunneling and a radically different character of excitations. A similar problem was studied by De Gennes[5] in the case of a ferromagnetic. Following De Gennes, let us term the new excitations collective tunneling modes.

To investigate the excitation spectrum, let us use the technique of the equation of motion, which includes the following steps: the part of Hamiltonian corresponding to the mean-field approximation is diagonalized; an equation of motion is derived by adding the fluctuational part; the equation of motion is linearized assuming that deviations of spins from their mean values are small. The linear equations of motion determine the spectrum of elementary excitations.[5,6] We begin carrying out this program with an identity transformation of Hamiltonian (1). Using the expressions for spins $\sigma_{zi}=\langle\sigma_{zi}\rangle+\delta\sigma_{zi}$ in Eq. (1), we obtain

$$H=-\Omega\sum_i\sigma_{xi}+J\sum_i h_i\sigma_{zi}-\frac{J}{2}\sum_{ij}\langle\sigma_{zi}\rangle\langle\sigma_{zj}\rangle+\frac{J}{2}\sum_{ij}\delta\sigma_{zi}\delta\sigma_{zj}, \tag{6}$$

where $h_i=\Sigma_{j\in i}\langle\sigma_{zj}\rangle$ is the effective mean field acting on spin *i* and generated by its nearest neighbors *j*. Next, as in the previous section, we divide all bonds into two groups — positive and negative. The mean values of the spins for these groups are $\pm\cos(2\theta)$, respectively. As a result, we have the Hamiltonian in the form

$$H=NJ\cos^2(2\theta)-\Omega\sum_i\sigma_{xi}\mp2J\times\cos(2\theta)\sum_i\sigma_{si}+\frac{J}{2}\sum_{ij}\delta\sigma_{zi}\delta\sigma_{zj}, \tag{7}$$

where the upper and lower signs correspond to the positive and negative bonds, respectively. The one-particle part of the Hamiltonian can be diagonalized through rotation in the pseudospin space, which is described by the change of variables:

$$\sigma_{xi}=\pm\sigma'_{xi}\cos(2\theta)+\sigma'_{zi}\sin(2\theta), \tag{8}$$

$$\sigma_{zi} = -\sigma'_{xi}\sin(2\theta) \pm \sigma'_{zi}\cos(2\theta), \tag{9}$$

where the parameter θ, as in the previous sections, is determined by the equation $\sin(2\theta)=\Omega/2J$. In new variables, the Hamiltonian takes the form

$$H = NJ\cos^2(2\theta) - [\Omega\sin(2\theta) + 2J\cos^2(2\theta)] \times \sum_i \sigma'_{zi} + \frac{J}{2}\sum_{ij} \delta\sigma'_{zi}\delta\sigma'_{zj}, \tag{10}$$

where $\delta\sigma_{zi} = -\sigma'_{xi}\sin(2\theta) \pm \sigma'_{zi}\cos(2\theta) \mp \cos(2\theta)$.

At first sight, it may seem that signs ± alternating randomly in the last term of Eq. (10) should lead to disordered equations for the operators σ'_{xi} and σ'_{yi}. This, however, is not so for the reasons given in the previous section (see the note about the choice of mean values of spins). A self-consistent derivation of the equation of motion for spin operators acting on both positive and negative bonds yields absolutely identical expressions. After linearization in small σ'_x and σ'_y variables, they take the form

$$\frac{d\sigma'_{xi}}{dt} = 4J\sigma'_{yi}, \tag{11}$$

$$\frac{d\sigma'_{yi}}{dt} = -4J\sigma'_{xi} - 2J\sin^2(2\theta)\sum_{j\in i} \sigma'_{xj}. \tag{12}$$

After eliminating the variables σ'_{yi} and taking $\sigma'_{xi} = a_i \exp(i\mathbf{kr} - i\varepsilon t)$, we obtain an equation for the amplitudes:

$$[\varepsilon^2 - (4J)^2]a_i = 8J\sin^2(2\theta)\sum_{j\in i} \exp(i\mathbf{kr}_{ij})a_j, \tag{13}$$

where $i=1,\ldots\nu$, and ν is the number of hydrogen bonds in one unit cell of the ice modification under investigation; $\nu=8$ for the hexagonal ice (Fig. 1).

4. CONCLUSION

The spectrum of collective tunneling modes is calculated by equating the determinant of the system of equations (13) to zero. Unfortunately, this equation cannot be solved analytically for all ice modifications, and numerical calculations are necessary. However, some essential properties of the spectrum can be determined without solving specific equations. Firstly, there is a gap in the excitation spectrum equal to approximately $4J$ and corresponding to the energy of formation of classical ionic defects. Secondly, the band width is proportional to Ω^2/J, but not Ω alone. Since $\Omega/J \ll 1$ holds, this means that the band is narrowed by the Coulomb interaction. Thirdly and finally, note that quantum analogs of classical ionic defects are more like spin waves than current carriers in conventional semiconductors. Like spin waves, the tunneling modes are ''almost'' bosons: the creation and annihilation operators, σ_i^+ and $\sigma_{i'}^-$, obey the commutation rules for different bonds i and i' and anticommutation rules for the same bond. An important point is that these statistical properties are only based on the commutation properties of their creation and annihilation operators, whereas distribution functions of ionic defects or their quantum analogues are undoubtedly of greater interest. As was shown previously,[8] the latter are more like fermion distribution functions.

The work was supported by the Russian Fund for Fundamental Research (Project No. 98-02-16642).

*)E-mail: ryzhkin@issp.ac.ru

---

[1] P. V. Hobbs, *Ice Physics*, Clarendon Press, Oxford (1974).

[2] I. A. Ryzhkin, Solid State Commun. **52**, 49 (1984).

[3] W. B. Holzapfel, J. Chem. Phys. **56**, 712 (1972).

[4] A. Polian and M. Grimsditch, Phys. Rev. Lett. **52**, 1312 (1984).

[5] P. G. De Gennes, Solid State Commun. **1**, 132 (1963).

[6] J. Hubbard, Proc. R. Soc. London, Ser. A **285**, 542 (1965).

[7] I. A. Ryzhkin, in *Proceedings of International Symposium on the Physics and Chemistry of Ice*, ed. by N. Maeno and T. Hondoh, Hokkaido University Press, Sapporo, Japan (1992), p. 141.

[8] V. F. Petrenko and I. A. Ryzhkin, J. Phys. Chem. **101**, 6285 (1997).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Nonuniform electron states near cleaved bismuth surface

A. M. Troyanovskiĭ

Institute of High-Pressure Physics, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

V. S. Édel'man*)

P. L. Kapitza Institute for Physical Problems, Russian Academy of Sciences, 117334 Moscow, Russia

(Submitted 14 December 1998)
Zh. Éksp. Teor. Fiz. **115**, 2214–2227 (June 1999)

Using scanning tunneling microscopy, we have detected nonuniform electron states on cleaved bismuth surfaces, which manifest themselves in the form of a random relief with a characteristic lateral size of 1–2 nm and a vertical size of a fraction of an angström. This relief is due to variations in current–voltage characteristics when the tip is moved over the sample surface. Features on the current–voltage characteristics associated with states near the surface have been observed in a voltage range of -0.4 to $+0.4$ V. The nonuniform states are tentatively associated with defects generated in the process of crystal cleavage.

1. INTRODUCTION

The cleaved surface of a brittle crystal has been a popular object of research using the scanning tunneling microscopy and scanning atomic-force microscopy. Most attention has been focused on cleaved surfaces of semiconductors, where one can observe atomically smooth surfaces with well-resolved atomic structures. In many cases features with lateral sizes of order several lattice constants and vertical displacements of a fraction of an ångström were observed. They are associated with impurities located at some depth below the upper surface layer.[1–3] The possibility that the cleaving process itself could generate local defects has not been discussed in this context. This issue can be raised, naturally, only when pure materials are studied. The experiments with atomically smooth surfaces described in this paper have demonstrated that their electronic parameters vary on the scale of several nanometers. This has led us to conclude that our concepts concerning the process of crystal cleavage should be revised.

Bismuth, studies of whose electronic properties have had considerable impact on the progress in metal physics, still attracts a lot of attention. But, whereas in the 1970–80s most effort was concentrated on studies of its bulk properties,[4] almost all current publications are dedicated to phenomena associated with the surface. The reason for this shift of attention is obvious: the concentration of current carriers in bismuth is low, of order 10^{-5} per atom, and their wavelength and Debye screening radius are much longer than the lattice constant and reach hundreds of ångströms, so the entire subsystem of conduction electrons should undergo a modification at such separations from the surface. In this connection, for example, quantum size effects and a transition to a semiconducting state in thin films have been mooted during recent decades.[5,6] In order to interpret experiments with thin epitaxial layers oriented in the basal trigonal plane, one has to assume the presence of a surface charge whose magnitude and sign depend on the substrate material and surface condition. This charge varies between -2.5×10^{12} (Ref. 5) and $+8\times10^{12}$ (Ref. 6) electron charges per square centimeter. If these numbers are scaled with respect to the volume in a layer of thickness equal to the lattice constant along the three-fold crystal axis, which is about 1.2 nm, we have $(2\text{–}6)\times10^{19}$ cm^{-3}, i.e., a quantity which is at least two orders of magnitude greater than the bulk density of both electrons and holes. In analyzing the reflection of current carriers from the surface and their transverse focusing in magnetic field one has to assume the presence of electron band bending of order 0.1 eV at the surface and the occurrence of an excess surface charge.[7]

Direct evidence of the presence of surface states or resonances lying at 0.4 eV below the Fermi level is provided by accurate measurements of the photoeffect (one of the latest studies in this field was published by Jezequel *et al.*[8]). The accuracy of these measurements, however, is limited: their resolution is 0.25 eV. Moreover, they yield parameters averaged over the entire surface.

Therefore, it seems interesting to study the tunneling spectra of bismuth. Tunneling measurements of layered structures were conducted long ago in studies of the bulk bismuth spectrum, when the impact of the surface had not been realized as clearly as at the present time. Note that disagreement between the results published by different authors is so great (these results were reviewed earlier[4]) that they are hardly worth considering. It seems highly probable that the decisive role in this case is played by interfaces in layered tunneling structures, whose properties strongly depend on technology. The technique of scanning tunneling spectroscopy, i.e., measuring local current–voltage characteristics using a scanning tunneling microscope, allows one to make specific statements about the state of the surface if samples are prepared in ultrahigh vacuum by cleaving crystals, or using a more expensive and complicated techniques of epitaxial film deposition in vacuum or ion etching of a crystal surface with subsequent annealing and concurrent monitoring of the surface composition using updated methods of analysis. The advantage of this technique is that the scanning tunneling spectroscopy allows one to measure the electron spectrum with a lateral resolution down to the interatomic separation. This is combined in a natural way with the scanning tunneling microscopy, which allows one to establish a correspondence between the spectroscopy data and the real surface structure, including certain linear or point defects.

Studies of cleaved bismuth surfaces using scanning tunneling microscopy were started in our earlier experiments,[9,10] when we investigated the structure of diatomic steps, thermal motion of their boundaries, and surface atomic corrugation at room temperature and above. Later studies were conducted at lower temperature, down to that of liquid helium. At these temperatures features such as a linear structure of terraces with straight, almost atomically smooth boundaries on surfaces cleaved at low temperatures[11] and twin layers of quantized widths with ideal boundaries[12,13] were discovered. Tunneling spectroscopy has indicated that one-dimensional electron states arise in the region of these twin layers. Below we describe our investigations of atomically smooth areas on cleaved bismuth surfaces, which have shown that electron states with typical energies of tens of meV with respect to the Fermi level are nonuniform in the surface plane over distances of several lattice constants.

2. EXPERIMENTAL TECHNIQUES

In our experiments, we used samples in the shape of long rods oriented along the [0001] axis with sizes of about $1\times 2\times 5\ \mathrm{mm}^3$. They were spark-cut from single crystals grown from melt by the technique described in Ref. 14 from a starting material with a purity of 99.99999%. At this degree of the material purity, the density of impurity atoms on a cleaved surface should be at a level of ≈ 1 atom/$\mu\mathrm{m}^2$. The dislocation density on the cleaved surface determined by counting pits after etching in dilute nitric acid was of order $0.05\ \mu\mathrm{m}^{-2}$. (Note that the same samples or those manufactured by the same technique were previously used in studies of the cyclotron resonance[4,15] and quantum oscillations of the quasistatic conductivity.[16] According to those measurements, the electron mean free path is of order one millimeter, so the high quality of initial bulk crystals is evident.)

In the middle of a sample (at half its height), a small notch was made when the sample was cut off from the initial crystal to define the cleaved surface position. Samples were etched in nitric acid in order to strip off the outer contaminated layer and washed in distilled water. Immediately before cleaving, the sample mounted in the scanning tunneling microscope was heated to 200–250 °C in vacuum in order to remove water and other volatile materials from its surface and avoid contamination of the cleaved surface in measurements of the heating effect on its properties.

In our experiments we have used the scanning tunneling microscope described in detail elsewhere.[17] Its basic characteristics are as follows: for voltages of 0–200 V applied to the x and y PZT drives and ± 24 V to the z drive, the scanning area is $1\times 1\times 0.05\ \mu\mathrm{m}^3$ at the liquid-helium temperature. Using three piezoinertial motors, we could move the tip of the scanning tunneling microscope along all three axes in steps of 0.1–1 μm and select an area to be tested within a range of several millimeters. When the tip approaches a sample, it is stopped automatically at the moment when a current between the tip and sample is detected. The samples were mounted in the scanning tunneling microscope holder[18] so that their upper halves jutted out and they could be fractured *in situ* by hitting with one end of a released spring. After that the tip was driven to the area selected for the experiment.

Low-temperature experiments were performed in a cryostat described elsewhere.[19] When gaseous helium is fed to the evacuated volume and a pressure of 10^{-3}–10^{-4} Torr builds up, the tunneling microscope and sample cool down to a temperature about 1 K above liquid-helium temperature. This is the temperature at which the reported measurements were performed. The presence of gaseous helium has no effect on recorded images. The sample and the tunneling microscope tip could be seen through a window transparent to visible light, which could be closed when necessary with a magnetically driven shutter. The sample could be heated *in situ* by feeding a current through a heater mounted in the holder.[18] One watt of dissipated electric power is sufficient to heat bismuth to its melting temperature even in the presence of the heat-exchange gas. The sample temperature was measured using a copper-constantan thermocouple.

The tips of the scanning tunneling microscope were platinum. They were fabricated either by cutting a wire or by electro-chemical etching of a wire with subsequent ablation of the surface layer using 2.5-keV argon ions. The tips were heated by an electron beam to the light-red color *in situ*.

The scanning tunneling microscope was driven by a computer with internal ADC and DAC cards, which allowed us to apply digital feedback in order to maintain the tunneling current and define all driving voltages required for the scanning microscopy and spectroscopy. Measurements were performed in the following modes.

1. Recording of $z(x,y)$ topograms when the tip was scanned over the xy plane (the x-axis was that of the line or frame scan, the y-axis was that of the frame or line scan, respectively, depending on the user's choice) with the feedback on and at prescribed tunneling current I and voltage U between the tip and sample.

2. Simultaneous recording of several frames at several prescribed values of the tunneling current I_k. In this case, each line is scanned several times and I_k is switched to a new value at the end of each line. After running through all prescribed values, the current returns to its initial value and the tip moves to a new line. The characteristic scan time of one line is about 0.1 s, so frames corresponding to different currents are not shifted with respect to each other because the temporal drift is within 0.1 Å/s.

3. Simultaneous recording of several frames at several prescribed voltages U_i between the tip and sample, which was performed similarly to the procedure described in the previous paragraph. Note that measurements performed in the latter two modes not only yielded frames recorded under

various experimental conditions, but also allowed us to derive from the shifts of recorded figures along the z-axis the effective work function W averaged over all points of the frame (usually 128×128 or 256×256). If W was about 4–4.5 V, the current undoubtedly flows in the vacuum-tunneling regime. At smaller W, the tip and/or sample surface were almost surely contaminated, had nonconducting coatings, and were in mechanical contact. In the experiments described below, either the condition of vacuum tunneling was satisfied, or a ''weak'' mechanical contact took place, which had no apparent effect on the results. Therefore, the condition of vacuum tunneling will be taken for granted in what follows.

4. Simultaneous recording of the topogram at a fixed voltage U_1 and of the map of current distribution at a different voltage U_2. In this mode, the tunneling gap was stabilized at each point at voltage U_1, and the corresponding value of z was measured, then the voltage driving the tip along the z-axis was fixed, the voltage between the sample and tip was switched to U_2, and after a delay of about 10 ms, which was required for termination of transient processes, the tunneling current was measured. Then the voltage was returned to the initial value U_1, the feedback was switched on, the tip moved to another point, and the procedure was repeated.

5. Measurements at given points of current–voltage characteristics within certain limits with accumulation of several measurements. In this case, the feedback system set a prescribed current I_0 at the initial voltage U_0, then the voltage driving the tip along the z-axis was fixed, the curve of $I(U)$ was recorded within a prescribed range of voltage U, the initial voltage U_0 was set, and the cycle was repeated if necessary. The typical recording time of one cycle was about 0.1 s. At the temperature of liquid helium the drift along the z-axis during this time was negligible.

6. Measurement of differential current–voltage characteristics at specified points within certain ranges with accumulation of sequential measurements. The difference from the previously described mode was that an ac component with a frequency of 20 kHz was added to the voltage U across the tunneling gap and the ac component of the tunneling current was measured using a lock-in amplifier.

3. EXPERIMENTAL RESULTS

Atomically smooth terraces with a characteristic size of a fraction of a micrometer are formed on bismuth cleaved surfaces both on the basal trigonal plane and on the surface of the twin layer, whose symmetry is characterized by a twofold axis (Fig. 1). On measurements made at the high resolution one can see an atomic-scale structure superposed on nonperiodic variations of the relief with the ''lemon peel'' shape with characteristic lateral sizes of about one nanometer, i.e., of order several lattice constants (Fig. 1b and 1c). On the surface of the trigonal plane, the lemon peel is isotropic in the sense that is has no specific direction. The structure on the surface of the twin layer is anisotropic because it shows valleys and ridges stretching along the twin boundary direction (Fig. 1b).

Neither the shape of the lemon-peel image nor its amplitude Δz depends on the tunneling current I (Fig. 2). Since the tunneling current is changed by varying the separation between the tip and sample, this experiment indicates that Δz is independent of this distance, at least over the range of 0.1 nm, which corresponds to changes in the tunneling current by a factor of ten.

The peak-to-trough variation Δz oof the relief superposed on the atomic structure varies with the voltage U following approximately the law $\Delta z\propto 1/U$ (Fig. 3a and 3c). When the voltage U is reversed, the image is inverted, i.e., we can see hills in place of valleys. When the voltage increases to the order of ten millivolts, the lemon-peel structure cannot be seen against the background of the noise. The atomic structure, whose amplitude varies more slowly,[10] can be seen at voltages of up to several hundred millivolts.

At low voltages of order one millivolt, the tunneling current becomes very unstable when the tip is scanned across the sample and even changes its sign. One could suppose that jumps in the current are due to the small separation between the sample and tip, so that sometimes the tip touches the surface because of the finite response time in the feedback loop with respect to noise. This, however, is not so, since one can observe changes in the current sign even when the current remains stable in the process of scanning. To this end, one should record simultaneously the topogram at the voltage U_1 (Fig. 3a) and the current distribution map (Fig. 3b) at the voltage U_2 so that the current averaged over the entire frame is zero.

Variations of the current and relief over a scale larger than the mean interatomic distance with the voltage U_1 correlate with one another, as one can see in Fig. 3. The distribution widths of the height Δz and current ΔI are shown in Fig. 3. These parameters were defined as the widths of appropriate histograms (Fig. 3c and 3d) at the level 0.1 so as to leave out overshoots. The choice of the cut-off level, however, has little effect on the ultimate result, which is that both ΔI and Δz are inversely proportional to U. Note also that $\Delta I\propto I$ holds to within 10–20%.

Current variations obviously indicate the presence of an additional voltage $\Delta U(x,y)$ applied between the tip and sample, whose amplitude is a function of the tip lateral coordinates x and y. The order of ΔU can be estimated using the obvious relation $\Delta U\approx\Delta I\cdot R$, where R is the resistance of the tunneling gap set when the feedback of the scanning tunneling microscope is turned on. For the case illustrated by Fig. 3, we have $\Delta U\approx0.6$ mV.

These results indicate that relief variations of the lemon-peel shape and current at zero voltage have much in common. Moreover, changes in z are mostly determined by variations in the voltage $U_1+\Delta U(x,y)$ between the tip and sample. The higher the relative contribution of the second term, the larger the tip displacement along the z-axis needed to maintain the prescribed value of the tunneling current. It is also clear that changing the sign of the voltage U_1 should invert the relief pattern. At the same time, a change in the value of the current I to be maintained constant at fixed U_1 should not change Δz, which is illustrated by Fig. 2. Numerically, a change in the current of about 0.2 nA (Fig. 3) should

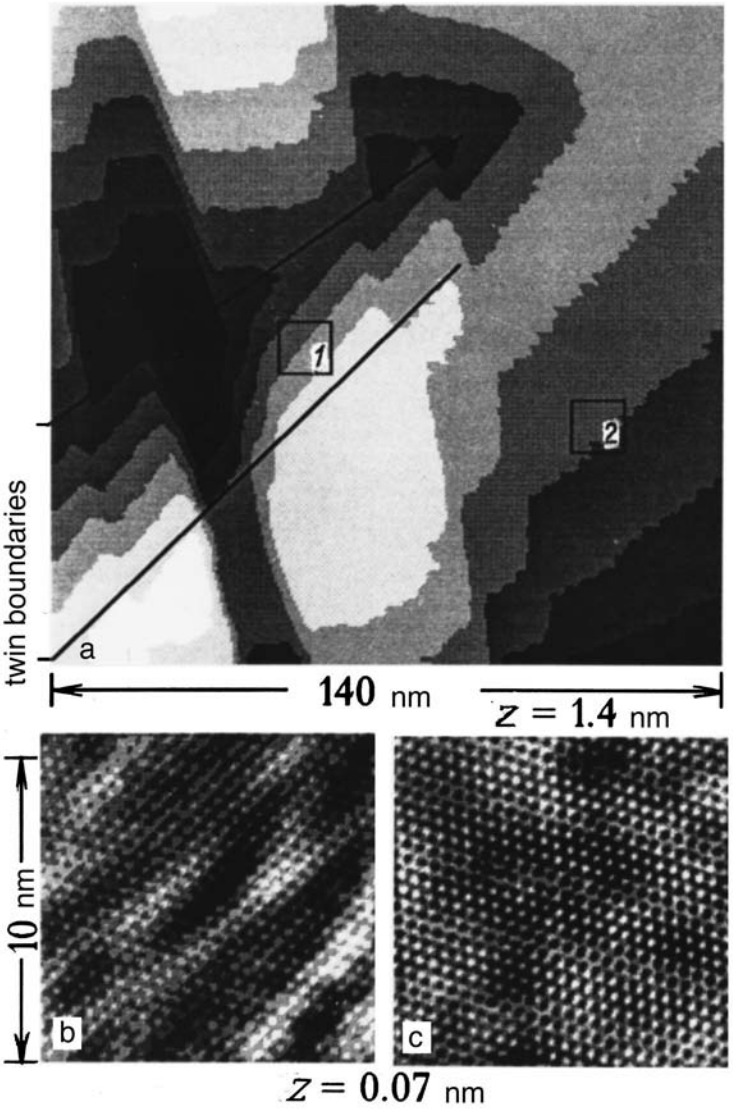

FIG. 1. (a) Topogram of bismuth surface in the region of the twin layer. Topograms of (b) area 1 on the surface of the layer with the two-fold axis and (c) area 2 on the basal plane with the trigonal symmetry. In order to emphasize atoms on the surface, the amplitude of atomic surface modulation is enhanced by a factor of four using a mathematical procedure (components of the two-dimensional Fourier transform corresponding to the atomic period are multiplied by four). The parameters of the experiment are the following: $I = 0.5$ nA, the voltage between the tip and sample (a) $U = 12$ mV; (b) 1 mV; (c) 3 mV.

correspond to $\Delta z \approx 0.04$–0.07 nm, judging by the measurements of the mutual displacements of topograms shown in Fig. 3 along the z-axis. This value is close to the measured $\Delta z \approx 0.08$ nm, which is also plotted in this graph. Given that the uncertainty is relatively large and ΔI can also be renormalized considerably owing to variations in z, the agreement between the numbers can be deemed quite satisfactory.

A very interesting fact is that point defects of the atomic scale do not produce features on current maps. Thus, on the surface relief shown in Fig. 4b one can clearly see two de-

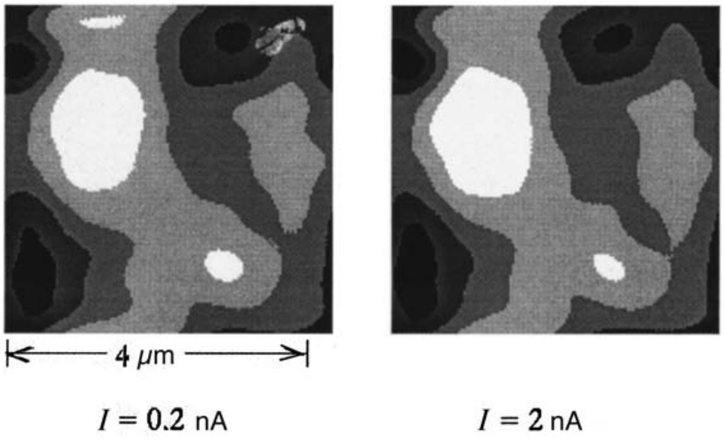

FIG. 2. Smoothed images of the same region on the surface on the trigonal plane recorded at different tunneling currents. The voltage $U = 3$ mV and $\Delta z = 0.06$ nm.

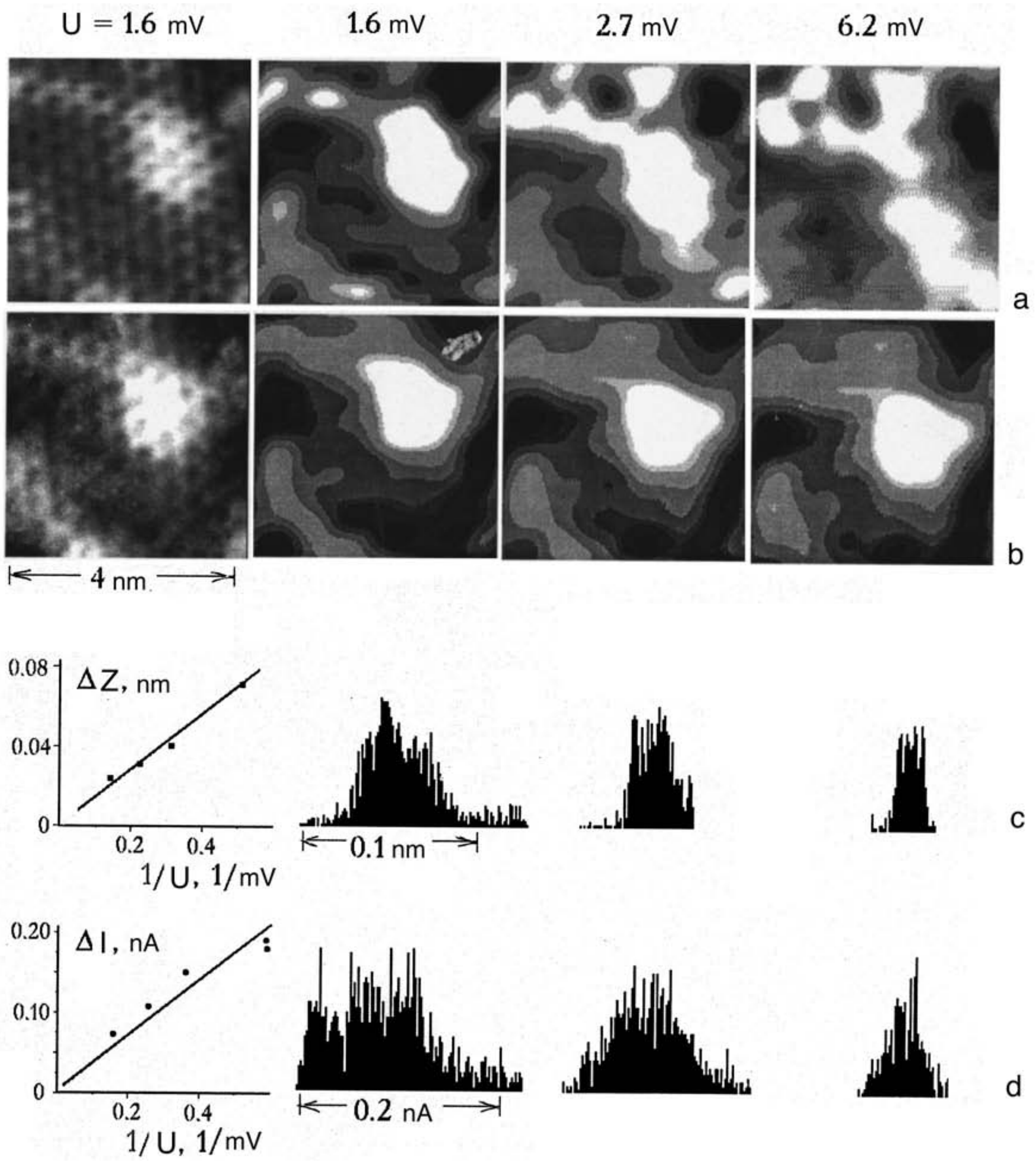

FIG. 3. (a) Original (left) and smoothed topograms and (b) current maps of the same region on the trigonal plane; (c, d) histograms for records (a) and (b), respectively. On the right of the histograms are curves of Δz and ΔI, respectively, as functions of voltage (which is shown over the graphs). The current at which the tunneling gap was fixed equaled 0.5 nA.

fects of the form of missing atoms. The lemon-peel structure is virtually unobservable in this graph, since U_1 is relatively high. It can be seen in the current map (Fig. 4c), where the point defects, in contrast, cannot be detected.

It follows from the results given above that on a seemingly ideal cleaved surface of a bismuth crystal, whose periodic atomic structure is clearly seen in topograms, nonuniform states with a characteristic sizes of one nanometer are established, and they appear in the form of an additional current initiation. They have been detected in tens of samples cleaved at low temperatures and on a sample cleaved at the room temperature (At temperatures above 350 K, bismuth single crystals become too plastic and we could not cleave them *in situ*.)

The mechanism of additional voltage initiation can be extracted from the fine structure of differential current–voltage characteristics measured at various arbitrary points on the atomically smooth sample surface. Three characteristics of this kind are plotted in Fig. 5. Each of them was recorded twice, and they show good reproducibility. The characteristics measured at different points are roughly similar, but at low voltages the differences between them becomes quite clear.

The changes in the current–voltage characteristics due to the tip displacement over the surface suggests that the additional voltage between the tip and sample is due to detection of some radiation by nonlinearity of the tunneling gap. In order to check out this hypothesis, we recorded current maps of the same tested area by varying the amplitude U_{200} of an ac voltage at a frequency of 200 kHz. Experiments were

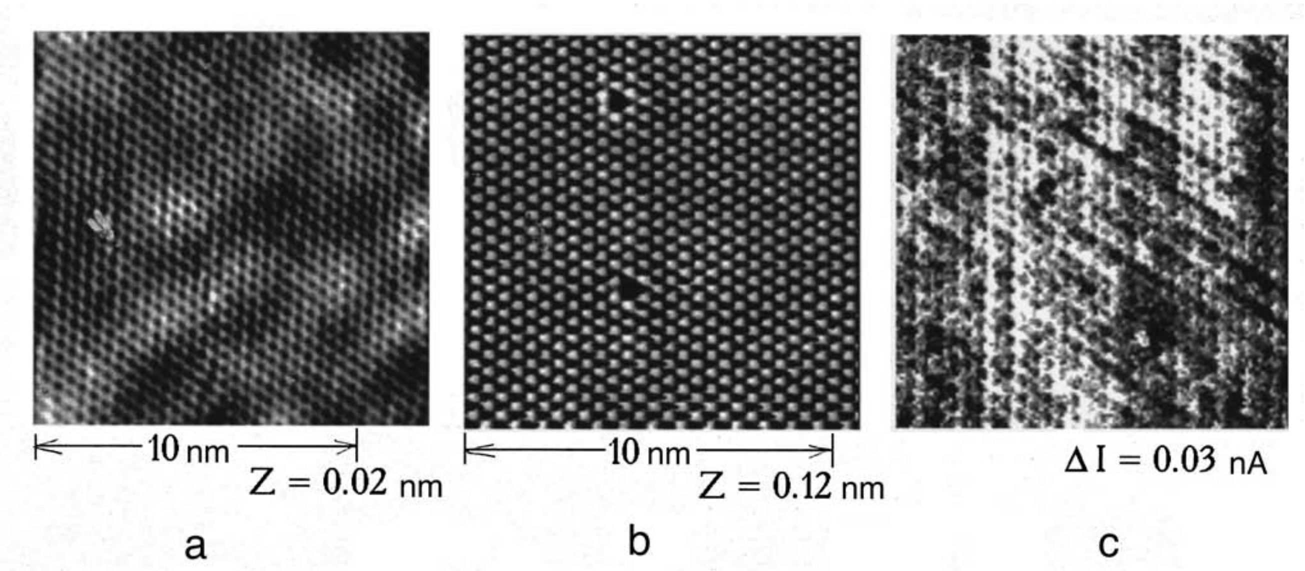

FIG. 4. Surface topograms of two different crystals recorded after heating them to about (a) 240 °C and (b) 200 °C. (c) Current map recorded simultaneously with topogram (b). Parameters of the experiment: (a) $I=1$ nA and $U=-1.5$ mV; (b, c) $I=1$ nA and $U=17$ mV.

carried out with several samples. This frequency is far beyond the cut-off frequency of the tunneling current preamplifier, which is less than 15–20 kHz. In recording these images, voltage U_1 was set at a fairly high level with a view to limiting variations of Δz and rule out their possible effect on measurements of ΔI.

Figure 6b shows three current maps at ac signal amplitudes $U_{200}=30$, 15, and 0 mV. It is clear that, although some fine structure of these maps is different (which is partly due to noise), there is correlation among the three maps in terms of the positions of the main features on the xy plane. At the same time, the range of variations in the current essentially broadens with the ac voltage amplitude, as can be seen in histograms of the current distribution (Fig. 6c).

From the current–voltage characteristics given in Fig. 5, one can estimate the expected current variations to order of magnitude. At low voltages, the rectified current component should be $(1/4)U_{200}^2 \cdot d^2I/dU^2$. Numerical differentiation of the characteristics yields at $U=0$ the values of d^2I/dU^2 ranging between -10 and 20 nA/V^2 (Fig. 5b). Hence we derive at $U_{200}=30$ mV the rectified currents ranging between -0.01 and $+0.02$ nA, i.e., the expected value can be $\Delta I \approx 0.03$ nA, which is close to the measurements in order of magnitude. (Our calculations take into account the fact that in recording the plots of Fig. 6 we set the stabilized current and voltage across the gap a factor of two higher and lower, respectively, than in measurements of the current–voltage characteristics (Fig. 5). For this reason, the values of d^2I/dU^2 should be multiplied by four.)

Figure 7 shows ΔI as a function of U_{200} and U_{200}^2. One can see that the measurement accuracy is insufficient for an unambiguous choice between $\Delta I \propto U_{200}$ and $\Delta I \propto U_{200}^2$. The latter function is compatible with the suggested mechanism, provided that, along with the voltage at a frequency of 200 kHz an additional voltage is applied to the gap between the

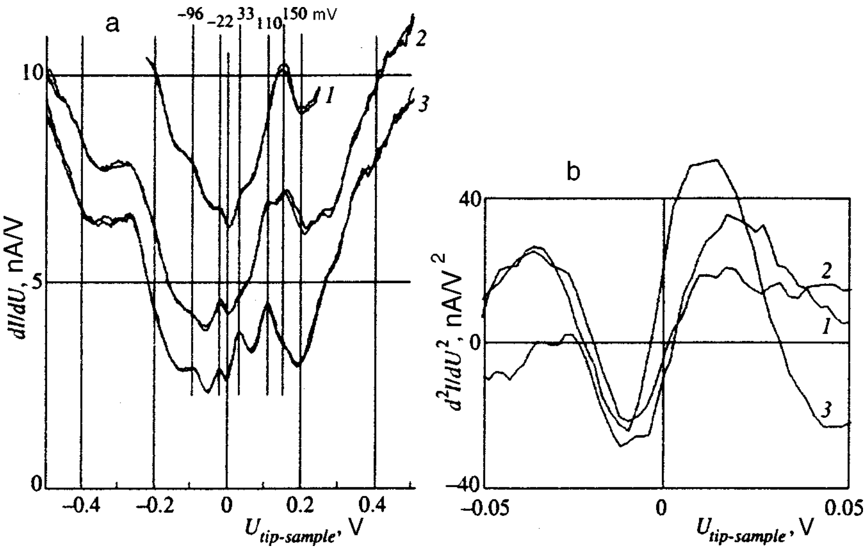

FIG. 5. (a) Differential current–voltage characteristics for three various points on the surface at a distance of ≈ 4 nm from one another. (b) Second derivative of current with respect to voltage obtained by numerical differentiation of curves (a). In measuring these characteristics, the tunneling gap was stabilized at $I=0.5$ nA and $U=0.125$ mV.

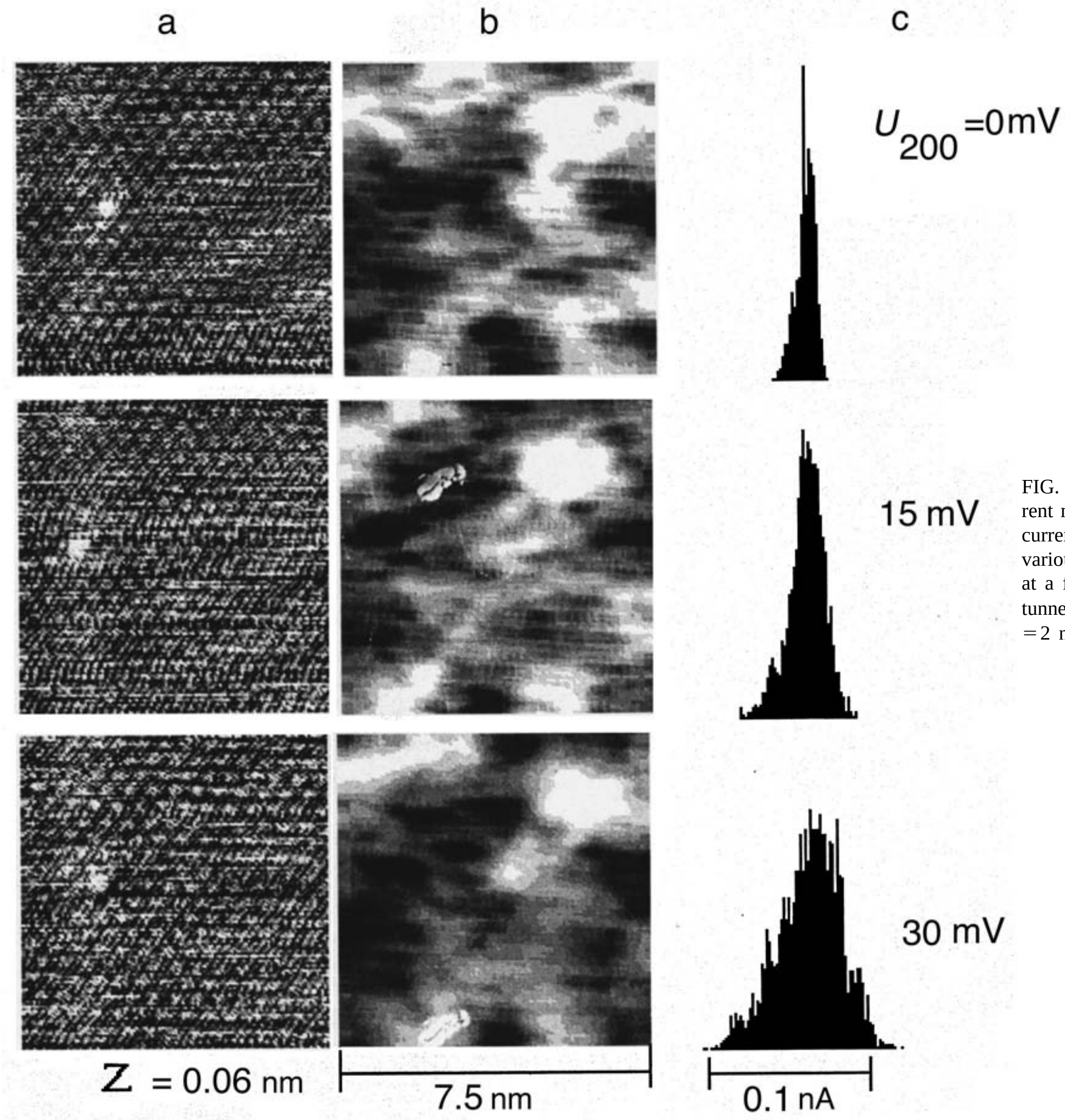

FIG. 6. (a) Topograms, (b) current maps, and (c) histograms of current distribution recorded at various amplitudes of ac voltage at a frequency of 200 kHz. The tunneling gap was stabilized at $I = 2$ nA and $U_1 = 0.06$ V.

tip and sample, and the powers of these two signals add. Note, however, that the ac signal amplitude is not much smaller than the typical voltages at which the nonlinearity parameter of current–voltage characteristics essentially changes (Fig. 5), so deviations from the quadratic curve should be noticeable.

The issue of the source of additional voltage has not been ultimately clarified. Figure 7 indicates that its amplitude is estimated to be about 10 mV, and the power dissipated in the tunneling gap at a level of 10^{-10}–10^{-12} W. One can say that the preamplifier noise of the scanning tunneling microscope in terms of the input voltage cannot have such an

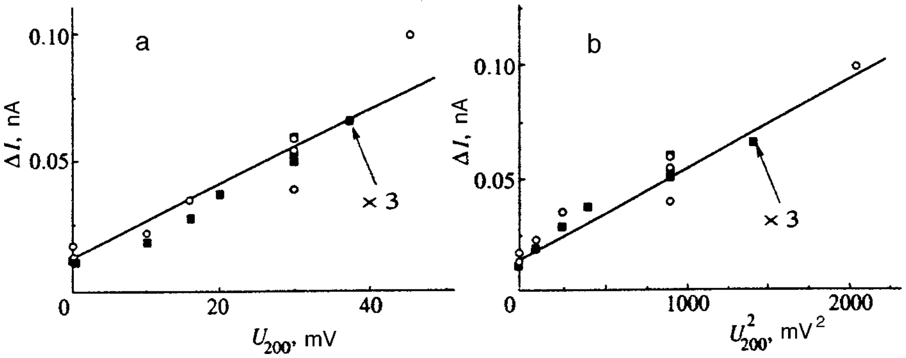

FIG. 7. Current variations versus (a) amplitude and (b) amplitude squared of the 200-kHz ac voltage component. Circles and squares correspond to different samples. (Values of ΔI for squares are multiplied by three.)

effect. Another plausible source of this voltage is light fed through the window, but it has no effect since the measurements do not depend on whether the window is open or closed by a copper foil shutter set on the helium screen of the scanning tunneling microscope on the side of the evacuated volume. Oscilloscope traces of the voltage at the preamplifier input indicated that the most probable source of this voltage is a spurious signal at the TV frequency of about 200 MHz. Note, however, that the uncertainty in this issue became inconsequential in the context of the reported experiments when we had amplified the effect under investigation by applying the definite ac voltage.

Based on the reasoning given above, we conclude that the mechanism leading to the nonuniformity in the relief and current maps has been established: these effects are due to differences among current–voltage characteristics at different points of the sample surface, i.e., nonuniformity of electron states. Obviously, such effects are out of question in an ideal lattice. Since the quality of the initial crystals was very high,[15,16] the only feasible explanation is generation of defects near the surface in the process of cleavage. In order to obtain additional information about their nature, we attempted to anneal crystals *in situ*. Various samples were heated from the room temperature to 200–240 °C (recall that the melting temperature of bismuth is 271 °C) and exposed to this temperature for about one minute. (The total time during which the sample temperature was higher than the room temperature was about ten minutes.) In all cases, nonuniform states persisted after cooling down, and features on the records were essentially the same as on freshly cleaved surfaces. Examples of records obtained on two different samples after annealing are shown in Fig. 4. Thus, a relatively high barrier had to be overcome to eliminate the defects. If the frequency of hits against the barrier is set at the upper frequency of acoustic phonons in bismuth, 1.5×10^{12} Hz,[20] we have a lower estimate of the barrier height of 1.5–2 eV.

Probably, a longer annealing could get rid of the lemon peel, but this would considerably complicate the experiment, in particular, this would sharply increase the liquid-helium consumption. Moreover, one should take into account that the surface composition can be changed by segregation of impurities from the bulk to surface, surface diffusion of foreign atoms from sample sides, and changes in the structure near the surface owing to the motion of terraces, which is noticeable even at the room temperature and greatly activated by heating.[9] Therefore the measurement of the activation energy of such defects is a separate complex task.

4. DISCUSSION

Many publications have been dedicated to the fracture dynamics of brittle crystals. To the best of our knowledge, however, none of these papers discussed formation of defects near cleaved surfaces and their structures. Nor is there a theory of this phenomenon available, so our discussion is limited to approximate estimates and qualitative considerations.

The lemon peel goes together with undisturbed atomic surface layers, so we can assert that it is an indication of crystal defects not associated with changes in the mean density. Such defects could be, for example, vacancies and interstitial atoms localized near the surface at depths of about one nanometer, which are comparable to characteristic lateral sizes of features in the structure. Probably these defects are generated when atoms are displaced from their initial positions in the process of cleavage. The formation and relaxation energies of vacancies and interstitial atoms are several electronvolts. (This is why they are not eliminated by annealing at about 500 K.) The surface density of defects observed in our experiments, which is a factor of several tens lower than the surface density of host atoms, and their random distribution indicates that these defect result from random events when momenta transferred from surface atoms add at some lattice sites. We do not try to describe the scenario of this process, but only indicate that a description in terms of phonons is hardly adequate since their maximal energy in bismuth is only about 8 meV for acoustic phonons and about 13 meV for optical phonons.[20]

In principle there can be two causes of the lemon-peel pattern: statistical variations of the defect density, which is fairly high, or sparsely distributed weakly interacting defects. The analysis of current–voltage characteristics (Fig. 5) reveals that they have different amplitudes of features located at the same voltages, i.e., there is a limited number of discrete features in the electron spectrum. This circumstance makes the second cause more probable.

The feature in the electron spectrum at -0.3 V is probably due to the T_6^+ level in the bulk with an energy of 0.38 eV above the Fermi level, which was calculated by Lin and Allen.[21] Other features that could be associated with defects in the bulk are peaks at -22 and $+33$ mV in Fig. 5 (the valence band top at 11 meV and levels at -27 and -40 meV corresponding to conduction electrons[4,21]). The dimensions of wave functions corresponding to these levels, however, are tens of nanometers, and their relative contribution could not change as a result of lateral translations through smaller distances. Therefore, all features in the spectrum at voltages between the tip and sample ranging between -0.2 and $+0.4$ V reflect properties of surface states.

A fraction of these features can be attributed, in principle, to surface levels of the two-dimensional lattice. Using the estimate $pr\approx\hbar$ and equating the size r to the interatomic distance on the surface of 0.45 nm, we obtain for these states the characteristic electron energy $p^2/2m\approx0.1$ eV. (The mass m is equated to the free electron mass.) Features at lower energies should correspond to smaller sizes, i.e., the structure dubbed lemon peel in this paper. Since the calculation techniques for electron spectra have been well developed, there is hope that the real structure of observed defects will be revealed by the theory in future if it is possible to compare experimental data with calculated by different models of defects.

The estimates given above explain why surface features of the atomic scale cannot be seen in current maps (Fig. 4b and 4c). They can be seen only at sufficiently high ac voltage amplitudes across the gap between the tip and sample. As follows from our preliminary experiments, this is true, and the amplitude of the atomic-scale pattern in the current map increases sharply when the ac voltage amplitude surpasses 100 mV, so that the contribution from surface defects of the atomic scale become dominant.

To sum up, we focus the reader's attention on the following: we have reported on the first ever observation of a nonuniform state on the cleaved crystal surface. This phenomenon can probably be detected by different techniques, and information about the structure of the corresponding defects can be obtained by studying the surface with the help of such techniques as grazing-incidence x-ray diffraction or electron diffraction. Another important aspect is theoretical investigation of both the mechanism of formation of the nonuniform state and the electron spectrum on the cleaved surface, as well as the search for similar effects in other materials. The work in this line of research will continue.

We are grateful to A. F. Andreev for support of this work, and to A. P. Volodin and I. N. Khlyustikov for helpful discussions. The work was supported by the Russian Fund for Fundamental Research (Grant No. 96-02-18991).

*)E-mail: edelman@kapitza.ras.ru

---

[1] M. B. Johnson, O. Albrektsten, R. M. Feenstra, and H. W. M. Salemink, Appl. Phys. Lett. **63**, 2923 (1993).

[2] J. F. Zheng, X. Liu, N. Newman, E. R. Weber, D. F. Ogletree, and M. Salmeron, Phys. Rev. Lett. **72**, 1490 (1994).

[3] J. Chao, A. R. Smith, and C. K. Shih, Phys. Rev. B **53**, 6935 (1996).

[4] V. S. Édel'man, Usp. Fiz. Nauk **123**, 257 (1977) [Sov. Phys. Usp. **20**, 819 (1977)].

[5] Yu. F. Komnik, U. I. Bukhshtab, Yu. V. Nikitenko, and V. V. Andrievskiĭ, Zh. Éksp. Teor. Fiz. **60**, 669 (1971) [Sov. Phys. JETP **33**, 364 (1971)].

[6] C. A. Hoffman, J. R. Meyer, and E. J. Bartoli, Phys. Rev. B **48**, 11431 (1993).

[7] V. S. Tsoi and I. I. Razgonov, JETP Lett. **23**, 92 (1975).

[8] G. Jezequel, J. Thomas, and I. Pollini, Phys. Rev. B **56**, 6620 (1997).

[9] A. M. Troyanovskiĭ and V. S. Édel'man, JETP Lett. **60**, 111 (1994).

[10] A. M. Troyanovskiĭ and V. S. Édel'man, JETP Lett. **60**, 300 (1994).

[11] A. M. Troyanovskiĭ and V. S. Édel'man, Kristallografiya **44** (2), 336 (1999) [Crystallogr. Rep. **44**, 300 (1999)].

[12] V. S. Édel'man, Phys. Lett. A **210**, 105 (1996).

[13] V. S. Édelman, D. Yu. Sharvin, I. N. Khlyustikov, and A. M. Troyanovskii, Europhys. Lett. **34**, 115 (1996).

[14] M. S. Khaĭkin, S. M. Cheremisin, and V. S. Édel'man, Prib. Tekh. Éksp. No. 4, 225 (1970).

[15] V. S. Édel'man and M. S. Khaĭkin, Zh. Éksp. Teor. Fiz. **49**, 107 (1965) [Sov. Phys. JETP **22**, 77 (1966)].

[16] V. S. Édel'man, Zh. Éksp. Teor. Fiz. **64**, 1734 (1973) [Sov. Phys. JETP **37**, 875 (1973)].

[17] V. S. Édel'man, A. M. Troyanovskii, M. S. Khaikin, G. A. Stepanyan, and A. P. Volodin, J. Vac. Sci. Technol. B, Second ser. **9**(2), Pt. II, 618 (1991).

[18] V. S. Édel'man, Prib. Tekh. Éksp. No. 4, 203 (1994).

[19] I. N. Khlyustikov and V. S. Édel'man, Prib. Tekh. Éksp. No. 1, 158 (1996).

[20] J. L. Yarnell, J. L. Warren, R. G. Wenzel, and S. H. Koenig, IBM J. Res. Dev. **8**, 234 (1964).

[21] Y. Lin and R. E. Allen, Phys. Rev. B **52**, 1566 (1995).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Nuclear magnetic resonance of ^{55}Mn in the antiferromagnet CsMnBr$_3$ in a variable longitudinal magnetic field

B. S. Dumesh*)

Institute of Spectroscopy, Russian Academy of Sciences, 142092 Troitsk, Moscow Region, Russia

M. I. Kurkin

Institute of Metal Physics, Urals Branch of the Russian Academy of Sciences, 620219 Ekaterinburg, Russia

S. V. Petrov and A. M. Tikhonov

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

(Submitted 20 January 1999)
Zh. Éksp. Teor. Fiz. **115**, 2228–2241 (June 1999)

The spectrum and intensities of NMR lines are investigated experimentally and theoretically for excitation by an alternating magnetic field $\mathbf{h}_{\|}$ parallel to a static field $\mathbf{H}$ in the quasi-one-dimensional, six-sublattice antiferromagnet CsMnBr$_3$. According to theory, two new NMR lines, which are not excited by a transverse magnetic field $\mathbf{h}_{\perp}$, are observed near the phase transition from triangular to collinear structure ($H=H_c$) [JETP **86**, 197 (1998)]. *© 1999 American Institute of Physics.* [S1063-7761(99)02506-8]

1. INTRODUCTION

The application of NMR methods to studies of the magnetic properties of quasi-one-dimensional, multiple-sublattice antiferromagnets has already helped to produce nontrivial results, such as the phenomenon of suppression of quantum fluctuations of electron spins[1] and a new type of magnetic structure in the easy-axis triangular antiferromagnet CsMnI$_3$.[2] Nonetheless, it has become increasingly obvious that the full potential of NMR in such investigations have yet to be fully exploited. In this paper we discuss results obtained in the excitation of NMR by a longitudinal alternating magnetic field $\mathbf{h}_{\|}$ parallel to a static magnetic field $\mathbf{H}$. This method has been successful in disclosing two additional NMR lines that are not excited by a transverse rf magnetic field $\mathbf{h}_{\perp}$. These lines are intriguing in that they exhibit a dynamic frequency shift near the phase transition from triangular to collinear structure.

In Secs. 2 and 3 of the present paper, we describe the magnetic properties of CsMnBr$_3$ and experiments on the observation of ^{55}Mn NMR for $\mathbf{h}_{\|}$. In Sec. 4 we give the results of calculations of the spectrum and intensities of NMR lines for various excitation techniques. In the Conclusion we discuss the suppression of steady-state NMR signals in the presence of large dynamic frequency shifts and the possibilities inherent in the parametric excitation of nuclear spins.

2. MAGNETIC PROPERTIES OF CsMnBr$_3$

The compound CsMnBr$_3$ is one of the family of halides of the type ABX$_3$, where A denotes an alkali metal, B is a $3d$ metal, and X is a halogen. The crystal structure of CsMnBr$_3$ is described by the spatial symmetry group D^4_{6h}, the Mn^{2+} ions forming a hexagonal grid in the basal plane (perpendicular to the $\mathbf{C}_6$ axis).[4] The crystal lattice determines the detailed magnetic structure of this compound.[5–9] The principal distinguishing feature of the lattice is that the distance between adjacent planes of magnetic ions is half the distance between nearest-neighbor ions within one plane. As a result, antiferromagnetic exchange interaction of the magnetic moments within chains running along the $\mathbf{C}_6$ axis is 10^3 times the interchain antiferromagnetic exchange. This quasi-one-dimensionality significantly affects the magnetic properties, thus accounting for the heightened interest in the study of this class of materials.

The easy-axis character of the magnetic anisotropy in conjunction with antiferromagnetic interchain exchange results in the formation of a noncollinear, six-sublattice magnetic structure (Fig. 1a). Intrachain exchange induces antiferromagnetic ordering of the magnetic moments $\mathbf{M}_j$ ($j=1-6$) of three pairs of electronic sublattices, which is described by the antiferromagnetism vectors

$$\mathbf{L}_1=\mathbf{M}_1-\mathbf{M}_4,\quad \mathbf{L}_2=\mathbf{M}_2-\mathbf{M}_6,\quad \mathbf{L}_3=\mathbf{M}_3-\mathbf{M}_5.$$

Because of the vanishingly weak magnetic anisotropy in the basal plane, the sublattices in a weak magnetic field $\mathbf{H}\perp\mathbf{C}_6$ are oriented in such a way that one of the indicated vectors $\mathbf{L}_i$, say $\mathbf{L}_1$, is perpendicular to $\mathbf{H}$ (Fig. 1a). The other two vectors, $\mathbf{L}_2$ $\mathbf{L}_3$, form angles close to 30° and 150° with $\mathbf{H}$.

As $\mathbf{H}$ is increased, the angle α between $\mathbf{L}_2$ and $\mathbf{L}_3$ varies according to the law[10]

$$\cos\frac{\alpha}{2}=\frac{1}{2-z},\qquad z=\frac{H^2}{H_c^2},\tag{1}$$

where $H_c=\sqrt{H_E H_{E'}}\approx 61\,\mathrm{kOe}$ (at $T=1.8\,\mathrm{K}$, Ref. 9), $H_E\approx 1500\,\mathrm{kOe}$, and $H_{E'}\approx 3\,\mathrm{kOe}$ are the effective fields of intrachain and interchain exchange interactions, respectively.

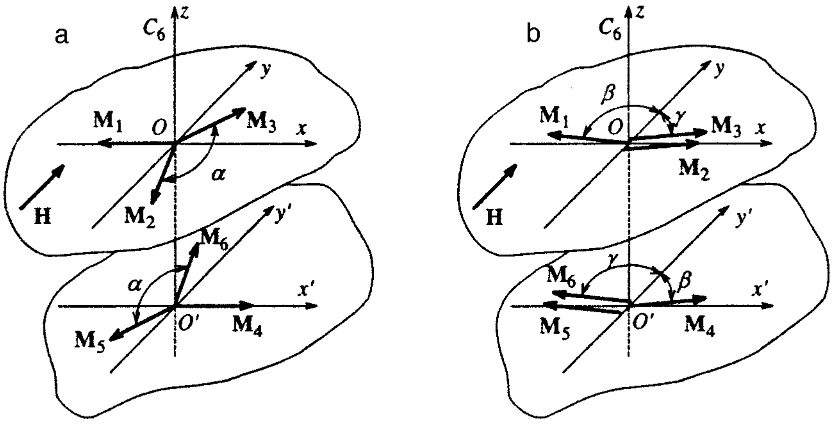

FIG. 1. Schematic representation of the magnetic structure of CsMnBr$_3$: a) $H \ll H_c$; b) $H > H_c$ ($\beta \simeq \gamma$).

In a field $H=H_c$ the above-described magnetic structure changes to a collinear configuration ($\alpha=0$), corresponding to a second-order phase transition (Fig. 1b).

The magnetic field of the nuclei of the jth sublattice is determined by the sum of the external field $\mathbf{H}$ and the hyperfine fields $\mathbf{H}_{nj}$:

$$H_j=|\mathbf{H}_{nj}+\mathbf{H}|=H_n\sqrt{1+\frac{H^2}{H_n^2}-2\frac{H}{H_n}\cos\theta_j}, \tag{2}$$

where $H_n=-AM_0$, M_0 is the average magnetic moment of the sublattice, A is the hyperfine interaction constant, and θ_j is the angle between $\mathbf{H}$ and $\mathbf{M}_j$:

$$\begin{aligned}
&\cos\theta_{1,4}=\frac{H}{H_E},\\
&\cos\theta_{2,5}=-\sin\frac{\alpha}{2}+\frac{H}{H_E}\cos^2\frac{\alpha}{2}+o\left(\frac{H}{H_E}\right),\\
&\cos\theta_{3,6}=\sin\frac{\alpha}{2}+\frac{H}{H_E}\cos^2\frac{\alpha}{2}+o\left(\frac{H}{H_E}\right).
\end{aligned} \tag{3}$$

Consequently, for $H<H_c$ there must be three twofold degenerate NMR branches $\omega_{nj}=\gamma_n H_j$.

In weak fields H the degeneracy is lifted by interaction with the Goldstone antiferromagnetic resonance (AFMR) mode. The frequencies of the three NMR branches (Ω_4, Ω_5, and Ω_6) decrease (this is the so-called dynamic frequency shift). Their spectrum has been investigated experimentally and theoretically.[1,3] A functional dependence $H_n(H)$ that differs for spins in sublattices 1, 4 and 2, 3, 5, 6 has also been reported in the cited papers, owing to the suppression of quantum fluctuations of the magnetic field and a corresponding increase in $M_j(H)$. The spectrum of all other branches is described by Eqs. (2)and (3) with the functional $H_{nj}(H)$ taken into account, but NMR signal amplification does not take place for these branches, and they have not been observed experimentally.

As $H \to H_c$, two of these branches begin to interact with the AFMR mode ω_5 (in the notation of Ref. 8), whose frequency tends to zero as phase transition is approached. The spectrum of these branches is deformed in this case, and the branches themselves become observable in excitation by an rf field $\mathbf{h}_\parallel$. These phenomena are the subject of the present article.

3. MEASUREMENT PROCEDURE AND DESCRIPTION OF THE EXPERIMENT

The objects of investigation were CsMnBr$_3$ single crystals grown and oriented as in Ref. 3, which also describes the wide-range continuous NMR spectrometer used to perform the measurements. The main difference is a modification of the cavity structure to impart the required polarization to the rf field $\mathbf{h}_\parallel$. A block diagram of the resonance circuit is shown in Fig. 2. A movable copper plate *2* with a dielectric coating *3* forms with the casing an additional variable capacitance, which is used to tune the cavity frequency. A narrow slot *7* forms the structural capacitance of the loop. The two-headed arrow indicates the directions of motion of the plate. The whole structure is positioned in a superconducting

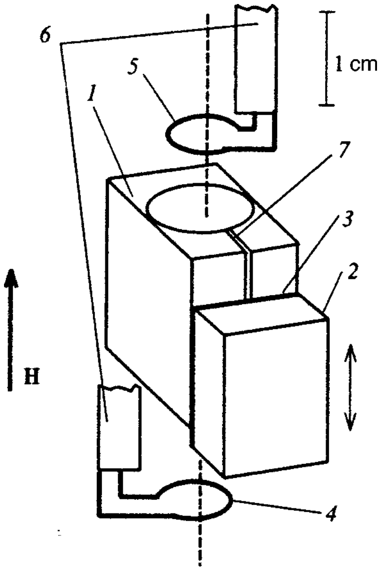

FIG. 2. Block diagram of the resonance circuit: (*1*) cavity; (*2*) movable copper plate; (*3*) thin insulating film; (*4, 5*) coupling loops; (*6*) coaxial leads; (*7*) narrow slot.

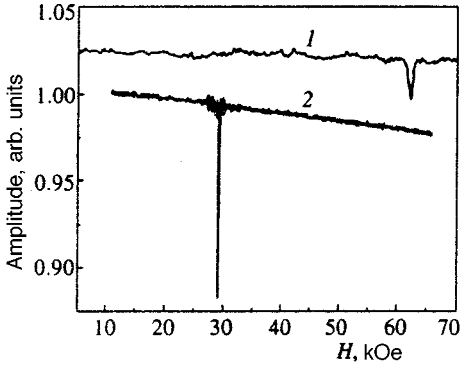

FIG. 3. Experimental absorption signal in CsMnBr$_3$ ($\mathbf{H}\perp C_6$) at $T=1.3$ K and 349.6 MHz with excitation by alternating magnetic fields $\mathbf{h}_\parallel$ (curve *1*) and $\mathbf{h}_\perp$ (curve *2*).

solenoid with an inside diameter of 25 mm, and the rf field is precisely parallel to the field of the solenoid **H**.

We used two cavities with $Q\approx 400$ at 4.2 K. The frequency tuning range of one cavity was from 390 MHz to 470 MHz, and the range of the other was from 310 MHz to 380 MHz. The cavity containing the investigated single-crystal sample was placed directly in a helium tank. An external magnetic field was applied perpendicular to the hexagonal $\mathbf{C}_6$ axis of the crystal. In all other respects the spectrometer and the measurement procedure were identical to those in Ref. 3.

Figure 3 shows the absorption signal in CsMnBr$_3$ for $T=1.3$ K, a frequency of 349.6 MHz, and fields $\mathbf{h}_\parallel$ (curve *1*) and $\mathbf{h}_\perp$ (curve *2*). It is evident that different NMR branches are excited in these two cases. The NMR spectrum in CsMnBr$_3$ for $\mathbf{h}_\parallel\perp\mathbf{C}_6$ at $T=1.3$ K is represented by light circles in Fig. 4. NMR is observed close to H_c over a broad frequency range, demonstrating the large dynamic frequency shift of NMR. As $|H-H_c|$ increases, the intensity of the signal decreases, and its position approaches the unshifted NMR spectrum, which is represented by dashed curves. The solid curves represent the NMR spectrum calculated from Eqs. (20) and (21) below. We have not used any fitting constants here. Satisfactory agreement is observed between the experimental and calculated spectra. The observed differences are attributable to the appreciable width of the AFMR line near phase transition.

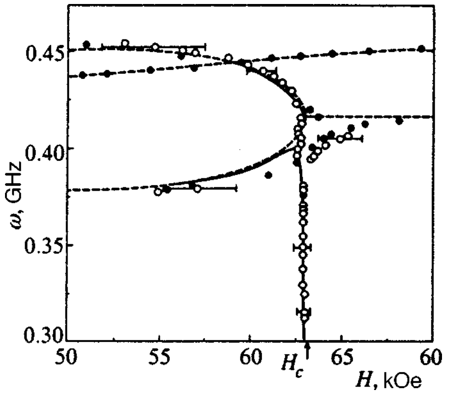

FIG. 4. NMR spectrum in CsMnBr$_3$ for $\mathbf{H}\perp C_6$ at $T=1.3$ K with excitation by alternating magnetic fields $\mathbf{h}_\parallel$ (light circles) and $\mathbf{h}_\perp$ (heavy dots, from Ref. 3).

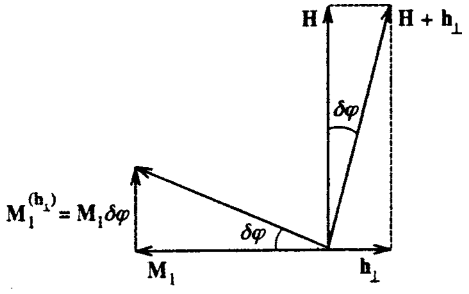

FIG. 5. Influence of $\mathbf{h}_\perp$ on the orientation of **M**.

It must be noted that NMR in parallel fields has heretofore been observed only in superfluid helium ^{3}He and in the domain walls of ferromagnets. In our case NMR is generated in a homogeneous sample, because the signal is observed at a large distance from the phase transition field.

Consequently, five NMR modes are observed in CsMnBr$_3$. Three of them are excited for $\mathbf{h}_\perp$, interact with the Goldstone AFMR mode, and are observed in the range of fields 20–80 kOe.[3] Two modes are generated for $\mathbf{h}_\parallel$, interact with the AFMR mode ω_5, and are observed in fields of 50–65 kOe.

4. THEORY

We analyze the intensities of NMR lines excited by various techniques, using the same equations for the magnetizations $\mathbf{m}_j$ ($j=1,\ldots,6$) of the nuclear sublattices as in Ref. 3. Now, however, in these equations we need to take into account, first, interaction with differently polarized alternating fields **h** and, second, nuclear magnetic relaxation, which takes the oscillations of $\mathbf{m}_j$ into the steady state. We treat relaxation processes in the relaxation time approximation, which corresponds to the Bloch equations[11] (see Appendix).

Figure 5 shows the influence of a transverse field $\mathbf{h}_\perp$ on the orientation of the vector $\mathbf{M}_1$ when the frequency of the alternating field is much lower than the AFMR frequency (at NMR frequencies this condition is easily satisfied at any H, owing to the hyperfine gap in the AFMR spectrum[12,13]). It is evident that for $h_\perp\ll H$ everything reduces to rotation through the angle $\delta\varphi=h_\perp/H$. All other vectors $\mathbf{M}_j$ rotate through the same angle, and their variations are therefore described by the equations

$$\mathbf{M}_j^{(h_\perp)}(t)=\mathbf{M}_j\delta\varphi=\chi_\perp\mathbf{h}_\perp(t), \tag{4}$$

where

$$\chi_\perp=M_j/H \tag{5}$$

is the magnetic susceptibility in a field $h_\perp$. It follows from Fig. 1 that a longitudinal field $\mathbf{h}_\parallel$ has scarcely any influence on the orientation of the vectors $\mathbf{M}_1$ and $\mathbf{M}_4$, but changes the angle α by

$$\delta\alpha=\begin{cases}4Hh_{\parallel}/\sqrt{(3H_c^2-H^2)(H_c^2-H^2)}, & H\leqslant H_c\,,\\ 0, & H>H_c\,.\end{cases}\qquad(6)$$

Equations (6) can be derived from Eq. (1) by writing the latter for the field $H+h_{\parallel}$ and expanding in powers of $h_{\parallel}$. The calculation of the variations of the vectors $\mathbf{M}_j$ for such a variation of the angle α leads to the equations

$$\mathbf{M}_j^{(h_{\parallel})}(t)=\chi_{\parallel j}\mathbf{h}_{\parallel}(t),\qquad(7)$$

where

$$\chi_{\parallel 1}=\chi_{\parallel 4}=0,\quad \chi_{\parallel 2}=\chi_{\parallel 6}=-\chi_{\parallel 3}=-\chi_{\parallel 5}=\chi_{\parallel}\,,\qquad(8)$$

$$\chi_{\parallel}(H)=\begin{cases}M_0\delta\alpha/2h_{\parallel}\,, & H\leqslant H_c\,,\\ 0, & H>H_c\,.\end{cases}\qquad(9)$$

We have thus calculated the NMR gains:

$$\eta_{\perp}=A\chi_{\perp}=H_n/H,$$

$$\eta_{\parallel j}=A\chi_{\parallel j}\,.\qquad(10)$$

In the Appendix we show that the use of Eqs. (6)–(9) in the equations for $\mathbf{m}_j$ permits them to be written in the form (A12):

$$\left[\left(\omega+\frac{i}{T_2}\right)^2-\omega_{nj}^2\right]m_{x_j}(\omega)$$

$$+\gamma_n\omega_{nj}Am_{z_j}\frac{H_n}{H_E'}\sum_{j}\lambda_{jj'}m_{x_{j'}}(\omega)$$

$$+\gamma_n\omega_{nj}m_{z_j}\eta_{\perp}h_{\perp}(\omega)+\gamma_n\omega_{nj}m_{z_j}\eta_{\parallel j}h_{\parallel}(\omega)=0.\qquad(11)$$

The Appendix also gives expressions for the quantities $\lambda_{jj'}$ (A7) and ω_{nj} (A13). The expressions for $m_{x_j}(\omega)$, $h_{\perp}(\omega)$, and $h_{\parallel}(\omega)$ are related to $m_{x_j}(t)$, $h_{\perp}(t)$, and $h_{\parallel}(t)$ by the Fourier transform (48).

The determinant of the system of equations (11) characterizes the six NMR frequencies. We note that the frequencies of only three of these lines, excited by a transverse field $\mathbf{h}_{\perp}$, were analyzed in Ref. 3. We now look into the feasibility of exciting all six lines.

Taking the amplification into account, we describe the intensity $I_n(\omega)$ of the absorption signal measured in the experiments of Ref. 3 by the equation

$$I_n(\omega)=\sum_j \eta_j\,\mathrm{Im}\,m_{x_j}(\omega),\qquad(12)$$

where $\mathrm{Im}\,m_{x_j}(\omega)$ is the imaginary part of the solution of the system (11) for the frequency ω. These equations have the simplest form in the cases $H\ll H_c$, $H\approx H_c$, and $H>H_c$. An analysis of the cases $H\ll H_c$ and $H>H_c$ yields results that agree with the curves in Refs. 1 and 3 to within the experimental errors. We therefore confine our discussion to the case $H\approx H_c$ only, as it is associated with new experimental results described in the preceding section of the article.

For $H\approx H_c$ the following expressions for $\lambda_{jj'}$ can be obtained from Eq. (A7):

$$\lambda_{22}=\lambda_{33}=\lambda_{55}=\lambda_{66}=\lambda_{26}=\lambda_{35}=-\lambda_{23}=-\lambda_{63}=-\lambda_{25}$$

$$=-\lambda_{56}=0.5\{2[1-(H/H_c)^2]+\epsilon\}^{-1},\qquad(13)$$

where $\epsilon=2m_0/H_{E'},\approx 10^{-2}$ [see Eq. (A5)].

The components λ_{1j} and λ_{4j} do not have singularities at $H=H_c$, and their influence can therefore be disregarded. A second procedure by which it is possible to substantially simplify the system of equations (11) involves the transformation to new variables $m_{k\pm}$ $(k=1,2,3)$:

$$m_{1\pm}=m_{x_1}\pm m_{x_4},\quad m_{2\pm}=m_{x_2}\pm m_{x_5},\quad m_{3\pm}=m_{x_3}\pm m_{x_6}.\qquad(14)$$

As a result, the system (11) is decomposed into four independent equations:

$$\left[\left(\omega+\frac{i}{T_2}\right)^2-\omega_{n1}^2\right]m_{1-}=0,$$

$$\left[\left(\omega+\frac{i}{T_2}\right)^2-\omega_{nk}^2\right]m_{k+}+2\omega_{nk}m_0\gamma_n\eta_{\perp}h_{\perp}=0,\quad k=1,2,3,\qquad(15)$$

and a system of two coupled equations:

$$\left[\left(\omega+\frac{i}{T_2}\right)^2-\omega_{n2}^2\right]m_{2-}+\frac{1}{2}\omega_{n2}\omega_{pc}(H)\frac{m_z}{m_0}(m_{2-}$$

$$-m_{3-})+2\omega_{n2}\gamma_n\eta_{\parallel}h_{\parallel}m_z=0,$$

$$\left[\left(\omega+\frac{i}{T_2}\right)^2-\omega_{n3}^2\right]m_{3-}+\frac{1}{2}\omega_{n3}\omega_{pc}(H)\frac{m_z}{m_0}(m_{3-}$$

$$-m_{2-})+2\omega_{n3}\gamma_n\eta_{\parallel}h_{\parallel}m_z=0,\qquad(16)$$

where the two quantities

$$\omega_{pc}(H)=\frac{\omega_{n2}(H_c)\epsilon}{[\epsilon+2(1-(H/H_c)^2)]}\qquad(17)$$

and

$$\eta_{\parallel}=H_n/\sqrt{H_c(H_c-H)},\qquad(18)$$

have singularities at $H=H_c$. According to Eqs. (15), the component in the NMR spectrum corresponding to m_{1-} is not excited by the variable field, and the m_{k+} components are excited by the transverse field $\mathbf{h}_{\perp}$. The spectrum of these components is represented by dots in Fig. 4, and their properties are discussed in Ref. 3. The new results, represented by the open circles in Fig. 4, are described by the solutions of Eqs. (16). Their form depends strongly on the ratio between the difference in the frequencies ω_{n2} and ω_{n3} (45),

$$\Delta=\omega_{n2}-\omega_{n3}=2H\sin(\alpha/2),\qquad(19)$$

and the quantity ω_{pc} (17). For $\Delta\gg\omega_{pc}$the frequencies of the components m_{2-} and m_{3-} differ from the frequencies of the components m_{2+} and m_{3+} excited by the field $\mathbf{h}_{\perp}$, consistent with the results shown in Fig. 4. For $\Delta\ll\omega_{pc}$ the following equations can be obtained for the frequencies of the components m_{2-} and m_{3-}, which are the roots of the secular equation for (16):

$$\Omega_2^2=0.5[\omega_{n2}^2(H)+\omega_{n3}^2(H)],\quad \Omega_3^2=\Omega_2[\Omega_2-\omega_{pc}(H)].\qquad(20)$$

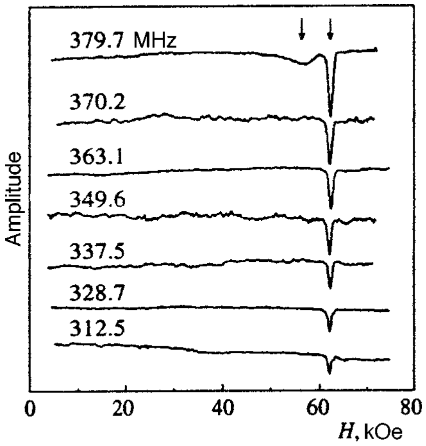

FIG. 6. Experimental traces of the low-frequency NMR branch $\Omega_3(H)$ at $T=1.3$ K and several frequencies. The arrows indicate the midpoints of the NMR lines at 379.7 MHz.

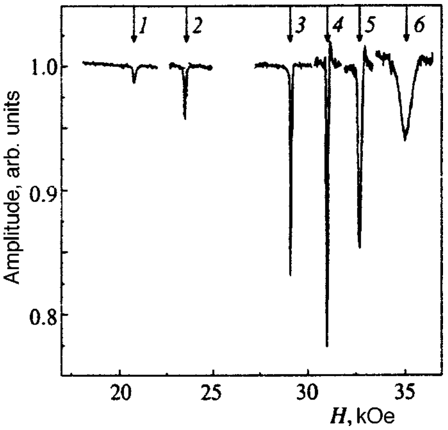

FIG. 7. Experimental low-frequency NMR branch $\Omega_4(H)$ at $T=1.3$ K and various frequencies: (*1*) 220.1 MHz; (*2*) 273.7 MHz; (*3*) 350.1 MHz; (*4*) 363.0 MHz; (*5*) 370.0 MHz; (*6*) 375.2 MHz.

As $H \rightarrow H_c$, the frequency Ω_2 approaches $\omega_{n2}(H_c) = \gamma_n A M_0$ (45), and Ω_3 (taking (17) into account) tends to zero:

$$\Omega_3^2(H) = \Omega_2^2(H_c)\frac{1-(H/H_c)^2}{\epsilon/2+1-(H/H_c)^2}. \tag{21}$$

The spectra calculated from these equations are represented by solid curves in Fig. 4.

Equations for the intensities $I_n(\omega)$ of the NMR lines at the frequencies Ω_2 and Ω_3 can be obtained by substituting the corresponding solutions of the system (16) into Eq. (12):

$$I_n(\Omega_3) = 2m_0 \eta_\parallel^2 \gamma_n h_\parallel T_2, \tag{22}$$

$$I_n(\Omega_2) = 4I_n(\Omega_3)\left[\frac{\gamma_n H}{\omega_{pc}(H)}\right]^2 \sin^2\frac{\alpha}{2}. \tag{23}$$

It follows from Eqs. (18), (21), and (22) that the intensity

$$I_n(\Omega_3) \propto \eta_\parallel^2 \propto \frac{1}{H_c-H} \propto \frac{1}{\Omega_3(H)} \tag{24}$$

should increase as $\Omega_3(H)$ decreases. It is evident from Fig. 6 that it decreases in the experiment. The reason for this disparity is that Eqs. (22) and (23) have been obtained from the solution of the system (16) in the approximation $m_z = m_0$. If this is not done, Eq. (20) has the form

$$\Omega_3^2 = \Omega_2[\Omega_2 - \omega_{pc}(H) m_z/m_0]. \tag{25}$$

Inasmuch as $m_0^2 = m_x^2 + m_y^2 + m_z^2$, the dependence of Ω_3 on m_z implies a dependence of Ω_3 on the amplitude of the oscillations of the magnetizations of the nuclear sublattices m_j. In other words, whereas the oscillations of m_j are in resonance with the alternating field at small amplitudes, the resonance conditions begin to break down as this amplitude increases, doing so more rapidly the higher the frequency $\omega_{pc}(H)$ and, accordingly, the lower the frequency Ω_3. Hence it follows that for sufficiently high $\omega_{pc}(H)$ the intensity $I_n(\Omega_3)$ begins to decrease as Ω_3 decreases.

The influence of various nonlinear effects on the steady-state NMR signals in the presence of a dynamic frequency shift has been analyzed previously.[11,13–18] If the equations given in the cited papers are used, the following relation between $I_n(\Omega_3)$ and Ω_3 can be obtained in the investigated frequency range of 310–380 MHz:

$$I_n(\Omega_3) \propto \Omega_3^{1/3}, \tag{26}$$

which agrees qualitatively with the results in Fig. 6. To make a quantitative comparison, the influence of inhomogeneities of the sample must also be taken into account in the theory, because we are now in the vicinity of the phase transition.

5. CONCLUSION

It follows from Fig. 6 that the decrease in the intensity of the steady-state NMR signals due to nonlinear effects strongly limits the frequency range in which such signals can be observed near H_c. A similar difficulty is encountered in weak fields $H \ll H_c$.[1,3] In this case the dynamic frequency shift is observed over a far broader range of fields H than merely in the vicinity of H_c, so that the NMR signals shown in Fig. 7 exhibit not only the values of the frequencies Ω, but also the fields H. It is evident from Fig. 7 that as H and Ω decrease (i.e., as the dynamic frequency shift increases), the signal intensity first increases and then decreases. If Eqs. (11), with nonlinear effects taken into account, are used to describe these signals, the following relation can be obtained for their intensities $I_n(\Omega)$ (12):

$$I_n(\Omega) \propto \eta_\perp^2 \propto H^{-2}, \tag{27}$$

where $\eta_\perp^2$ is given by Eq. (10). Like Eq. (24), this relation describes the increase in $I_n(\Omega)$ as Ω decreases. For larger dynamic frequency shifts, such that nonlinear effects become appreciable, it is necessary to use the same expressions for I_n as those from which the relation (26) is obtained. Then in place of (27) we can obtain a dependence of the form

$$I_n(\Omega_4) \propto H^{10/3}, \tag{28}$$

which dictates that I_n decreases as H decreases.

There are three ways to approach the investigation of the nuclear spin properties at lower frequencies (for larger dynamic frequency shifts). First, the amplitude of the exciting field can be raised to levels such that hysteresis effects begin to set in as a result of the nonlinearity of the dynamic frequency shift.[13–16] Tulin[17] investigated these effects experimentally for three-dimensional antiferromagnets ($MnCO_3$ and $CsMnF_3$), but he worked with weak fields H of the order of 1 kOe or less, for which the gain $\eta_\perp$ (10) is large. At $H \approx 20$ kOe, $\eta_\perp$ is much smaller, so that the investigation of hysteresis effects in the case of $CsMnBr_3$ requires more powerful rf field generators.

Second, NMR pulse techniques can be used, and echo signals in particular. Owing to the mechanism by which such signals are generated as a result of modulation of the NMR frequency, a certain reserve is available for increasing their amplitude in the presence of a large dynamic frequency shift. The feasibility of using NMR pulse signals to study the properties of nuclear spins with a large dynamic frequency shift has been discussed in detail.[18]

The third possibility for the investigation of NMR signals in the presence of a large dynamic frequency shift involves the parametric excitation of nuclear spins by parallel pumping.[19] This method is based on the fact that under the conditions of a dynamic frequency shift, the precession of the nuclear magnetic moments normally becomes elliptical, with an eccentricity that depends on the magnetic field. As a result, an alternating magnetic field $\mathbf{h}_\parallel$ at twice the NMR frequency imparts parametric instability to such precession if the amplitude $h(t)$ exceeds the threshold level h_0. Equations (11) then lead to an expression for h_0,

$$h_0(\omega) = \frac{2}{T_2 |\partial \omega_p / \partial H|}, \tag{29}$$

where ω_p is the dynamic frequency shift parameter. In $CsMnBr_3$ a large dynamic frequency shift occurs in two cases: a) for $H \approx H_c$. when $\omega_p(H_c) = \omega_{pc}$ (17); b) for $H \ll H_c$, when $\omega_p(H)$ is given by

$$\omega_p = \frac{12\epsilon\omega_{n1}}{9(H/H_c)^6 + 12\epsilon}. \tag{30}$$

Equation (29) is transformed as follows for these two cases:

a) for $H \approx H_c$ and $\epsilon \ll |1-(H/H_c)^2|$

$$h_0 = \frac{H_c|1-(H/H_c)^2|}{\omega_{pc}(H_c)T_2}; \tag{31}$$

b) for $H \ll H_c$ and $\epsilon \ll H/H_c$

$$h_0 = \frac{H}{3\omega_p(H)T_2}. \tag{32}$$

Equations (31) and (32) can also be used to analyze the threshold amplitude in the parametric pumping of nuclear spin waves.[20–22] It is sufficient here to replace T_2 and ω_p by $T_2(q)$ and $\omega_p(q)$, where q is the wave vector.

It is clear from these equations that the threshold amplitude h_0 drops as ω_p increases. This means that when the dynamic frequency shift increases, the conditions for the observation of parametric NMR improve, rather than deteriorating as in the case of steady-state NMR.

In closing, the authors wish to extend their sincere thanks to N. M. Kreĭnes, L. A. Prozorov, A. I. Smirnov, and I. A. Fomin for a productive discussion and critical remarks. This work received financial support from the Russian Fund for Fundamental Research (Projects 96-02-16489 and 98-02-16572).

APPENDIX

As mentioned in the Introduction, the equations used in Ref. 3 to analyze the spectrum of NMR frequencies can be applied to the calculation of the intensities of steady-state NMR signals. To do so, however, the equations must be modified to account for interactions with the external alternating magnetic field and with the fluctuating fields responsible for nuclear magnetic relaxation. If direct interaction of the magnetizations $\mathbf{m}_j$ of the nuclear sublattices with the alternating field is ignored (by virtue of amplitude effects[11]) and if interaction with the fluctuating fields is taken into account in the relaxation time approximation, it is possible to obtain a system of equations that coincides with the usual Bloch equations[23] in the external magnetic field $\mathbf{H}$ and the hyperfine field $\mathbf{H}_{nj} = A\mathbf{M}_j$ (Ref. 11):

$$\frac{dm_{x_j}}{dt} = \gamma_n(AM_{z_j} + H_{z_j})m_{y_j} - \frac{m_{x_j}}{T_2},$$
$$\frac{dm_{y_j}}{dt} = -\gamma_n(AM_{z_j} + H_{z_j})m_{x_j} + \gamma_n m_{z_j} AM_{x_j} - \frac{m_{y_j}}{T_2}, \tag{A1}$$

where (x_j, y_j, z_j) denotes the coordinate systems associated with equilibrium orientations of the vectors $\mathbf{m}_j \| \mathbf{H}_{nj} \| \mathbf{M}_j$ (we disregard the deviation of the orientation of the field $\mathbf{H}_{nj} + \mathbf{H}$ from the hyperfine field $\mathbf{H}_{nj}$, since $H_{nj} \gg H$), A is the hyperfine interaction constant, γ_n is the nuclear gyromagnetic ratio, and T_2 is the transverse nuclear magnetic relaxation time.

For small oscillations of m_j the components m_{zj} can almost always be replaced by the equilibrium value m_0. An exception is encountered for oscillations with a large dynamic frequency shift, when nonlinear effects become significant, and the variation of the components m_{z_j} must be taken into account. In $CsMnBr_3$, however, as opposed to two-sublattice, three-dimensional antiferromagnets, these effects do not exhibit any specific attributes, so that the equations derived in Refs. 13–18 can be used.

In the description of NMR experiments the magnetization components M_{z_j} of the electronic sublattices can always be replaced by the equilibrium values M_j (by virtue of the weak influence of hyperfine interaction on the oscillations of M_j).

The components M_{y_j} do not occur in Eqs. (A1). As in Ref. 3, they are disregarded, because strong uniaxial anisotropy prevents the sublattices from leaving the basal plane ($M_{y_j} \ll M_{x_j}$). The components M_{x_j} in Eq. (A1) can be written as a three-term sum

$$M_{x_j} = M_{x_j}^{(h_\perp)} + M_{x_j}^{(h_\parallel)} + M_{x_j}^{(m)}, \tag{A2}$$

where $M^{(h_\perp)}$ and $M^{(h_\parallel)}$ are attributable to interaction with the transverse ($\mathbf{h}_\perp$) and longitudinal ($\mathbf{h}_\parallel$) alternating magnetic fields, and the components $M_{x_j}^{(m)}$ are attributable to hyperfine interaction with the oscillations of $\mathbf{m}_j$. The equations used to calculate $M_{x_j}^{(m)}$ in Ref. 3 take into account only one low-frequency AFMR branch. Here we have made use of the fact that the NMR frequencies are very much lower than the AFMR frequencies (with allowance for the hyperfine gap in the spectrum of magnons[12]), so that the kinetic energy can be neglected in the expression for the Lagrangian describing the behavior of $\mathbf{M}_j$ (Ref. 24). In this approximation, equations for M_{x_j} can be obtained by minimizing the potential energy with hyperfine interaction taken into account. We can now obtain equations that are valid for any fields H, despite the difference in the symmetries of the soft modes for $H=0$ and $H=H_c$:

$$\begin{gathered} a_1 M_{x_1}^{(m)} - b_1 (M_{x_2}^{(m)} + M_{x_3}^{(m)}) = (H_n/H_{E'})(m_{x1}+m_{x4}), \\ -b_1 M_{x_1}^{(m)} + a_2 M_{x_2}^{(m)} + b_2 M_{x_3}^{(m)} = (H_n/H_{E'})(m_{x2}+m_{x6}), \\ -b_1 M_{x_1}^{(m)} + b_2 M_{x_2}^{(m)} + a_2 M_{x_3}^{(m)} = (H_n/H_{E'})(m_{x3}+m_{x6}), \end{gathered} \tag{A3}$$

$$M_{x_1}^{(m)} = M_{x_4}^{(m)}, \quad M_{x_2}^{(m)} = M_{x_6}^{(m)}, \quad M_{x_3}^{(m)} = M_{x_5}^{(m)}, \tag{A4}$$

where $H_n = AM_0$,

$$\begin{gathered} b_1 = \cos\frac{\alpha}{2} = \frac{1}{2-(H/H_c)^2}, \quad b_2 = \cos\alpha = 2b_1^2 - 1, \\ a_1 = \left(\frac{H}{H_c}\right)^2 + 2b_1 + \epsilon, \quad a_2 = b_1 - \left[1 - \left(\frac{H}{H_c}\right)^2\right] b_2 + \epsilon, \\ \epsilon = \frac{2Am_0}{H_{E'}}, \end{gathered} \tag{A5}$$

Am_0 is the static hyperfine field, which is conveyed by the nuclei to the electrons and is responsible for the hyperfine gap in the AFMR spectrum,[11], $H_{E'} = 3$ kOe is the effective in-plane exchange field, α is the angle between the vectors $\mathbf{M}_2$ and $\mathbf{M}_3$ (Fig. 1), and H_c is the critical field for transition to the collinear phase. At $T > 1$ K we have

$$\epsilon = \frac{2Am_0}{H_{E'}} = \frac{2H_n}{H_{E'}} \frac{\gamma_n}{\gamma_e} \frac{\hbar \gamma_n H_n}{kT} \leqslant 10^{-2},$$

so that only linear corrections in ϵ will be taken into account below. The solution of the system (A3), (A4) has the form

$$M_{x_j}^{(m)} = \frac{H_n}{H_{E'}} \sum_{j'} \lambda_{jj'} m_{x_{j'}}, \tag{A6}$$

where

$$\begin{gathered} \lambda_{11} = \lambda_{44} = (a_2^2 - b_2^2)/\mathscr{D}, \quad \lambda_{1j} = \lambda_{4j} = b_1(a_2 - b_2)/\mathscr{D}, \\ \lambda_{22} = \lambda_{33} = \lambda_{55} = \lambda_{66} = \lambda_{26} = \lambda_{35} = (a_1 a_2 - b_1^2)/\mathscr{D}, \\ \lambda_{23} = \lambda_{63} = \lambda_{25} = \lambda_{65} = (b_1^2 - a_1 b_2)/\mathscr{D}, \quad \lambda_{jj'} = \lambda_{j'j}; \end{gathered} \tag{A7}$$

and

$$\mathscr{D} = (a_2 - b_2)[a_1(a_2 + b_2) - 2b_1^2] \tag{A8}$$

is the determinant of the system (A3). For $H \leqslant H_c$, on the basis of (A5) and (A8), we have

$$\begin{aligned} \mathscr{D}(H) = [b_1(H)+1][1-(H/H_c)^2+\epsilon][3b_1(H)\epsilon \\ + (H/H_c)^6 b_1(H)(b_1(H)+1)^2], \end{aligned} \tag{A9}$$

and for $H \geqslant H_c$

$$\mathscr{D} = [(H/H_c)^2 - 1 + \epsilon](H/H_c)^2[(H/H_c)^2 + 3]. \tag{A10}$$

At the points $H=0$ and $H=H_c$ the determinant $\mathscr{D}(H)$ attains the minimum values

$$\mathscr{D}(0) = 9\epsilon/4, \quad \mathscr{D}(H_c) = 4\epsilon, \tag{A11}$$

and the components $M_{x_j}^{(m)}$ attain maxima.

Taking Eqs. (A2), (A6), (7), (9), and (10) into account, we can write Eq. (A1) in the form

$$\begin{aligned} &\left[\left(\omega + \frac{i}{T_2}\right)^2 - \omega_{nj}^2\right] m_{x_j}(\omega) \\ &\quad + \gamma_n \omega_{nj} A m_{z_j} \frac{H_n}{H_{E'}} \sum_{j'} \lambda_{jj'} m_{x_j}(\omega) \\ &\quad + \gamma_n \omega_{nj} m_{z_j} \eta_\perp h_\perp(\omega) + \gamma_n \omega_{nj} m_{z_j} \eta_{\|j} h_\|(\omega) = 0, \end{aligned} \tag{A12}$$

where

$$\begin{gathered} \omega_{n1} = \omega_{n4} = \gamma_n A M_1, \\ \omega_{n2} = \omega_{n5} = \gamma_n A M_2 + H\sin(\alpha/2), \\ \omega_{n3} = \omega_{n6} = \gamma_n A M_2 - H\sin(\alpha/2), \end{gathered} \tag{A13}$$

are the unshifted (not perturbed by dynamic frequency shift) NMR frequencies, while

$$\eta_\perp = A\chi_\perp = H_n/H \tag{A14}$$

and

$$\eta_{\|2} = \eta_{\|6} = -\eta_{\|3} = -\eta_{\|5} = \eta_\| = A\chi_\|, \quad \eta_{\|1} = \eta_{\|4} = 0 \tag{A15}$$

are the gains for the field $h_\perp$ and $h_\|$; the variables $m_{x_j}(\omega)$, $h_\perp(\omega)$, and $h_\|(\omega)$ are related to $m_{x_j}(t)$, $h_\perp(t)$, and $h_\|(t)$ by the Fourier transform

$$m_{x_j}(\omega) = \int_{-\infty}^{+\infty} \exp(i\omega t)\, m_{x_j}(t)\, dt. \tag{A16}$$

Analogous relations hold for $h_\perp(\omega)$ and $h_\|(\omega)$.

*)E-mail: dumesh@isan.troitsk.ru

---

[1] A. S. Borovik-Romanov, S. V. Petrov, A. M. Tikhonov, and B. S. Dumesh, JETP Lett. **66**, 759 (1997).
[2] B. S. Dumesh, S. V. Petrov, and A. M. Tikhonov, JETP Lett. **67**, 1046 (1998).
[3] A. S. Borovik-Romanov, B. S. Dumesh, S. V. Petrov, and A. M. Tikhonov, Zh. Éksp. Teor. Fiz. **113**, 352 (1998) [JETP **86**, 197 (1998)].
[4] J. Goodyear and D. J. Kennedy, Acta Crystallogr. Sect. B **28**, 1640 (1974).
[5] M. Eibischutz, R. C. Sherwood, F. S. L. Hsu, and D. E. Cox, in *Proceedings of the 18th Annual Conference on Magnetism and Magnetic Materials*, AIP Conf. Proc., No. 10, American Institute of Physics, New York (1973), p. 684.

[6] B. D. Gaulin, T. E. Mason, M. F. Collins, and J. Z. Larese, Phys. Rev. Lett. **62**, 1380 (1989).
[7] B. D. Gaulin, M. F. Collins, and W. J. L. Buyers, J. Appl. Phys. **61**, 3409 (1987).
[8] I. A. Zaliznyak, L. A. Prozorova, and S. V. Petrov, Zh. Éksp. Teor. Fiz. **97**, 359 (1989) [Sov. Phys. JETP **70**, 203 (1989)].
[9] S. I. Abarzhi, A. N. Bazhan, L. A. Prozorova, and I. A. Zaliznyak, J. Phys.: Condens. Matter **4**, 3307 (1992).
[10] A. V. Chubukov, J. Phys. C Solid State Phys. **21**, 441 (1988).
[11] M. I. Kurkin and E. A. Turov, *Nuclear Magnetic Resonance in Magnetically Ordered Materials and Its Applications* [in Russian], Nauka, Moscow (1990).
[12] I. A. Zaliznyak, N. N. Zorin, and S. V. Petrov, JETP Lett. **64**, 473 (1996).
[13] P. G. De Gennes, P. Pincus, F. Hartmann-Bourtron, and J. M. Winter, Phys. Rev. **129**, 1105 (1963).
[14] M. I. Kurkin, JETP Lett. **28**, 628 (1978).
[15] M. I. Kurkin, Yu. G. Raĭdugin, V. N. Sedyshkin, and F. P. Tankeev, Fiz. Tverd. Tela (Leningrad) **32**, 1577 (1990) [Sov. Phys. Solid State **32**, 923 (1990)].
[16] M. I. Kurkin, Fiz. Tverd. Tela (Leningrad) **33**, 1805 (1991) [Sov. Phys. Solid State **33**, 1014 (1991)].
[17] V. A. Tulin, Zh. Éksp. Teor. Fiz. **55**, 831 (1968) [Sov. Phys. JETP **28**, 431 (1968)]; Zh. Éksp. Teor. Fiz. **78**, 149 (1980) [Sov. Phys. JETP **51**, 73 (1980)].
[18] A. S. Borovik-Romanov, Yu. M. Bun'kov, B. S. Dumesh *et al.*, Usp. Fiz. Nauk **142**, 537 (1984) [Sov. Phys. Usp. **27**, 235 (1984)].
[19] A. G. Gurevich and G. A. Melkov, *Magnetic Oscillationsand Waves* [in Russian], Nauka, Moscow (1994), p. 285.
[20] P. M. Richards and L. M. Hinderks, Phys. Rev. **183**, 575 (1969).
[21] V. T. Adams, L. W. Hinderks, and P. M. Richards, J. Appl. Phys. **41**, 931 (1970).
[22] V. I. Ozhogin and A. Yu. Yakubovsky, Phys. Lett. **43**, 505 (1973).
[23] F. Bloch, Phys. Rev. **40**, 460 (1946).
[24] A. F. Andreev and V. I. Marchenko, Usp. Fiz. Nauk **130**,39 (1980) [Sov. Phys. Usp. **23**, 21 (1980)].

Translated by James S. Wood

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Nonlinear microwave response of YBCO single crystal in constant magnetic field

V. V. Bol'ginov, V. M. Genkin, G. I. Leviev,*) and L. V. Ovchinnikova

Institute of Solid State Physics, Russian Academy of Sciences, 142432 Chernogolovka, Moscow Region, Russia

(Submitted 20 January 1999)
Zh. Éksp. Teor. Fiz. **115**, 2242–2253 (June 1999)

The generation of a microwave second harmonic by a YBCO single crystal in a dc magnetic field was studied. We found that the signal existed only when there was a direct screening current. As a result, the pinning current as a function of magnetic field can be derived directly from the second harmonic intensity versus the magnetic field. The experimental data were interpreted in terms of a generalized model of the critical state taking into account diffusion of vortices and the absence of a barrier stopping vortices from leaving the sample. We have shown that, in a decreasing dc magnetic field, the current density is considerably lower than both the critical and screening current densities in an increasing dc field. Our experiments indicate that vortices are not the sources of radiation at the double frequency. A relation between the mechanism of harmonic generation in the Meissner phase and modulation of the order parameter by the microwave magnetic field (Ginzburg–Landau nonlinearity) is discussed. It is remarkable that, by measuring the second harmonic intensity in the Meissner state versus temperature, one can obtain the magnetic field penetration depth as a function of temperature with fairly good accuracy. [S1063-7761(99)02606-2]

1. INTRODUCTION

The behavior of superconductors in strong high-frequency fields has been extensively studied in recent years. The incentive for this research is not only the interest in the fundamental aspects of the nonlinear electrodynamics of superconductors, but also the promise shown by these materials in the field of development of microwave devices, and the fact that the nonlinearity of their microwave response determines the characteristics of prospective devices. The interaction between high-temperature superconductors and microwave fields has been studied by many researchers. Gor'kov and Éliashberg[1] showed that harmonic generation in superconductors in the framework of the nonstationary superconductivity theory is determined by the order-parameter relaxation time, so this time can be derived from measurements of the harmonic intensity. Such an experiment was performed by Amato and McLean[2] on LaCaSn samples. The relaxation time was calculated by means of equations taking into account the actual configuration[3] of the experiment at a radiation frequency of 11 GHz. Later this technique was used to measure the order-parameter relaxation time in a YBCO superconducting single crystal.[4,5] At a lower frequency harmonic generation in a YBCO single crystal was studied by Ciccarello *et al.*[6,7] Their results were interpreted using the two-liquid model with the densities of superconducting and normal electrons characterized by special functions of the electromagnetic wave amplitude. In order to calculate a harmonic intensity, the requisite Fourier component of the nonlinear crystal magnetization was identified. By combining results of nonlinear experiments and measurements of single-crystal surface impedance under magnetic field, the authors determined the temperature behavior of the upper and lower critical fields and their numerical values.[7,8] In view of possible applications, experiments with films are most important. The main mechanism of nonlinearity in film samples is associated with Josephson junctions between structural irregularities.[9–14] Harmonic generation in structures with weak links was the subject of publications by Jeffries *et al.*[15,16] It is noteworthy that such ideas have also been developed in the context of low-temperature superconductors.[17]

The investigation of the high-frequency nonlinear response, in particular, generation of even harmonics, in a magnetic field is of special interest.[18–26] In the absence of a dc magnetic field, this generation is forbidden by the symmetry laws, so the harmonic intensity is determined by the field configuration (of the screening current) in the surface layer and contains information concerning this configuration. The model of the critical state is most often used in describing nonlinear effects in a magnetic field. The mechanism of nonlinear absorption or emission at frequencies of higher harmonics is associated with vortices generated in a sample in the mixed state. The nonlinearity of the electromagnetic response of the vortex system is determined by several factors, such as the pinning potential nonparabolicity,[27] its dependence on the current,[28] and others.[29] In our experiments, no nonlinearity associated with vortices has been detected. This conclusion derives from the curves of harmonic intensity versus temperature and magnetic field. This paper does not deal with details of the radiation mechanism in the mixed state. As for the Meissner phase, we will discuss an interpretation of the harmonic radiation in terms of the Ginzburg–Landau nonlinearity, namely, modulation of the order param-

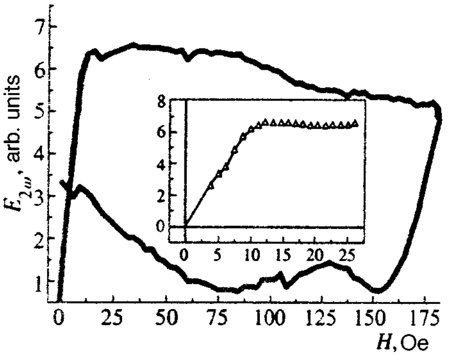

FIG. 1. Second harmonic amplitude $E_{2\omega}$ as a function of dc magnetic field H in the range of 0 to 180 Oe. The inset shows the range of 0 to 25 Oe. The temperature $T=78$ K, the field scan starts at 0 Oe.

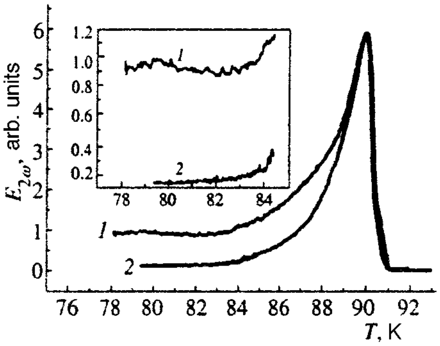

FIG. 3. Second harmonic amplitude $E_{2\omega}$ versus temperature T. Curve *1* was recorded by warming the sample to the normal state in a magnetic field of 250 Oe introduced when the sample was in the superconducting state at the lowest temperature; curve *2* was recorded by cooling the sample in the same field (250 Oe). The inset shows fragments of curves *1* and *2* on the extended scale.

eter by the microwave magnetic field. Irrespective of the mechanism, the intensity of the double-frequency radiation is controlled by currents near the surface, which screen the dc magnetic field, and the current can be derived from the second harmonic intensity. Below, our experimental results on generation of the second harmonic in a YBCO single crystal will be interpreted after a description of the experimental procedure.

2. EXPERIMENT

A YBCO single crystal shaped as a plate with dimensions of $4\times3\times0.2$ mm^3 was fabricated by G. A. Emel'chenko in the Institute of Solid State Physics, Russian Academy of Sciences. The superconducting transition occurs, according to impedance measurements at a frequency of 3 GHz (Ref. 30), at a temperature of 91 K and has a width of about 1 K. The sample was placed on the bottom of a bimodal cylindrical cavity resonating in the TM_{010} mode at the incident wave frequency $\omega/2\pi=9.2$ GHz and in the TE_{111} mode at the second harmonic frequency. The sample was placed at the site where the magnetic fields of both these modes were aligned with one another over the sample area. The dc magnetic field was also aligned with the microwave fields. The source of incident radiation was a magnetron operating in the pulsed mode with a pulse duration of 1 μs. The relative pulse duration was chosen so that the results of measurements were independent of the repetition rate. The intensity of the emitted second harmonic was recorded as a function of the dc magnetic field, temperature, and amplitude of the incident radiation.

The second harmonic amplitude as a function of dc magnetic field is plotted in Figs. 1 and 2. The sample was cooled down to the required temperature at zero magnetic field, then the field was slowly scanned to its maximum value and the scanning direction was reversed. The field scan rate was selected so that the shape of the hysteresis loop remained the same when the rate was reduced severalfold. The curves were recorded at several temperatures, and the field H_x at which the harmonic amplitude saturated proved to be almost independent of temperature.

Figure 3 shows the harmonic intensity versus temperature. The magnetic field of 250 Oe was applied at a temperature of 78 K, then the sample was heated above T_c and again cooled without changing the magnetic field. One can see that the harmonic intensity after the heating-and-cooling cycle is several times lower. The second harmonic intensity in the Meissner phase (at a magnetic field of 5.3 Oe) versus temperature is plotted in Fig. 4.

The harmonic amplitude can be treated as a quadratic function of the incident wave amplitude throughout the interval of studied microwave intensities (Fig. 5).

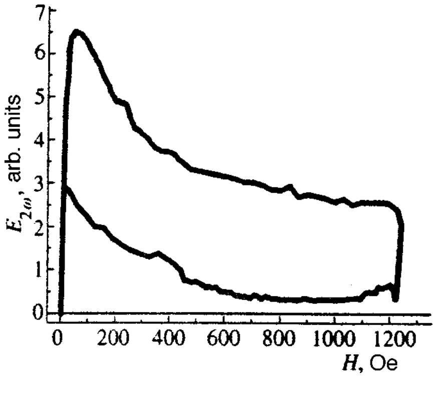

FIG. 2. Second harmonic amplitude $E_{2\omega}$ versus dc magnetic field H in the range of 0 to 1200 Oe. The temperature is $T=78$ K, the field scan starts at 0 Oe.

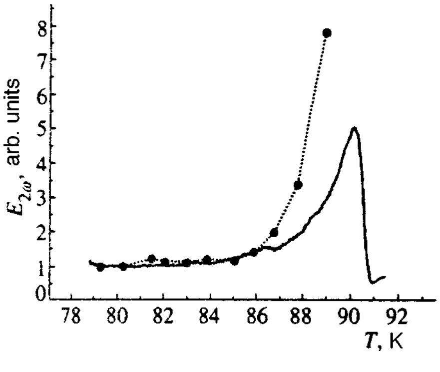

FIG. 4. Second harmonic amplitude versus temperature in the Meissner state (under a field of 5.3 Oe). Filled symbols connected by the dashed line follow the λ^5 law.

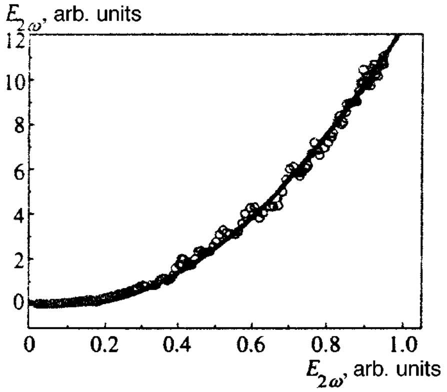

FIG. 5. Second harmonic amplitude $E_{2\omega}$ versus the incident wave amplitude E_ω. The solid line follows the quadratic function. The temperature is $T=78$ K.

3. DISCUSSION

Harmonic amplitude. A second harmonic cannot be generated in the absence of a magnetic field in a crystal with an inversion center, which directly follows from symmetry considerations.[31] As the magnetic field increases, a harmonic can already be detected in the Meissner state, when the field is surely lower than the field of vortex penetration into the sample. In the presence of a magnetic field, a constant screening current is generated near the surface and, in accordance with our data, we will attribute the second harmonic radiation to this current. In accordance with the phenomenological approach, let us write an expression for the microwave current j_{mw} carried by superconducting electrons in the presence of the screening current j_{dc} and microwave fields at frequencies ω and 2ω characterized by the vector potential components $A_\omega e^{-i\omega t}$ and $A_{2\omega}e^{-i2\omega t}$:

$$j_{\mathrm{mw}}=-\frac{c}{4\pi\lambda^2}A_\omega e^{-i\omega t}-\frac{c}{4\pi\lambda^2}A_{2\omega}e^{-i2\omega t}+\chi(j)A_\omega e^{-i\omega t}A_\omega e^{-i\omega t}. \tag{1}$$

Here $\chi(j)$ is the phenomenological nonlinearity factor, which vanishes together with the direct screening current j. The linear component of the microwave current is written in the conventional London form, λ is the penetration depth of magnetic field, and c is the speed of light. These terms are also nonvanishing in the absence of a dc magnetic field, and in analyzing the harmonic generation one can assume that they are independent of the field and limit the analysis to the function determined by the nonlinear term. In the experiment, one can measure the power fed to the microwave cavity, its Q-factor, coupling factors for both resonant modes, geometrical sizes of the sample and cavity. Using these parameters and the constitutive equation (1), one can determine the second harmonic amplitude $E_{2\omega}$ using the theory describing a cavity excited by a prescribed current.[32] The calculation details can be found in Ref. 33. This technique used in calculating the radiation intensity, unlike those used by other authors,[6,7] yields a correct asymptotic result in the limit of infinite sample thickness (in reality, for thicknesses much larger than the field penetration depth λ). The wave amplitude at the second harmonic frequency is determined by the integrated interaction between a predetermined nonlinear source

$$j^{\mathrm{nl}}(2\omega)=\chi(j)A_\omega e^{-i\omega t}A_\omega e^{-i\omega t}$$

and the resonant mode field, to which the radiation is coupled:

$$E_{2\omega}\propto Q_{111}\int j^{\mathrm{nl}}(2\omega)e_{111}(\mathbf{r})d^3r. \tag{2}$$

Integration in Eq. (2) is performed over the sample volume, $e_{111}(\mathbf{r})$ is the electric field of the TE_{111} mode inside the crystal, Q_{111} is the Q-factor of this mode. The integral (2) strongly depends on temperature because the integrand includes the field within the sample. The magnetic field dependence is incorporated only in the form of the nonlinearity factor $\chi(j)$, i.e., it is determined in the long run by the screening current in the skin. It is more convenient to assume that χ is a function of current (rather than the field H) because there is a hysteresis effect (Figs. 1 and 2). The harmonic amplitude versus the screening current is assumed to be a single-valued function, without hysteresis. It follows from our experimental data (Fig. 5) that the radiation intensity at the double frequency is a quadratic function of the incident wave amplitude throughout the range of studied microwave intensities (at temperatures notably different from T_c). Therefore the nonlinear source in Eq. (1) is a quadratic function of the vector potential at the fundamental frequency. A phenomenological relation between the nonlinear source and vector potential does not specify a mechanism of nonlinearity. We will reconsider the issue of the harmonic generation mechanism in discussing the temperature dependence of the generation intensity, but first let us discuss the intensity versus magnetic field (curves with hysteresis loops in Figs. 1 and 2), which requires a detailed analysis of the surface current in increasing and decreasing magnetic fields. An increase in the signal and switching to a slower function of the field can be easily described in terms of the Bean model modified by Portis for analysis of the microwave absorption.[34,35] But according to this model, the harmonic signal should drop to zero when the magnetic field scan is reversed and then rapidly recover to the level corresponding to the field scanned in the upward direction. This is not so in the experiment: the signal remains small over a wide range of magnetic fields. Later we will discuss a model that will allow us to interpret this behavior of the harmonic signal in terms of the screening current as a function of magnetic field, generation of vortices on the surface, their diffusion inside the sample, and their exit from the sample without any barrier.

Direct screening current and hysteresis loop on the curve of the second harmonic intensity. The magnetization, alongside the screening current in a superconductor in an equilibrium state, is controlled by the applied magnetic field and temperature. The Meissner state is stable in a plate of a type-II superconductor under a magnetic field parallel to its surface if $H\leqslant H_{c1}$, where H_{c1} is the first critical field. If the field exceeds H_{c1}, the Meissner state becomes metastable, and vortices can penetrate into the sample. In a magnetic field $H = H_{c1}$ the surface current density j_s is approximately $j^*\xi/\lambda$, where j^* is the Ginzburg–Landau decoupling current density and ξ is the coherence length. The Meissner state becomes absolutely unstable in a magnetic field $H = H_s$, where the surface current density satisfies $j_s \simeq j^*$. Here H_s is the field of the Bean–Livingston surface barrier.[36] Thus, one can say that on the microscopic level the process of vortex penetration into a superconductor is controlled by the current density. It is usually supposed that the main parameter is the magnetic field strength.[37,38]

In a real experiment, a vortex can enter a superconductor at a field weaker than H_s. This process can take place because of a surface irregularity where the local current density achieves j^*. From this viewpoint, vortices can penetrate into a sample when the current density on the surface is higher than a certain critical value $\tilde{j}$. This parameter is determined by the real surface morphology and satisfies the condition $j^*\xi/\lambda \leqslant \tilde{j} \leqslant j^*$. In real samples, pinning should be taken into account, which was done in the theory of the critical state.[39] The magnetization is assumed to be equilibrium, and surface currents are not taken into account. Nevertheless, they play an important part in the magnetic moment and microwave response.

The pinning critical current density $j_{\rm p}$ is controlled by the interaction between vortices and static defects in the bulk. We can identify two possibilities: $j_{\rm p} > \tilde{j}$ and $j_{\rm p} < \tilde{j}$. In the former case the strong pinning masks the surface barrier, and one can say that vortices penetrate into the sample only if $j_s > j_c$. In reality this is the case treated by the critical state theory. The latter possibility is more interesting.

The vortex density near the surface can be small because of surface currents, which effectively drive vortices into the sample. A similar inhomogeneous vortex distribution was considered in connection with the "geometrical" barrier for a plate in a normal magnetic field.[40] For a thermodynamically equilibrium state, the vortex density ρ near the surface is different from that in the bulk. In the sample interior we have $\rho < H/\phi_0$, where ϕ_0 is the magnetic flux quantum, because the superconductor magnetic susceptibility is $\mu < 1$, and the magnetic induction jumps on the superconductor surface. The resulting surface current drives vortices from the surface.

Let us discuss the distributions of current and vortex densities in a plate in a parallel magnetic field. Suppose that the vortex penetration into the sample is controlled by the current density. We will prove that the difference between surface currents in increasing and decreasing magnetic fields is due to redistribution of vortices in the sample without a change in their total number. Consider a superconducting plate of thickness $2d$ in an external parallel magnetic field H and introduce a reference frame with the x-axis normal to the plate and z-axis aligned with the applied magnetic field. Let us define the $x = 0$ plane in the middle of the plate. In the linear approximation, we have an equation for the vector potential:

$$\Delta\mathbf{A} = \left[\mathbf{A} - \phi_0 \sum_i \nabla\phi(\mathbf{r}-\mathbf{r}_i)/2\pi\right] \Big/ \lambda^2, \tag{3}$$

where $\nabla\phi(\mathbf{r}-\mathbf{r}_i)$ is the phase gradient generated by the ith vortex at point $\mathbf{r}(x,y)$. Equation (3) is linear, so one can average it over the plane $x = \text{const}$ and introduce the average density of vortices $\rho(x)$:

$$\partial^2 A/\partial x^2 = (A - \phi_0 N(x))/\lambda^2, \tag{4}$$

where

$$N(x) = \int_0^x \rho(x)\,dx.$$

Equation (4) does not imply averaging along the x-axis. The solution of this equation with the boundary conditions $\partial A/\partial x = H$ at $x = d$ and $-d$ has the form

$$\begin{aligned} A = {} & \left[H\lambda + \phi_0 \exp\left(-\frac{d}{\lambda}\right)\right. \\ & \left.\times \int_0^d \sinh\left(\frac{x}{\lambda}\right)\frac{N(x)}{\lambda}\,dx\right]\frac{\sinh(x/\lambda)}{\cosh(d/\lambda)} \\ & + \phi_0 \int_0^d \left[\exp\left(-\frac{|x-x'|}{\lambda}\right)\right. \\ & \left. - \exp\left(-\frac{|x+x'|}{\lambda}\right)\right]\frac{N(x')}{2\lambda}\,dx'. \end{aligned} \tag{5}$$

If $N(x)$ changes slowly over the London length and $2d \gg \lambda$ $(x > 0)$, the vector potential A and superconducting current density j can be expressed by approximate formulas:

$$A = \phi_0 N(x) + \phi_0 \lambda^2 \frac{\partial\rho}{\partial x} + \lambda[H - \phi_0\rho(d)]\exp\left(-\frac{d-x}{\lambda}\right), \tag{6}$$

$$j = -\frac{c}{4\pi}\left[\phi_0\frac{\partial\rho}{\partial x} + \left(\frac{H}{\lambda} - \phi_0\frac{\rho(d)}{\lambda}\right)\exp\left(-\frac{d-x}{\lambda}\right)\right]. \tag{7}$$

The first term on the right of Eq. (7) yields the current density, which is used in the theory of critical state if $\phi_0\rho$ is treated as the magnetic induction. Given the current density, one can calculate the sample magnetic moment:

$$\begin{aligned} M = {} & \phi_0\left(1 + \tanh\frac{d}{\lambda}\right)\exp\left(-\frac{d}{\lambda}\right)\int_0^d \frac{N(x)}{2\pi\lambda} \\ & \times \sinh\left(\frac{x}{\lambda}\right)dx - \frac{H}{2\pi}\left(d - \lambda\tanh\frac{d}{\lambda}\right). \end{aligned} \tag{8}$$

Direct calculation of the functional derivative $-\partial(MH)/\partial\rho(x)$ indicates that this parameter vanishes at $x = d$. This means that there is nothing to stop vortices from leaving the sample, so elimination of a vortex decreases the free energy. Thus, if there is no surface current acting as a wall restraining the vortices they will be ejected from the sample. As a result, the vortex density on the surface will vanish in a decreasing magnetic field, when the Lorentz force expels them from the sample. This differs from the theory of the critical state, where the surface density of vortices can take arbitrary values.

For $N(x)$ we have a continuity condition:

$$\partial N/\partial t + V\partial N/\partial x + D\partial^2 N/\partial x^2 = Q(j), \tag{9}$$

where V is the average vortex density, D is the diffusion coefficient, and $Q(j)$ is responsible for generation of vortices by superconducting currents in the sample. The term containing the second derivative on the left-hand side of Eq. (9) accounts for the vortex diffusion. We describe $Q(j)$ in terms of a model in which vortices are generated when the current density is higher than a certain critical value $\tilde{j}$ (Ref. 41):

$$Q(j)=-\gamma\int_0^x(|j|-\tilde{j})\Theta(|j|-\tilde{j})\,\mathrm{sign}(j)\,dx, \tag{10}$$

where γ is a phenomenological parameter and $\Theta(x)$ is the Heaviside's step-function.

In order to obtain a closed equation system, we need a relation between the vortex velocity V and current density j. We assume a relation between V and j in the spirit of the Bean model:

$$V=0,\quad j<j_p;\quad V=\mathrm{const}(j-j_p),\quad j>j_p, \tag{11}$$

where j_{p} is the pinning current density.

There is a solution of Eqs. (4) and (9) with zero current and a constant vortex density $\rho=H/\phi_0$ throughout the sample. This is stationary but not an equilibrium state. For a thermodynamically equilibrium state with $\rho_0\phi_0<H$, the vortex density on the surface is zero owing to the surface currents driving vortices into the sample interior. Here ρ_0 is the equilibrium vortex density in the volume. The formation of a vortex-free zone in a nonequilibrium state was discussed by Clem[38] and Burlachkov.[42]

There are discontinuities in the current and vortex densities because all diffusion processes have been neglected. One can write an approximate expression for the current density on the surface for an equilibrium state in an applied magnetic field H:

$$j_s=cH(3.06976-2.78577\mu-0.428182\mu^{-1}+0.216499\mu^{-2})/4\pi\lambda \tag{12}$$

for $0.35<\mu<0.95$ and $\lambda/d<0.05$.

In this model, the difference between the intensities of the second harmonic in increasing and decreasing magnetic fields can be interpreted in the following manner. Experimental data indicate that the harmonic generation is controlled by the direct surface current. In an increasing magnetic field the surface current density is $\tilde{j}$ because vortices do not penetrate into the sample at smaller currents. This current drives vortices into the sample and forms a vortex-free zone near the surface. In a decreasing field, the surface current changes its sign, ejects vortices from the sample, and forms a vortex-free zone near the surface. There is a jump in the vortex density in both these cases. The currents have opposite signs, but almost equal magnitudes. Vortex diffusion eliminates jumps and causes the surface current density to vanish for a decreasing magnetic fields because only in this case do the diffusion processes and Lorentz force act in one direction. In an increasing magnetic field, diffusion drives vortices out of the sample, i.e., it acts against the Lorentz force. Figure 6 shows current distributions in the skin in increasing and decreasing fields with and without inclusion of diffusion. The difference between the currents in increasing and decreasing fields leads to the difference in the second harmonic intensity. The theoretical curve in Fig. 7 was calculated in the following manner. The second harmonic amplitude, in accordance with Eq. (2), is proportional to

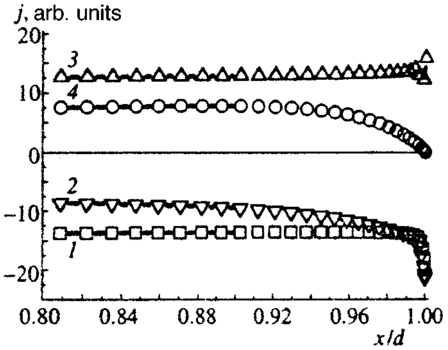

FIG. 6. Distribution of screening current calculated by solving model equations. Curves *1* and *2* are calculated for the increasing magnetic field with and without taking account of diffusion, respectively. Curves *3* and *4* show calculations for the decreasing field with and without taking account of diffusion.

$$P=\int_d^0 j(x)\exp\bigl[3(x-d)/\lambda\bigr]dx, \tag{13}$$

where the penetration depth of the high-frequency field is assumed to be constantas a function of the frequency, and the nonlinear susceptibility is proportional to the direct current in the sample. First P was calculated without taking diffusion into account, $D=0$, but the pinning current was treated as a function of magnetic field. We assume that measurements of the second harmonic amplitude versus an increasing magnetic field faithfully reproduce the pinning current versus magnetic field, $j_{\mathrm{p}}(H)$. After substituting this function in Eq. (13) we have good agreement between the theory and experiment for increasing magnetic fields, but a large discrepancy for decreasing fields. In the next stage the vortex distribution in a certain magnetic field was treated as a rough approximation, and a more accurate distribution function was calculated for $D\neq 0$. It is clear that the diffusion has little effect on the second harmonic amplitude, whereas its effect is quite

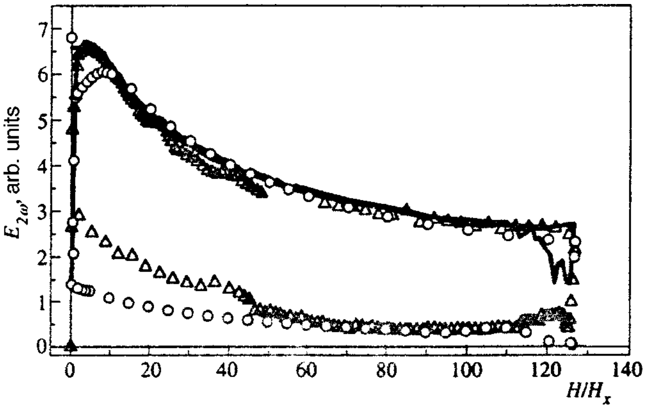

FIG. 7. Comparison between measurements of the harmonic amplitude versus magnetic field and curves with hysteresis loops calculated by model equations. Triangles plot experimental data, the solid line plots calculations without diffusion, open circles calculations taking diffusion into account.

considerable in a decreasing magnetic field, and one can see that the agreement between calculations and experimental data is better if diffusion is taken into account. An estimate of the diffusion coefficient yields $D \leqslant 10^{-12}$–10^{-14} cm^2/s.

Harmonic intensity versus temperature. The second harmonic intensity plotted against the temperature in Fig. 3 allows us to draw an important conclusion: vortices are not sources of the second harmonic radiation. Specifically, Fig. 3 allows one to compare between radiation intensities at the same temperature and magnetic field magnitude, but with different magnetic prehistories of the sample, namely, with the magnetic field introduced in the normal and superconducting states. One can see that the harmonic intensity is notably higher in the latter case, although the vortex density generated near the surface is lower in this case. (A similar conclusion about the insignificant contribution of vortices to radiation derives from the analysis of the harmonic amplitude versus magnetic field.) The weak radiation intensity generated in the sample cooled in a magnetic field indicates that the surface current is low. Although it stretches the point a little bit, we can say that this current is suppressed because the Meissner current is compensated for by the vortex current (strictly speaking, one cannot separate the Meissner and vortex currents).

Let us discuss the harmonic intensity in a weak magnetic field as a function of temperature, when there is no penetration of vortices over a wide temperature range (Fig. 4). In the microwave frequency band one can use the relation between the current and vector potential obtained in the Ginzburg–Landau theory for a static homogenous configuration:

$$j_j = -\frac{c}{4\pi\lambda^2}A\left(1-\frac{c^2}{108\pi^2\lambda^4 j^{*2}}A^2\right). \tag{14}$$

Let us express the vector potential in the form of two terms: one at the frequency of the incident radiation ω and the other at zero frequency: $A = A_0 + A_\omega$. We obtain a nonlinear source j^{nl} in the expression for the harmonic intensity (2):

$$j^{\mathrm{nl}} \propto K A_0 A_\omega^2 .$$

The proportionality factor K in this expression is independent of temperature. Given that $A_\omega \propto \zeta \propto \lambda$, where ζ is the sample surface impedance in the microwave band, and $A_0 \propto \lambda$, we obtain the harmonic amplitude

$$E_{2\omega} \propto \lambda^5 . \tag{15}$$

Figure 4 shows that the measurements of the harmonic intensity and calculations by Eq. (15) are in fairly good agreement at temperatures quite different from T_c. The penetration depth λ versus temperature was taken from Ref. 43, where it was derived from measurements of the surface impedance of a YBCO single crystal. Around T_c Eq. (15) fails, at least because the generation regime is not quadratic. These considerations lead us to a conclusion that the mechanism of harmonic generation in the Meissner phase is due to modulation of the order parameter by microwave magnetic field, i.e., Ginzburg–Landau nonlinearity. If this is the case, Eq. (15) can be used in calculating the penetration depth versus temperature.

4. CONCLUSION

Let us summarize the basic results of our work. We have studied experimentally and theoretically generation of the second harmonic in a YBCO superconducting single crystal under an intense microwave field. The second harmonic is generated only in a dc magnetic field. The curve of the harmonic intensity versus applied magnetic field has a hysteresis loop, which can be interpreted in terms of a model taking into account diffusion of vortices, although vortices themselves do not act as sources of the second harmonic. There are good reasons to assume that the harmonic generation mechanism for the Meissner phase is due to the Ginzburg–Landau nonlinearity.

We are grateful to E. S. Borovitskaya, M. V. Feigel'man, and D. V. Shovkun for detailed discussions of results of this study, and to M. V. Golubkov for help in conducting the experiment. The work was supported by the Russian Fund for Fundamental Research (Project No. 97-02-17981).

[*)] E-mail: leviev@issp.ac.ru

---

[1] L. P. Gor'kov and G. M. Éliashberg, Zh. Éksp. Teor. Fiz. **54**, 612 (1968) [Sov. Phys. JETP **27**, 328 (1968)].
[2] J. S. Amato and W. L. McLean, Phys. Rev. Lett. **37**, 930 (1976).
[3] W. L. McLean and J. S. Amato, Phys. Lett. A **45**, 289 (1973).
[4] G. I. Leviev, A. V. Rylyakov, and M. R. Trunin, JETP Lett. **50**, 88 (1989).
[5] M. Trunin and G. Leviev, J. Phys. III **2**, 355 (1992).
[6] I. Ciccarello, M. Guccione, M. Li Vigni *et al.*, Phys. Rev. B **49**, 6280 (1994).
[7] A. Gallitto and M. Li Vigni, Physica C **305**, 75 (1998).
[8] A. Gallitto, I. Ciccarello, M. Guccione *et al.*, Phys. Rev. B **56**, 5140 (1997).
[9] G. Hampel, B. Batlog, K. Krishana *et al.*, Appl. Phys. Lett. **71**, 3904 (1997).
[10] P. P. Nguen, D. E. Oates, G. Dresselhaus *et al.*, Phys. Rev. B **51**, 6686 (1995).
[11] J. McDonald, J. R. Clem, and D. E. Oates, Phys. Rev. B **55**, 11823 (1997).
[12] Y. M. Habib, C. J. Lehner, D. E. Oates *et al.*, Phys. Rev. B **57**, 13833 (1998).
[13] A. Philipp, Solid State Commun. **86**, 719 (1993).
[14] Yu. V. Artemov, V. M. Genkin, G. I. Leviev *et al.*, Supercond. Sci. Technol. **10**, 590 (1997).
[15] C. D. Jeffries, Q. H. Lam, Y. Kim *et al.*, Phys. Rev. B **37**, 9840 (1988).
[16] Q. H. Lam and C. D. Jeffries, Phys. Rev. B **39**, 4772 (1989).
[17] B. P. Wilfley, H. Suhl, and S. Schults, Phys. Rev. B **30**, 2649 (1984).
[18] I. Ciccarello, C. Fazio, M. Guccione *et al.*, Physica C **159**, 769 (1989).
[19] L. Ji, R. H. Sohn, and M. Tinkham, Phys. Rev. B **40**, 936 (1989).
[20] I. F. Voloshin, N. M. Makarov, L. M. Fisher *et al.*, JETP Lett. **51**, 255 (1990).
[21] G. I. Leviev, R. S. Papikyan, and M. R. Trunin, Zh. Éksp. Teor. Fiz. **99**, 357 (1991) [Sov. Phys. JETP **72**, 201 (1991)].
[22] A. S. Shaulov, Phys. Rev. B **43**, 3760 (1991).
[23] P. Chaddah, S. B. Roy, and S. Kumar, Phys. Rev. B **46**, 11737 (1992).
[24] M. Coffey, Phys. Rev. B **47**, 15298 (1993).
[25] S. Sridhar, Appl. Phys. Lett. **65**, 1055 (1994).
[26] F. Gunea and Yu. Pogorelov, Phys. Rev. B **53**, 6725 (1996).
[27] C. J. van der Beek, V. B. Geshkenbein, and V. M. Vinokur, Phys. Rev. B **48**, 3393 (1993).
[28] G. Blatter, M. Feigel'man, V. Geshkenbein *et al.*, Rev. Mod. Phys. **66**, 1147 (1994).
[29] M. Golosovsky, M. Tsindlekht, and D. Davidov, Semicond. Sci. Technol. **9**, 1 (1996).
[30] A. A. Golubov, M. R. Trunin, S. V. Shulga *et al.*, Physica C **213**, 139 (1993).
[31] N. Bloembergen, *Nonlinear Optics*, Benjamin, New York (1965).

[32] L. A. Vainshtein, *Electromagnetic Waves* [in Russian], Nauka, Moscow (1988).
[33] E. S. Borovitskaya, V. M. Genkin, and G. I. Leviev, Zh. Éksp. Teor. Fiz. **110**, 1081 (1996) [JETP **83**, 597 (1996)].
[34] A. M. Portis, K. W. Blasey, K. A. Muller *et al.*, Europhys. Lett. **5**, 467 (1988).
[35] A. M. Portis, K. W. Blasey, and F. Waldner, Physica C **153–155**, 308 (1988).
[36] C. P. Bean and J. D. Livingston, Phys. Rev. Lett. **12**, 14 (1964).
[37] F. F. Ternovskii and L. N. Shekhata, Zh. Éksp. Teor. Fiz. **62**, 2297 (1972) [Sov. Phys. JETP **35**, 1202 (1972)].
[38] J. R. Clem, J. Appl. Phys. **50**, 3518 (1979).
[39] C. P. Bean, Rev. Mod. Phys. **36**, 31 (1964).
[40] E. Zeldov, A. I. Larkin, V. B. Geshkenbeim *et al.*, Phys. Rev. Lett. **73**, 1428 (1994).
[41] E. S. Borovitskaya, V. M. Genkin, and G. I. Leviev, IEEE Trans. Appl. Supercond. **7**, 1220 (1997).
[42] L. Burlachkov, Phys. Rev. B **47**, 8056 (1993).
[43] M. R. Trunin, Usp. Fiz. Nauk **168**, 931 (1998).

Translation provided by the Russian Editorial office.

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

Toward a theory of nuclear relaxation in dielectric glasses at ultralow temperatures

Yu. Kagan, L. A. Maksimov, and I. Ya. Polishchuk*)

Russian Scientific Center ''Kurchatov Institute''

(Submitted 28 January 1999)
Zh. Éksp. Teor. Fiz. **115**, 2254–2262 (June 1999)

The temperature and frequency dependence of the nuclear relaxation rate in dielectric glasses is investigated. It is shown that at low and ultralow temperatures nuclear relaxation is due to an interaction between the nuclear quadrupole moment and fluctuations of the electric field created by dipole moments of two-level systems. Fluctuations of this field can be associated with the background relaxation or are due only to the dipole–dipole interaction between two-level systems. It is shown that at lower temperatures the second relaxation mechanism begins to dominate. Expressions are obtained for the temperature and frequency of crossover between different nuclear relaxation regimes. The possibility of experimental confirmation of our results is discussed. © *1999 American Institute of Physics.* [S1063-7761(99)02706-7]

1. More than 25 years ago Zeller and Pohl[1] demonstrated experimentally that at temperatures below 1 K dielectric glasses exhibit a universal anomalous behavior of their specific heat and thermal conductivity in comparison with their crystalline counterparts. To explain these anomalies, the authors of Ref. 2 advanced the hypothesis that additional degrees of freedom (in addition to the phonon degrees of freedom) exist in such glasses which can be described phenomenologically within the framework of the model of noninteracting two-level tunneling systems (TLS). In this model an isolated two-level system is described by the standard pesudospin Hamiltonian

$$h=-\Delta S^z-\Delta_0 S^x, \tag{1}$$

where the level detuning Δ and the tunneling transition amplitude Δ_0 are governed by the distribution function

$$P(\Delta,\Delta_0)=\overline{P}/\Delta_0. \tag{2}$$

This model of a dielectric glass has afforded a satisfactory explanation of experimental results at temperatures below 1 K, all the way down to 100 mK.[3,4] Thus, for example, in this temperature range experimental studies have been conducted on the relaxation properties of thermal two-level systems with parameters

$$\Delta\simeq\Delta_0\simeq T, \tag{3}$$

whose concentration is equal to

$$c_T=a^3(\overline{P}T), \tag{4}$$

where a is the interatomic distance. For the rate of longitudinal relaxation of two-level systems (inverse lifetime) the following dependence has been revealed[4]:

$$\tau_{\mathrm{ph}}^{-1}=U_0(T/v)^3, \tag{5}$$

which is associated with single-phonon relaxation. Here

$$U_0=\frac{\gamma^2}{\rho v^2}, \tag{6}$$

and γ is the constant of the deformational interaction between the two-level tunneling system and the phonons, and ρ and v are respectively the density and speed of sound in the glass. Note that for all known dielectric glasses it has been empirically determined that

$$\overline{P}U_0\ll 1. \tag{7}$$

The interaction of two-level tunneling systems with phonons gives rise to an indirect interaction between two-level systems, which can be written as (see, e.g., Ref. 5)

$$\hat{V}=\frac{1}{2}\sum_{ij} U(R_{ij})S_i^z S_j^z, \quad U(R_{ij})=\frac{U_0}{R_{ij}^3}. \tag{8}$$

The first systematic experimental studies indicating the importance of taking this interaction into account were begun in 1976.[3] The dipole–dipole nature of this interaction and the associated spectral diffusion between the two-level systems[6] made it possible to explain the experimentally revealed quadratic temperature dependence of the transverse relaxation rate

$$\tau_2^{-1}\sim T^2. \tag{9}$$

2. For a long time it was assumed that spectral diffusion is the only manifestation of the interaction between two-level systems,[7] and the question has remained open, to what degree the properties of dielectric glasses at ultralow temperatures are due to the dipole–dipole interaction. The possibility that this interaction plays a defining role in the formation of the universal properties of such glasses was indicated in Refs. 8 and 9.

The influence of the dipole–dipole interaction on the properties of dielectric glasses at ultralow temperatures was most consistently investigated in the series of papers by Burin *et al.*[10–12] In these papers it was shown that a new class of collective low-energy excitations arises in dielectric glasses due to the dipole–dipole interaction, which bring about the relaxational properties of these glasses at ultralow temperatures. This class of excitations is associated with transitions in pairs of two-level systems whose parameters are close to thermal (3) and whose excitation energy is $E=\sqrt{\Delta_0^2+\Delta^2}\simeq T$, where the transitions themselves are created by the indirect interaction (8).

Any pair of two-level systems can be found, generally speaking, in one of four states: $(+-)$, $(-+)$, $(++)$, $(--)$. Here the symbols $+$ and $-$ correspond to the excited state and the ground state, respectively. The amplitude of the transition between states belonging to the flip-flop configuration, i.e., between the states $(+-)$ and $(-+)$, is given by[10]

$$\Delta_{0p}(R)\approx\frac{U_0}{R^3}\,\frac{\Delta_0\Delta_0'}{EE'}. \tag{10}$$

As was shown in Refs. 10–12, if

$$U_0/R^3\ll T, \tag{11}$$

then the probability of a transition between the states of the flip-flop configuration is much greater than the probability of a transition from these states to any of the two other states: $(++)$ and $(--)$. Therefore it may be assumed that provided condition (11) is satisfied, transitions in pairs of two-level systems take place only between states of the flip-flop configuration, and the energy of these transitions is $\Delta_p=|E-E'|$. Thus, a pair of thermal two-level tunneling systems for which condition (11) is fulfilled is in essence a two-level system described by the Hamiltonian (1) with parameters Δ_p and Δ_{0p}. An important role in such relaxation processes is played by the so-called resonance pairs, for which

$$\Delta_p\leqslant\Delta_{0p}\approx\Delta_*=T(\overline{P}U_0)^2.$$

The concentration of such resonance pairs is equal to[10]

$$c_*=c_T(\overline{P}U_0)\ll c_T, \tag{12}$$

and the relaxation rate

$$\tau_*^{-1}\approx\Delta_*=T(\overline{P}U_0)^2. \tag{13}$$

Note that the characteristic dimension of the resonance pairs, R_*, is equal in order of magnitude to

$$R_*=ac_*^{-1/3}. \tag{14}$$

It is important to note that because of the interaction with phonons the coherent coupling between two-level systems belonging to the same resonance pair is destroyed. For this reason, as was shown in Ref. 10, the concept of a resonance pair, considered as a carrier of the elementary excitation, loses meaning at temperatures

$$T>T_0=(\overline{P}U_0)\sqrt{v^3/U_0}, \tag{15}$$

for which $\tau_{\mathrm{ph}}^{-1}=\tau_*^{-1}$. It is important that for $T<T_0$

$$\tau_*<\tau_{\mathrm{ph}}.$$

Despite the fact that taking relaxation of resonance pairs into account enables one to understand peculiarities of the behavior of dielectric glasses at ultralow temperatures (e.g., it explains the linear temperature dependence of the internal friction coefficient[10] and the rate of phase memory loss[12]), in some works the fundamental role of the dipole–dipole interaction in the formation of the universal properties of glasses at ultralow temperatures continues to be put in doubt.[13,14] In this regard, in our opinion it is important to identify alternative properties of dielectric glasses that might be determined primarily by relaxation of resonance pairs.

3. A study of relaxation of nuclei having a quadrupole electric moment in dielectric glasses in external magnetic fields at low temperatures has a more-than-two-decades-long history.[15–17] At low temperatures spin-lattice relaxation $(\sim T^7)$ is suppressed against the background of the relaxation associated with the presence of a nuclear quadrupole moment, when the relaxation of nuclear spins is due to their interaction with dipole moments of the two-level systems $(\sim T^3$ or $T)$. In an external electric field, the nuclear spin relaxes in the fluctuating electric field created by the dipole moments of the surrounding two-level systems. Depending on whether the thermal two-level system fluctuates because of phonon relaxation or as a component of a resonance pair, different temperature, temporal, and frequency characteristics of nuclear relaxation will be observed, an experimental study of which will allow us to judge the role of the dipole–dipole interaction in the formation of the dynamic properties of dielectric glasses.

To start with, let us specify the relaxation mechanism of the two-level systems. If we neglect anisotropy, the energy of the quadrupole–dipole interaction between a nucleus and a two-level system located a distance R from it has the form

$$E=A/R^4, \tag{16}$$

where A is the interaction constant. Note that in the case when a two-level system existing as a component of a resonance pair relaxes, we take R to mean the distance from the nucleus to the nearest two-level system of this resonance pair.

We assume that we are considering a nuclear transition between levels separated by the energy interval ω. It is well known that in a time-varying external electric field $E(t)$ the probability of such a transition is proportional to the corresponding Fourier transform of the correlator $\langle E(t)E(0)\rangle$. In our case the time dependence of the field (16) is linked with fluctuations of the dipole moments of the two-level systems. In order of magnitude the characteristic correlation time τ of these fluctuations coincides with one of the times (5) and (13). Therefore the relaxation rate of a nucleus located a distance R from a two-level system is equal to[18]

$$\Gamma_R=\left(\frac{A}{R^4}\right)^2\frac{\tau^{-1}}{\tau^{-2}+\omega^2}=\left(\frac{a}{R}\right)^8\left(\frac{A}{a^4}\right)^2\frac{\tau^{-1}}{\tau^{-2}+\omega^2}=\left(\frac{a}{R}\right)^8\Gamma_a, \tag{17}$$

where

$$\Gamma_a=\left(\frac{A}{a^4}\right)^2\frac{\tau^{-1}}{\tau^{-2}+\omega^2}. \tag{18}$$

If at first the projection of the magnetic moment on some axis is equal to M_0, then because of the interaction with a fluctuating two-level system located a distance R from the nucleus the temporal evolution of this projection is given by

$$M = M_0 \exp(-t\Gamma_R). \tag{19}$$

As a consequence of the short-range nature of the quadrupole-dipole interaction (16), relaxation of the nucleus is completely determined by the two-level system nearest to it. The probability $W(R)$ that the fluctuating two-level system nearest the nucleus is found at the distance R is

$$W(R) = c \exp[-(R/R_c)^3], \quad R_c = ac^{-1/3}, \tag{20}$$

where c coincides with either c_T or c_*. Therefore, averaging expression (19) over the configuration of the two-level system about the nucleus, we obtain

$$\overline{M}(t) = M_0 \sum_R W(R) \exp(-t\Gamma_R). \tag{21}$$

We rewrite relation (21) in the form

$$\overline{M}(t) - M_0 = -M_0 c \sum_R \exp\left[-\left(\frac{R}{R_c}\right)^3\right] \times \left\{1 - \exp\left[-\left(\frac{R_t}{R}\right)^8\right]\right\},$$

$$R_t = a(t\Gamma_a)^{1/8}. \tag{22}$$

Let us consider the behavior of $\overline{M}(t)$ on different time scales. At early times $t \ll \Gamma_a^{-1}$ [see Eq. (18)] nuclei relax that have a fluctuating two-level system at one of their nearest sites. Equation (22) then yields

$$\overline{M}(t) \simeq M_0 \exp(-t\gamma_1), \tag{23}$$

where

$$\gamma_1 = zc\Gamma_a, \tag{24}$$

and z is the number of nearest neighbors. Thanks to the smallness of the low-temperature relaxation the early-time stage is extremely important.

At intermediate times, satisfying the condition

$$a \ll R_t \ll R_c, \tag{25}$$

nuclei relax, for which the nearest fluctuating two-level system is found at $R \simeq R_t$.

Neglecting the first factor in Eq. (22) inside the sum and replacing the sum by an integral, we find

$$\overline{M}(t) \simeq M_0 \exp[-(t\gamma_2)^{3/8}], \tag{26}$$

where

$$\gamma_2 = c^{8/3}\Gamma_a = c^{5/3}\gamma_1 \ll \gamma_1. \tag{27}$$

And, finally, at late times, when

$$R_c \ll R_t, \tag{28}$$

changing over in Eq. (21) from a sum to an integral, we obtain

$$\overline{M}(t) \simeq M_0 c \frac{4\pi}{a^3} \int R^2 dR \exp\left[-\left(\frac{R_t}{R}\right)^8 - \left(\frac{R}{R_c}\right)^3\right]. \tag{29}$$

From this formula it follows that in the considered time interval, thanks to spatial inhomogeneity of the configuration of the subsystem, nuclei relax that have the nearest fluctuating TLS at a distance

$$R = R_t^{8/11} R_c^{3/11} \gg R_c.$$

Estimation of the integral (29) by the method of steepest descent gives

$$\overline{M}(t) \simeq M_0 \exp[-(t\gamma_2)^{3/11}]. \tag{30}$$

Note that, as for the case of intermediate times (26), relaxation at late times does not have a simple exponential character and occurs much more slowly than at early times.

4. Let us turn now to an analysis of the obtained dependences in the two following cases:

1) when nuclear relaxation is caused by single-phonon fluctuations of all the thermal two-level systems;

2) when nuclear relaxation is caused by fluctuations of only thermal two-level systems existing as components of resonance pairs.

Thus, in expressions (24) and (27) in the first case instead of τ and c we must substitute respectively τ_{ph} and c_T [see Eqs. (5) and (4)], and in the second case τ_* and c_* [see Eqs. (12) and (13)]. For the analysis that is to follow, note that in dielectric glasses the numerical values of τ_{ph} and τ_* are such that for the characteristic radio frequencies at which nuclear resonance experiments are usually performed, the condition

$$1 \ll \omega\tau_* \ll \omega\tau_{\text{ph}} \tag{31}$$

is satisfied and the quantity (18) is equal to

$$\Gamma_a = \left(\frac{A}{a^4\omega}\right)^2 \tau^{-1}.$$

Let the characteristic times be such that $t \ll (\Gamma_a^*)^{-1}$. Then, the nuclear relaxation goes according to the exponential law (23), (24). In this case, in the first of the above two cases, when nuclear relaxation is stimulated by fluctuations of thermal two-level systems, we obtain

$$\gamma_1^{\text{ph}} = c_T \Gamma_a^{\text{ph}} = (\overline{P}T)\left(\frac{A}{a^4\omega}\right)^2 \tau_{\text{ph}}^{-1} \sim T^4, \tag{32}$$

while in the second case, when nuclear relaxation is caused by fluctuations of resonance pairs,

$$\gamma_1^* = c_* \Gamma_a^* = (\overline{P}T)(\overline{P}U_0)\left(\frac{A}{a^4\omega}\right)^2 \tau_*^{-1} \sim T^2. \tag{33}$$

Comparison of expressions (32) and (33) shows that $\gamma_1^{\text{ph}} = \gamma_1^*$ at the temperature

$$T' = (\overline{P}U_0)^{3/2} \sqrt{\frac{v^3}{U_0}} = (\overline{P}U_0)^{1/2} T_0. \tag{34}$$

Thus, at temperatures $T > T'$ the nuclei relax on thermal two-level systems, and in the opposite case $T < T'$ the nuclei relax on resonance pairs. Note that T' is the temperature at which crossover takes place in the rate of longitudinal and transverse relaxation of two-level systems.[10,12]

The cases of late times [see Eqs. (26) and (30)] are treated analogously. When

$$t \gg (\Gamma_a^{\mathrm{ph}})^{-1}, \tag{35}$$

we obtain

$$\gamma_2^{\mathrm{ph}} = c_T^{8/3} \Gamma_a^{\mathrm{ph}} = (\bar{P}T)^{8/3} \left(\frac{A}{a^4 \omega} \right)^2 \tau_{\mathrm{ph}}^{-1} \sim T^{17/3}, \tag{36}$$

while

$$\gamma_2^{*} = c_*^{8/3} \Gamma_a^{*} = [(\bar{P}T)(\bar{P}U_0)]^{8/3} \left(\frac{A}{a^4 \omega} \right)^2 \tau_*^{-1} \sim T^{11/3}. \tag{37}$$

Comparing formulas (36) and (37), we find that $\gamma_2^{\mathrm{ph}} = \gamma_2^*$ at the temperature

$$T'' = (\bar{P}U_0)^{5/6} T' < T', \tag{38}$$

and, thus, for temperatures $T > T''$ the nuclei relax on thermal two-level systems, and in the opposite case, at temperatures $T < T''$ the nuclei relax on resonance pairs. Note that in the case (35) crossover in the nuclear relaxation rate occurs at the temperature T'', which is lower than T'. Also note that in the time interval

$$(\Gamma^*)^{-1} < t < (\Gamma^{\mathrm{ph}})^{-1}$$

the relaxation processes γ_1^{ph} and γ_2^* compete.

For completeness, let us consider the hard-to-realize case of low temperatures when

$$\omega \tau_* \ll 1 \ll \omega \tau_{\mathrm{ph}}. \tag{39}$$

The relaxation times grow as the temperature is lowered, and the first of conditions (39) in fact imposes a lower bound on the temperature:

$$T_\omega = \frac{\omega}{(\bar{P}U_0)^2} < T, \tag{40}$$

when nuclear relaxation can still be stimulated by a resonance pair. Thus, for a low enough frequency in the case of early times we have

$$\gamma_1^{**} = c_* \left(\frac{A}{a^4} \right)^2 \frac{1}{\tau_*^{-1}} \sim T^0 \tag{41}$$

and the nuclear relaxation rate is independent of the temperature as long as $\gamma_1^{**} \geq \gamma_1^{\mathrm{ph}}$. If this condition is satisfied for $T = T_\omega$, then it continues to be satisfied as the temperature is raised up to the temperature

$$T'_\omega = \frac{1}{(\bar{P}U_0)} \sqrt{\omega T'} = \sqrt{\frac{\omega T_0}{(\bar{P}U_0)^{3/2}}} = \sqrt{T_\omega T'} \tag{42}$$

[for γ_1^{ph} see expression (32)]. This expression is meaningful as long as

$$T_\omega \leq T', \quad \omega < \omega'_* = T'(\bar{P}U_0)^2 = T_0 (\bar{P}U_0)^{5/2}. \tag{43}$$

At late times we have

$$\gamma_2^{**} = c_*^{8/3} \left(\frac{A}{a^4} \right)^2 \frac{1}{\tau_*^{-1}} \sim T^{5/3}.$$

Recall that the resonance pairs control relaxation at late times if $\gamma_2^{**} \geq \gamma_2^{\mathrm{ph}}$. Estimates (42) and (43) in this case are replaced by

$$T''_\omega = T'_\omega (\bar{P}U_0)^{5/12} = (\bar{P}U_0)^{5/12} \sqrt{T_\omega T'},$$

$$T_\omega \leq (\bar{P}U_0)^{5/24} T',$$

$$\omega < \omega''_* = T'(\bar{P}U_0)^{2+5/24} = T_0 (\bar{P}U_0)^{2.7}.$$

5. For the most widely studied dielectric glasses, such as SiO_2, the characteristic temperature region in which crossover of the relaxation mechanism can be expected to occur is $T < 100$ mK. For such temperatures the above numerical estimates show that the nuclear relaxation rate is on the order of $10^{-6}\,\mathrm{s}^{-1}$.

More suitable objects in which crossover of the nuclear relaxation mechanism can be observed are the mixed crystals $KBr_{1-x}KCN_x$—compounds which have recently been widely studied as systems which quite adequately model the properties of dielectric glasses at low temperatures. Estimates show that in such systems the threshold temperature T_0 [see Eq. (15)] at which the coherent coupling between the two-level systems of a resonance pair has still not been destroyed by phonons exceeds 1 K. For such temperatures in the indicated mixed crystals the nuclear relaxation rate can reach values of $10^{-4} - 10^{-3}\,\mathrm{s}^{-1}$. Note that nuclear relaxation rates of such orders of magnitude have been experimentally recorded in amorphous As_2S_3 (Ref. 19). Note that for this compound at temperatures near 10 K a power-law dependence T^α was discovered for the nuclear relaxation rate, where it was established that $1 \leq \alpha \leq 2$. This is in clear contradiction with the assumption of a phonon mechanism of nuclear relaxation (32), but may be evidence in favor of the mechanism of non-phonon nuclear relaxation (33) proposed in the present paper. However, the existence of coherent coupling between the two-level systems in this temperature range cannot be taken as an undisputed fact and requires separate study.

This work was carried out with the support of the Russian Fund for Fundamental Research (Grant No. 98-02-16729), the Program of the Russian Ministry of Education in the field of fundamental natural science (Grant No. 97-0-14.0-80), and INTAS.

*)E-mail: iyp@lkv.kiae.su

---

[1] R. C. Zeller and R. O. Pohl, Phys. Rev. B **4**, 2029 (1971).
[2] P. W. Anderson, B. I. Halperin, and C. M. Varma, Philos. Mag. **25**, 1 (1972); W. A. Phillips, J. Low Temp. Phys. **7**, 351 (1972).
[3] S. Hunklinger and A. K. Raychaudhari, Prog. Inorg. Chem. **9**, 267 (1986).
[4] W. A. Phillips, Rep. Prog. Phys. **50**, 1657 (1987).
[5] J. Joffrin and A. Levelut, J. Phys. (Paris) **36**, 811 (1975).
[6] J. L. Black and B. I. Halperin, Phys. Rev. B **16**, 2819 (1968).
[7] D. Osheroff, S. Rogge, and D. Natelson, Czech. J. Phys. **46**, Suppl. S6, 3295 (1996).
[8] M. V. Klein, B. Fischer, A. C. Anderson, and P. J. Anthony, Phys. Rev. B **34**, 5887 (1978).

[9] C. C. Yu, Phys. Rev. B **32**, 4220 (1988).
[10] A. L. Burin and Yu. Kagan, Zh. Éksp. Teor. Fiz. **106**, 633 (1994) [JETP **79**, 347 (1994)].
[11] A. L. Burin, L. A. Maksimov, I. Ya. Polishchuk, JETP Lett. **49**, 784 (1989).
[12] A. L. Burin, Yu. Kagan, L. A. Maksimov, and I. Ya. Polishchuk, Phys. Rev. Lett. **80**, 2945 (1998).
[13] B. I. White, Jr. and R. O. Pohl, Phys. Rev. Lett. **75**, 4437 (1995).
[14] S. N. Coppersmith, Phys. Rev. Lett. **67**, 2315 (1991).
[15] M. Rubinstein and P. C. Taylor, Phys. Rev. B **9**, 4258 (1974).
[16] O. Kanert, J. Steinert, H. Jain, and K. J. Ngai, J. Non-Cryst. Solids **131**, 1001 (1991).
[17] F. Devreux and L. Malier, Phys. Rev. B **51**, 11344 (1995).
[18] C. P. Slikter, *Principles of Magnetic Resonance*, 3rd ed. (Springer-Verlag, Berlin, 1990).
[19] J. Szeftel and H. Alloul, J. Non-Cryst. Solids **29**, 253 (1978).

Translated by Paul F. Schippnick

JOURNAL OF EXPERIMENTAL AND THEORETICAL PHYSICS VOLUME 88, NUMBER 6 JUNE 1999

ERRATA

Erratum: Quasi-collinear and partially degenerate four-wave mixing: an alternative explanation of the phase-conjugation property of backward stimulated scattering [JETP 88, 235–245 (February 1999)]

Dun Liu

Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109, USA

Guang S. He

Photonics Research Laboratory, State University of New York at Buffalo, New York 14260-3000, USA

[S1063-7761(99)02806-1]

Figures 5 and 6 should read as follows:

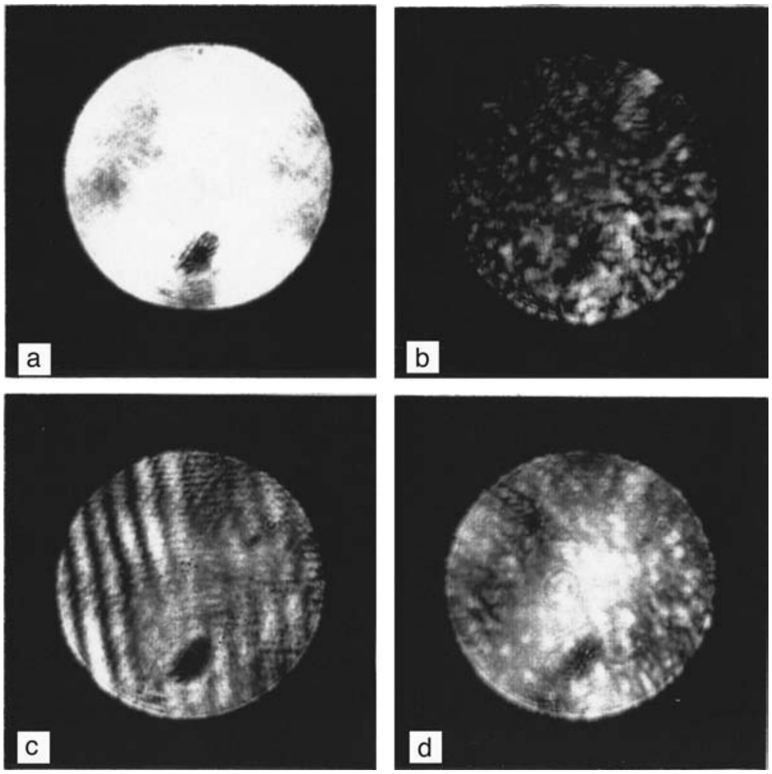

FIG. 5. Photographs of (a) the pump beam II, (b) the pump beam I after passing through an aberration plate, (c) the interference pattern of the two pump beams with no aberration plate, and (d) the interference pattern of the two pump beams with aberration plate in position A shown in Fig. 4.

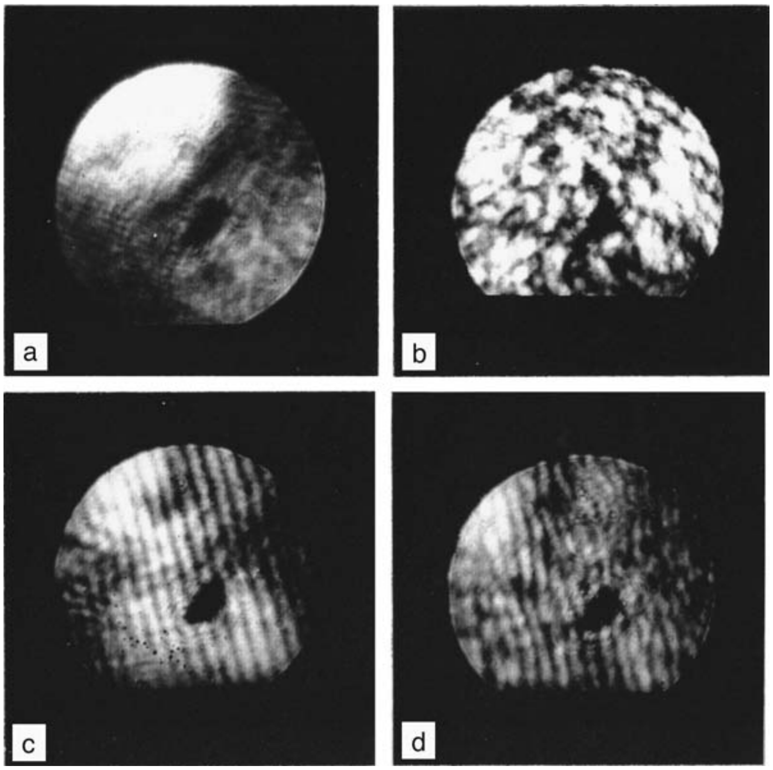

FIG. 6. Photographs of (a) the BSBS beam excited by the pump beam II, (b) the BSBS beam excited by the pump beam I passing through an aberration plate, (c) the interference pattern of the two BSBS beams with no aberration plate, and (d) the interference pattern of the two BSBS beams with an aberration plate in position B shown in Fig. 4.